

Detection of false data injection attacks in smart grid cyber-physical systems

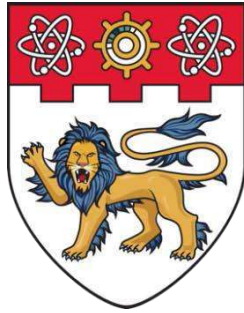
Li, Beibei

2019

Li, B. (2019). Detection of false data injection attacks in smart grid cyber-physical systems. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/88855>

<https://doi.org/10.32657/10220/47640>



**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**DETECTION OF FALSE DATA INJECTION ATTACKS IN
SMART GRID CYBER-PHYSICAL SYSTEMS**

LI BEIBEI

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2019

Detection of False Data Injection Attacks in Smart Grid Cyber-Physical Systems

Li Beibei

School of Electrical & Electronic Engineering
Nanyang Technological University

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

February 2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

07/02/2019

Li Beibei

.....
Date

.....
Li Beibei

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

07/02/2019

.....
Date



.....
Xiao Gaoxi

Authorship Attribution Statement

This thesis contains material from three papers published in and one paper submitted to the following peer-reviewed journals where I was the first author.

Chapter 3 has been published as B. Li, R. Lu, K.-K. R. Choo, W. Wang, S. Luo. "On Reliability Analysis of Smart Grids under Topology Attacks: A Stochastic Petri Net Approach". *ACM Transactions on Cyber-Physical Systems*, vol. 3, no. 1: 10, Jan. 2019. DOI: 10.1145/3127021.

The contributions of the co-authors are as follows:

- Assist. Prof. Lu provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Assist. Prof. Lu, Assoc. Prof. Choo, Dr. Wang and Dr. Luo.
- All the design, testing, simulation, and data collection were conducted by me at the School of Electrical and Electronic Engineering. I co-analyzed the data with Dr. Wang and Dr. Luo.

Chapter 4 has been published as B. Li, R. Lu, W. Wang, K.-K. R. Choo. "Distributed Host-based Collaborative Detection for False Data Injection Attacks in Smart Grid Cyber-Physical System". *Journal of Parallel and Distributed Computing*, vol. 103, pp. 32 - 41, May 2017. DOI: 10.1016/j.jpdc.2016.12.012.

The contributions of the co-authors are as follows:

- Assist. Prof. Lu provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Assist. Prof. Lu, Assoc. Prof. Choo, and Dr. Wang.
- I co-designed the study with Assist. Prof. Lu.
- All the testing, simulation, data collection, and data analysis were conducted by me at the School of Electrical and Electronic Engineering.

Chapter 5 has been published as B. Li, R. Lu, W. Wang, K.-K. R. Choo. "DDOA: A Dirichlet-based Detection Scheme for Opportunistic Attacks in Smart Grid Cyber-Physical System". IEEE Transactions on Information Forensics and Security, vol. 11, no. 11, pp. 2415 - 2425, Nov. 2016. DOI: 10.1109/TIFS.2016.2576898.

The contributions of the co-authors are as follows:


- Assist. Prof. Lu provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Assist. Prof. Lu, Assoc. Prof. Choo, and Dr. Wang.
- I co-designed the study with Assist. Prof. Lu.
- All the testing, simulation, data collection, and data analysis were conducted by me at the School of Electrical and Electronic Engineering.

Chapter 6 is submitted as B. Li, G. Xiao, R. Lu, R. Deng, H. Bao. "On Feasibility and Limitations of Detecting False Data Injection Attacks on Smart Grids Using D-FACTS Devices". Submitted to IEEE Transactions on Industrial Informatics.

The contributions of the co-authors are as follows:

- Assoc. Prof. Xiao provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Assoc. Prof. Xiao, Assist. Prof. Lu, Assist. Prof. Deng, and Assoc. Prof. Bao.
- I co-designed the study with Assoc. Prof. Xiao and Assist. Prof. Deng.
- All the testing, simulation, data collection, and data analysis were conducted by me at the School of Electrical and Electronic Engineering.

07/02/2019



.....
Date

.....
Li Beibei

Dedication

This thesis is wholeheartedly dedicated to my beloved parents and respectable supervisors.

Acknowledgements

There have been many people who have walked alongside me during my Ph.D. journey. They have guided, supported, and accompanied me. They placed opportunities in front of me and showed me the doors that might be useful to open. I would like to, hereby, thank each and every one of them sincerely.

First and foremost, I would like to express my deepest gratitude to my respectable supervisors - Dr. Xiao Gaoxi at Nanyang Technological University (NTU), Singapore, and Dr. Lu Rongxing at University of New Brunswick (UNB), Canada - for their unwavering support and constructive guidance throughout this thesis. They, upon whose shoulders I stand, explored and paved the path before me. Without them, this thesis would simply not have been possible. Such academic rigour as may be found in this thesis is largely due to Dr. Xiao's refusal to let me get away with things, while his unerring sense of when and how to intervene has taught me not only as a good researcher but also a potential good tutor. He is always willing to take time to listen, and usually provide insightful questions and comments as well as clear instructions as feedback. Dr. Lu is a renowned expert in cyber security domain, whose passion for doing research and teaching has set a new standard for everyone involved. His unstinting support and encouragement have driven me to strive to excellence. Having also a friend figure, Dr. Lu is really a nice guy who cares about his students not only on research career but also on daily lives. Many thanks are also due to Dr. Wang Licheng at Beijing University of Posts and Telecommunications (BUPT), China, who started me down this road with selfless support, encouragement, and guidance.

I would especially acknowledge my Thesis Advisory Committee (TAC) members - Dr. Zhang Jie and Dr. Ma Maode at NTU. Thanks for their faith in my ability and continuous support ever since I joined NTU. I hope this research opens up opportunities for us to do research together in the future.

I would extend my heartfelt gratitude to Dr. Ali A. Ghorbani, Dr. Kim-Kwang Raymond Choo, Dr. Bao Haiyong, Dr. Deng Ruilong, Dr. Wang Wei, and Dr. Luo Sheng who contributed to the making of this thesis. Thanks for their constructive criticism, which enabled me to improve my research and writing skills. Particular thanks must also be recorded to Dr. Xu Chang, Dr. Liu Yali, Dr. Kong Qinglei, Dr. Zhao Ming, Dr. Meng Min, Dr. Lin Changlu, Dr. Liu Ximeng, Dr. Li Chen, Dr. Li Lichun, Dr. Hu Hao, Dr. Zhai Chao, Mr. Huang Cheng, Mr. Wang Guoming, Mr. Katuwal Rakesh, Mr. Cheng Shuo, Mr. Hao Changyu, Mr. Zhang Hehong, and Mr. Li Xiang who offered collegial guidance and support over the years.

There are many of my friends to name individually, however special thanks are given to Dr. Yang Rong, Dr. Li Dan, Dr. Yang Ming, Dr. Bi Hui, Dr. Ma Lijia, Dr. Zhang Heng, Dr. Wang Zeng, Dr. Chen Chunyang, Ms. Sun Meng, Ms. Gao Yumeng, Ms. Gong Bo, Ms. Huang Yi, Ms. Wang Zhenzhen, Ms. Huang Rui, Ms. Xin Jian, Ms. Chen Qian, Ms. Wang Yongheng, Ms. Li Yanan, Ms. Chen Shi, Ms. Chen Qiyin, Ms. Zhang Lili, Ms. Ouyang Qingling, Ms. Zhang Ran, Mr. Cheng Yanyu, Mr. Liu Yunxiang, Mr. Pan Zihan, Mr. Wang Yang, Mr. Shao Hongxin, Mr. Li Xiaochen, Mr. Zhang Yunlong, Mr. Guo Zhihong, Mr. Zhang Songze, Mr. Gao Li at NTU, and Dr. Yang Haomiao, Dr. Huang Junlin, Ms. Yang Xue, Ms. Zheng Yandong, Ms. Li Huixia, Ms. Deepigha S. V. Babu, Mr. Guo Wei, Mr. Xu Chenghao, Mr. Hassan Mahdikhani, Mr. Saeed Shafeiee Hasanabadi, Mr. Tao Xi, Mr. Zhang Xichen, Mr. Xiao Hongtao, and Mr. Zhang Yongcan at UNB. Thanks for making me a wonderful and memorable life at both NTU and UNB. In addition, my wholehearted thanks are given to Mrs. Han Yuhong, Ms. Xie Wanlun, Ms. Yu Pan, Mr. Du Wuhang, Mr. Chen Dawen at BUPT for their generous support, understanding, and encouragement given

in many moments of crisis over the years. I cannot list all the names here, but you guys hold a special place in my heart.

Finally, and most importantly, my most heartfelt and forever gratitude goes to my family and girlfriend who have always been a constant source of support and encouragement. Thanks to my parents and elder sister for putting me through the best education possible, and giving me strength to reach for the stars and chase my dreams. I appreciate their sacrifices and unending support, and I wouldn't have been able to get to this stage without them. Special and grateful thanks are given to my girlfriend Ms. Wang Xiaotong for her continuous loving support. Without her companion, this road would have been a lonely place.

Abstract

Building an efficient, green, and multifunctional smart grid cyber-physical system (CPS) while maintaining high reliability and security is an extremely challenging task, particularly in the ever-evolving cyber threat landscape. This challenge is also compounded by the increasing pervasiveness of information and communications technologies across the power infrastructure, as well as the growing availability of advanced hacking tools in the hacker community. One of the most critical security threats in smart grid CPSs lies in the high-profile false data injection (FDI) attacks, where attackers attempt to inject either fabricated measurement data to mislead power grid state estimation & bad data detection or tampered command data to misguide power management & control. Accordingly, FDI attacks can be subdivided into false measurement data injection (FmDI) attacks and false command data injection (FcDI) attacks, respectively.

Detection techniques for FDI attacks have been a significant research focus for smart grid CPSs to withstand these security threats and further protect the power infrastructure. However, conventional state estimation based bad data detection approaches have been proved vulnerable to the evolving FDI attacks. To meet this gap, this thesis introduces four creative research works to analyze and detect FDI attacks in smart grid CPSs.

First, a stochastic Petri net based analytical model is developed to evaluate and analyze the system reliability of smart grid CPSs, specifically against topology attacks with system countermeasures (i.e., intrusion detection systems and malfunction recovery techniques). Topology attacks are evolved from FmDI attacks, where attackers initialize FmDI attacks by

tempering with both measurement data and grid topology information. This analytical model is featured by bolstering both transient and steady-state analysis of system reliability.

Second, a distributed host-based collaborative detection scheme is proposed to detect FmDI attacks in smart grid CPSs. It is considered in this work that phasor measurement units (PMUs), deployed to measure the operating status of power grids, can be compromised by FmDI attackers. Trusted host monitors (HMs) are assigned to each PMU to monitor and assess PMUs' behaviors. Neighboring HMs make use of the majority voting algorithm based on a set of predefined normal behavior rules to identify the existence of abnormal measurement data collected by PMUs. In addition, an innovative reputation system with an adaptive reputation updating algorithm is designed to evaluate the overall operating status of PMUs, by which FmDI attacks as well as the attackers can be distinctly observed.

Third, a Dirichlet-based detection scheme for FcDI attacks in hierarchical smart grid CPSs are proposed. In the future hierarchical paradigm of a smart grid CPS, it is considered that the decentralized local agents (LAs) responsible for local management and control can be compromised by FcDI attackers. By issuing fake or biased commands, the attackers anticipate to manipulate the regional electricity prices with the purpose of illicit financial gains. The proposed scheme builds a Dirichlet-based probabilistic model to assess the reputation levels of LAs. This probabilistic model, used in conjunction with a designed adaptive reputation incentive mechanism, enables quick and efficient detection of FcDI attacks as well as the attackers.

Last, we systematically explore the feasibility and limitations of detecting FmDI attacks in smart grid CPSs using distributed flexible AC transmission system (D-FACTS) devices. Recent studies have investigated the possibilities of proactively detecting FmDI attacks on smart grid CPSs by using D-FACTS devices. We term this approach as proactive false data detection (PFDD). In this work, the feasibility of using PFDD to detect FmDI attacks are investigated by considering single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. It is proved that PFDD can detect all these three

types of FmDI attacks targeted on buses or super-buses with degrees larger than 1, as long as the deployment of D-FACTS devices covers branches at least containing a spanning tree of the grid graph. The minimum efforts required for activating D-FACTS devices to detect each type of FmDI attacks are respectively evaluated. In addition, the limitations of this approach are also discussed, and it is strictly proved that PFDD is not able to detect FmDI attacks targeted on buses or super-buses with degrees equalling 1.

Table of Contents

Dedication	xi
Acknowledgements	xiii
Abstract	xvii
List of Figures	xxvii
List of Tables	xxxi
Acronyms	xxxiii
1 Introduction	1
1.1 Background	1
1.1.1 Cyber Security Events Relating to Electric Grids	2
1.1.2 Recent Actions and Investments in Security and Reliability of Electric Grids	3
1.2 Brief Introduction of A Smart Grid CPS	4
1.2.1 Cyber-Physical System	5
1.2.2 Smart Grid CPS	6
1.2.3 The Architecture of SCADA	8
1.2.4 The Architecture of WAMS	10
1.3 Security Requirements, Challenges, and Research Motivations	12

1.3.1	Security Requirements	12
1.3.2	Security Challenges	13
1.3.3	Research Motivations	14
1.4	Research Contributions	15
1.5	Thesis Outline	17
2	Fundamentals and Related Literature	19
2.1	State Estimation and Bad Data Detection	19
2.1.1	State Estimation Formulation	19
2.1.2	Estimated System States and Measurements	20
2.1.3	Bad Data Detection	21
2.2	FmDI Attacks Against State Estimation	22
2.3	State-of-the-Art Literature	23
2.3.1	A Taxonomy of IDSs	24
2.3.2	FDI Attacks Detection	26
2.3.3	Insider Threats Detection	27
3	SPNTA: Reliability Analysis Under Topology Attacks: A Stochastic Petri Net Approach	29
3.1	Introduction	30
3.2	Preliminaries	33
3.2.1	Petri Net Modeling	33
3.2.2	Coordinated Attacks	34
3.3	Models and Design Goals	35
3.3.1	System Model	35
3.3.2	Adversary Model	37
3.3.3	Design Goals	41
3.4	Proposed Analytical Model	41

3.4.1	Construction of the Proposed SPN Model	41
3.4.2	Maximum Spanning Tree Based Unreliability Enabling Scheme	47
3.5	Performance Evaluation	55
3.5.1	Performance Metrics	55
3.5.2	Numerical Results	58
3.6	Summary	69
4	DHCD: Distributed Host-Based Collaborative Detection for FmDI Attacks	71
4.1	Introduction	72
4.2	Models and Design Goals	74
4.2.1	System Model	74
4.2.2	Threat Model	75
4.2.3	Design Goals	76
4.3	Proposed DHCD Method	76
4.3.1	Collaborative FDD	76
4.3.2	Determination of Compromised PMU	81
4.4	Performance Evaluation	87
4.4.1	Efficacy of FDD Algorithm	88
4.4.2	Identification of Compromised PMUs with Our Reputation System	93
4.5	Summary	96
5	DDOA: A Dirichlet-Based Detection Scheme for Opportunistic Attacks	97
5.1	Introduction	97
5.2	Models and Design Goals	101
5.2.1	System Model	102
5.2.2	Threat Model	103
5.2.3	Design Goals	105
5.3	Preliminaries	105

5.3.1	State Estimation	105
5.3.2	Real-Time LMP	106
5.3.3	Dirichlet Distribution	107
5.4	Proposed DDOA Scheme	108
5.4.1	Behavior Rule Specifications	108
5.4.2	Dirichlet-Based Reputation Model	110
5.4.3	Description of DDOA	113
5.5	Performance Evaluation	116
5.5.1	Data Collection in PowerWorld	116
5.5.2	Data Analytics in MATLAB	117
5.6	Summary	122
6	PFDD: On Feasibility and Limitations of Detecting FmDI Attacks Using D-FACTS	125
6.1	Introduction	126
6.2	Overview and Models	128
6.2.1	Overview of D-FACTS Devices	128
6.2.2	System Model	129
6.2.3	Adversary Model	131
6.3	The Feasibility of PFDD	132
6.3.1	The Framework for PFDD Approach and Its Rationale	133
6.3.2	Evaluation of the Minimum Efforts Required for D-FACTS Devices to Detect Effective FmDI Attacks	135
6.3.3	Minimum Deployment Requirements of D-FACTS Devices to Detect FmDI Attacks	144
6.4	Discussions on PFDD Limitations	148
6.4.1	Limitations of Detecting Effective FmDI Attacks Using PFDD	148
6.4.2	Case Study	153

6.5	Conclusions	156
7	Conclusions and Future Work	157
7.1	Summary of Contributions	157
7.2	Future Work	159
	Appendix A Author's Publications	161
A.1	Book Chapters	161
A.2	Journal Papers	161
A.3	Conference Papers	162
	References	165

List of Figures

1.1	A timeline of recently reported significant cyber security events on electric grids	3
1.2	A summary of recent significant actions and investments in security and reliability of electric grids	4
1.3	A common layered architecture of CPSs	5
1.4	The architecture of a smart grid CPS	7
1.5	The architecture of SCADA in electric grids	8
1.6	The architecture of a WAMS	10
2.1	A taxonomy of IDSs	24
3.1	System model: power system state estimation	36
3.2	Adversary model	38
3.3	Analytical SPN model	42
3.4	System state transitions triggered by the second event	44
3.5	System state transitions triggered by the third event	44
3.6	System state transitions triggered by the fourth event	45
3.7	System state transitions triggered by the fifth event	45
3.8	System state transitions triggered by the sixth event	47
3.9	System state transitions triggered by the seventh event	47
3.10	System state transitions triggered by the eighth event	47

3.11	The MxST of the IEEE 14-bus system	50
3.12	The workflow of calculating reliability	57
3.13	The MTTD and MTTF vs. the compromising rate for IEEE 14-bus system ($\lambda_{rlm} = 0.08$)	59
3.14	The MTTD and MTTF vs. the detection rate for IEEE 14-bus system ($\lambda_{rlm} =$ 0.08)	59
3.15	The MTTD and MTTF vs. the compromising rate for various IEEE testing power systems ($\lambda_{dblm} = 0.02$ and $\lambda_{rlm} = 0.08$)	60
3.16	The MTTD and MTTF vs. the detection rate for various IEEE testing power systems ($\lambda_{clm} = 0.001$ and $\lambda_{rlm} = 0.08$)	60
3.17	Steady-state probability distribution of the number of tokens in place P_UBLM under different compromising rates for IEEE 14-bus system ($\lambda_{dblm} = 0.04$ and $\lambda_{rlm} = 0.08$)	63
3.18	Steady-state probability distribution of the number of tokens in place P_UBLM under different detection rates for IEEE 14-bus system ($\lambda_{clm} = 0.01$ and $\lambda_{rlm} = 0.08$)	63
3.19	System reliability of IEEE 14-bus system vs. the compromising rate ($\lambda_{rlm} =$ 0.08)	64
3.20	System reliability of IEEE 14-bus system vs. the detection rate ($\lambda_{rlm} = 0.08$)	64
3.21	System reliability vs. the compromising rate for various IEEE testing power systems ($\lambda_{dblm} = 0.04$ and $\lambda_{rlm} = 0.08$)	65
3.22	System reliability vs. the compromising rate for various IEEE testing power systems ($\lambda_{clm} = 0.02$ and $\lambda_{rlm} = 0.08$)	65
3.23	Steady-state probability distribution of the number of tokens in place P_UBLM under different recovery rates for IEEE 14-bus system	67
3.24	System reliability vs. the recovery rate for IEEE 14-bus system	67
4.1	The hierarchical architecture of WAMS	75

4.2	Comparison of contouring maps describing the distribution of current amplitude on transmission lines: (a) before open circuit and (b) after open circuit on line from Bus 16 to 17 (marked by a red dashed elliptical circle) in IEEE-39 bus system.	77
4.3	The distributed host-based collaborative FDD system	80
4.4	An example of the conjunctive results transmitted between HMs	81
4.5	IEEE 39-bus power system	88
4.6	Four different cases of the distribution of PMUs with inserted false measurement data: single, sparse, random, and dense	90
4.7	The average iterations needed for FDD algorithm vs. different numbers of PMUs with false measurement data	92
4.8	The detection rate of FDD algorithm vs. different numbers of PMUs with false measurement data	92
4.9	The reputation level of a PMU under various values of ω ($T_h = 0.8, D_{th} = 0.6, S_b = 10, \lambda_b^0 = 0.5$)	94
4.10	The reputation level of a PMU under different under various values of D_{th} ($T_h = 0.8, \omega = 0.4, S_b = 10, \lambda_b^0 = 0.5$)	94
4.11	The reputation level of a PMU under various values of λ_b^0 ($T_h = 0.8, \omega = 0.4, D_{th} = 0.6, S_b = 10$)	95
4.12	The reputation level of a PMU under various values of S_b ($T_h = 0.8, \omega = 0.4, D_{th} = 0.6, \lambda_b^0 = 0.5$)	95
5.1	Three-tier hierarchical flocking-based framework for future smart grids . . .	98
5.2	The contouring map of electricity price distribution on IEEE 39-bus power system	104
5.3	IEEE 39-bus power system with example flocking areas	117
5.4	Reputation level vs. different values of ϵ with $P_n = 0.1$	118
5.5	Reputation level vs. different values of P_n with $\epsilon = 0.75$	118

5.6	Reputation level vs. different values of ϵ with daily dynamic P_n	119
5.7	Reputation level with an aggregative attacker with $\epsilon = 0.75$	119
5.8	Reputation level with an inserted temporal system fault with $\epsilon = 0.75$	121
5.9	Detection rate vs. different lengths of observation window with $P_n = 0.1$. .	121
6.1	The system model - DC state estimation in smart grids	130
6.2	The relationship between the minimum $ \Delta b_{25} $ and c_2	141
6.3	The relationship between the minimum $ \Delta b_{25} $ and c_2 under various values of c_5	142
6.4	The relationship between the minimum efforts and the injected voltage phase angle	143
6.5	Detection of <i>effective</i> FmDI attacks by using PFDD approach under vari- ous D-FACTS deployment strategies: (a) no <i>effective</i> FmDI attack when <i>unknown branches</i> contain a spanning tree; (b) <i>effective</i> single-bus (on \mathcal{V}^2) or coordinated multiple-bus (on \mathcal{V}^1) FmDI attacks when <i>unknown branches</i> fail to contain a spanning tree; and (c) <i>effective</i> single-bus (on \mathcal{V}^1), unco- ordinated multiple-bus (on \mathcal{V}^1), or coordinated multiple-bus (on \mathcal{V}^2) FmDI attacks when <i>unknown branches</i> fail to contain a spanning tree - less <i>unknown</i> <i>branches</i> compared to (b)	145
6.6	An illustrative 8-bus power system with D-FACTS deployment covering a spanning tree.	153
6.7	An illustrative 39-bus power system with D-FACTS deployment covering a spanning tree.	155

List of Tables

1.1	The variables that PMUs can measure	11
1.2	Some representatives of cyber attacks	13
2.1	A summary of example IDSs in recent years	26
3.1	Places in the SPN model	43
3.2	Transitions in the SPN model	43
3.3	Weights assigned to each bus	48
3.4	The bus type and total weight assigned in IEEE 14-bus system	52
3.5	Weights assigned for each branch in IEEE 14-bus system (g denotes generator and l denotes load)	53
4.1	Rules specifications for PMUs in stable status	78
4.2	Parameter settings	89
4.3	The detection rate and the average iterations of FDD algorithm with single rule violated false measurement data under four different distribution types. The number of PMUs with false measurement data is 6	89
4.4	The detection rate and the average iterations of FDD algorithm with multiple rules violated false measurement data under four different distribution types. The number of PMUs with false measurement data is 6	90
5.1	Rule specifications	109

Acronyms

CC	Control Center
CPS	Cyber-Physical System
DDoS	Distributed Denial-of-Service
DoS	Denial-of-Service
D-FACTS	Distributed Flexible AC Transmission System
FcDI	False Command Data Injection
FDD	False Data Detection
FDI	False Data Injection
FmDI	False Measurement Data Injection
GPS	Global Positioning System
HM	Host Monitor
HMI	Human Machine Interface
ICS	Industrial Control System
ICT	Information and Communications Technology
IDS	Intrusion Detection System
IoT	Internet-of-Things
IPS	Intrusion Prevention System
LA	Local Agent
LAN	Local Area Network
LMP	Locational Marginal Price

MLE	Maximum Likelihood Estimation
MTTD	Mean Time to Disturbance
MTTF	Mean Time to Failure
MTTM	Mean Time to Malfunction
MxST	Maximum Spanning Tree
PDC	Phasor Data Concentrator
PLC	Programmable Logic Controller
PMU	Phasor Measurement Unit
RF	Radio Frequency
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Networking
SPN	Stochastic Petri Net
WAMS	Wider Area Measurement System
WAN	Wide Area Network

Chapter 1

Introduction

1.1 Background

With energy being a premium resource for economy, society, and national security, ensuring an accurate, reliable, and efficient power generation, transmission, distribution, and consumption is of prime concern in the 21st century [1]. Unfortunately, the massive blackout in northeastern North America in 2003 uncovered the ease with which the electric grids could be taken down. It is, indeed, not the end of the story for the security and reliability breaches of existing electric grids. To build a fully automated, resilient, and self-healing smart grid, a series of advanced technologies including information and communications technologies (ICTs), automation, distributed control, wide area monitoring and control, smart metering, to name a few, are rapidly incorporated into the existing electric grid over the recent years [2, 3]. Due to lack of strong and diligent security measures in place, however, these newly introduced technologies - exposing a great number of access points to the public - have been opening up possibilities for malignant penetrations [4].

1.1.1 Cyber Security Events Relating to Electric Grids

Cyber attacks on electric grids are no longer a theoretical concern. The summer of 2010 stroke the world in an unprecedented way by discovering the world's first digital weapon - Stuxnet [5]. Unlike any other computer virus or worm that came before, Stuxnet escaped the digital realm to wreak physical destructions on the equipment that computers controlled. By infiltrating the Windows computers at the Natanz nuclear plant in Iran, Stuxnet destroyed an estimate number of 984 uranium enriching centrifuges in total [6]. The impacts of this event go beyond the immense damages caused to Iran. A great deal of ideas of copying and re-purposing Stuxnet from the hacking community as well as, correspondingly, research studies focusing on detection and mitigation of such cyber attacks from the academia and industry have quickly emerged thereafter.

Started by Stuxnet, a huge wave of cyber security events on electric grids have been observed since then (see Fig. 1.1 a timeline of these events). It is reported in 2011 that a cyber campaign involving a Trojan Horse based malware, also notorious as BlackEnergy, compromised the industrial control systems (ICSs) of numerous national critical infrastructures in the U.S. [7]. In August 2012, a self-replicating computer virus named Shamoon infected three quarters of Windows-based corporate PCs at Saudi Aramco, one of the world's largest oil companies [8]. A great amount of invaluable data including documents, spreadsheets, emails, files was eventually erased and replaced with an image of a burning American flag. An analogous attack on Saudi Aramco was initiated in August 2017, where a malware called TRITON created operational disruptions towards critical infrastructures in Saudi Arabia [9]. In February 2013, JEA, the seventh-largest community-owned electric utility in the U.S. was hit by a distributed denial-of-service (DDoS) attack, which led to a crash of online and telephone payment systems for a few days [10].

The year 2015 has witnessed the world's most sophisticated and most successful cyber security event in electric grids to date [11]. It was a Saturday night just two days before Christmas in 2015, an orchestrated cyber attack simultaneously hijacked several power

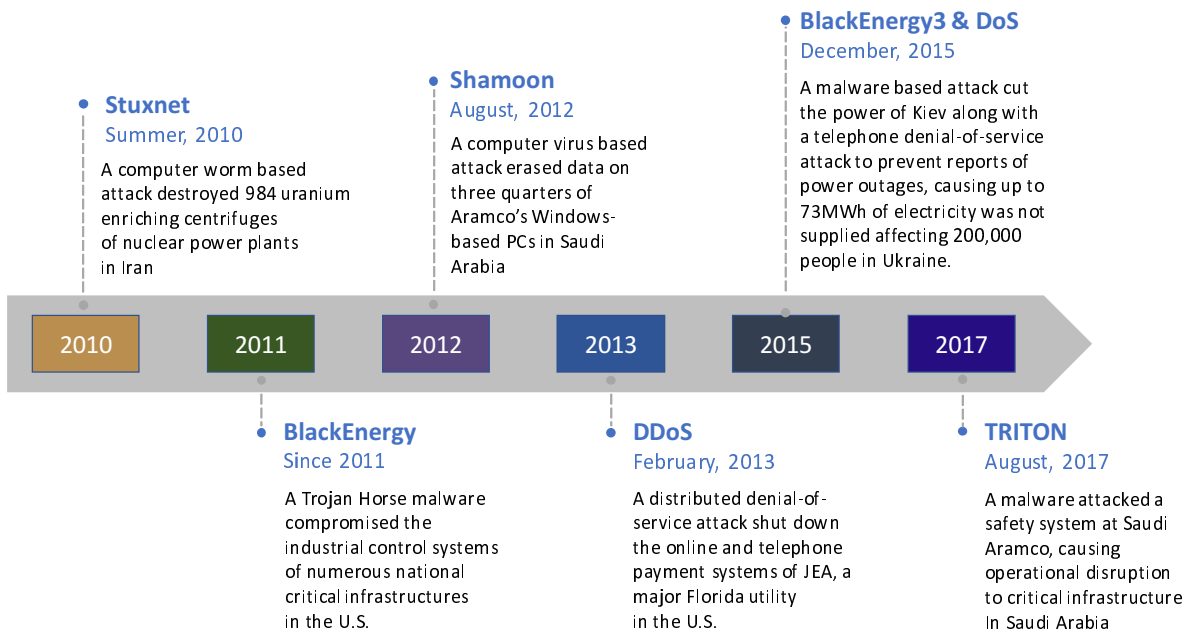


Fig. 1.1 A timeline of recently reported significant cyber security events on electric grids

distribution centers at the Ivano-Frankivsk region of Western Ukraine. Approximately 30 substations were eventually taken offline in this assault, leaving more than 230,000 Ukrainians in the dark for a period of one to six hours. This assault was launched in a well-choreographed dance, where well-trained hackers synchronously switched off a number of substations, disabled IT infrastructures, destroyed files stored on the servers, as well as initiated a telephone denial-of-service (DoS) to deny customers' reports of power blackouts. It was exactly the "brilliant" plan for launching such a real cyberattack that shaped the research landscape of both the power and security community.

1.1.2 Recent Actions and Investments in Security and Reliability of Electric Grids

The frequent and horrible cyber attacks in recent years have emerged as a driving factor to promote the advancements of existing electric grids. It becomes a common sense for nation governors that a secure and reliable power delivery network is of utmost importance

Nation	Actions and Investments
The U.S.	Since 2010, the U.S. has invested more than \$210 million in cybersecurity research, including advancing the resilience of the Nation's energy delivery systems; The American Recovery and Reinvestment Act prompted more than \$4.5 billion investments in grid modernization involving improving the grid reliability and resilience; The Joint United States-Canada Electric Grid Security and Resilience Strategy was issued in 2016 to promote the development of a security and resilience strengthened North American electricity grid.
Canada	An investment of estimated CAD \$25-40 billion by 2020 was announced to refurbish, rebuild, replace the grid infrastructure to ensure a reliable power transmission system; In January, Canada's Minister of Natural Resources announced a \$100-million call for proposals to fund more smart grid systems to fight climate change, create clean jobs and ensure safer power delivery for Canadians; The Joint United States-Canada Electric Grid Security and Resilience Strategy was issued in 2016 to promote the development of a security and resilience strengthened North American electricity grid.
China	A total amount of RMB 286.11 billion was invested in upgrading smart substations and smart meters during the 12-th Five-Year Plan (2011-2015); At least RMB 2 trillion will be spent to improve the reliability of power transmission over the 13-th Five-Year Plan (2016-2020).
The U.K.	Up to £6.1 billion of electricity network investment in Scotland and around £11.6 billion in England and Wales were approved in 2013 to build reliable power delivery networks.

Fig. 1.2 A summary of recent significant actions and investments in security and reliability of electric grids

to support a functioning society. Mindful of this, they have been making action plans and directing investments to reinforce the security and reliability of their electric grids.

Figure 1.2 summaries several recent significant actions and investments in security and reliability of electric grids in the U.S., Canada, China, and the U.K. [12–16]. As we see, each nation has announced enormous investments as well as necessary joint strategies to improve the electric grids' security and reliability.

1.2 Brief Introduction of A Smart Grid CPS

With such a research background in mind, in this section, we brief the concepts and architectures of cyber-physical systems (CPSs), the smart grid CPS, the supervisory control and data acquisition (SCADA) system, as well as the wider area measurement system (WAMS), respectively.

1.2.1 Cyber-Physical System

A CPS is an integrated, hybrid networks of cyber and engineered physical elements. It is co-designed and co-engineered by experts from various domains, including control & automation, computer science, communications, mechanics, etc., to create an adaptive, flexible, situation aware, and predictive hybrid system.

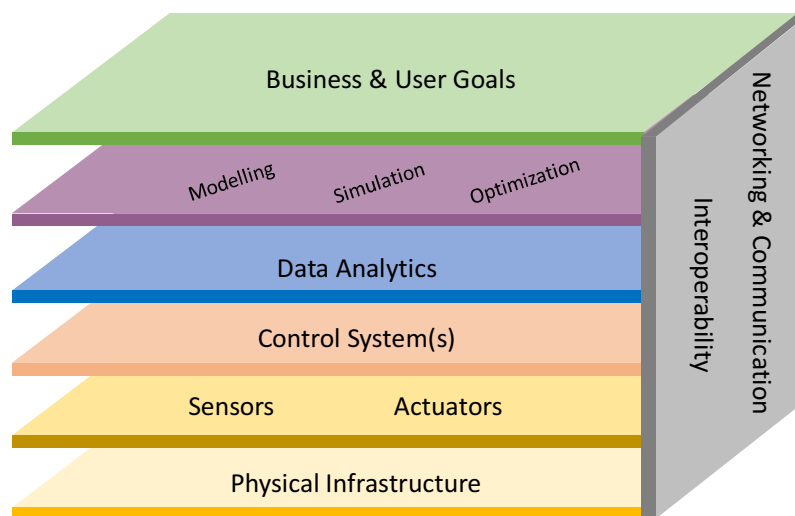


Fig. 1.3 A common layered architecture of CPSs

A common layered architecture of CPSs are presented in Fig. 1.3. In this reference architecture, the bottom layer is the large-scale physical infrastructure. Widespread sensors and actuators are deployed over the physical infrastructure to measure its operating status and execute given control commands towards it. Data collected by sensors will be reported to the control systems, who also issue the control commands to the actuators. On top of the control systems, data analytic techniques will be employed to analyze these reported data to further support various applications such as system modelling, simulation, and optimization. Owners of the CPSs will make use of these invaluable data as well as corresponding applications for business purposes and user goals. Note that the networking & communication technologies running throughout all the layers making its interoperability possible.

1.2.2 Smart Grid CPS

A smart grid CPS, whose physical system is the electric grid, incorporates advanced digital technologies, automation, computer and control to perform a duplex two way communications between the customers and utilities. Also, a smart grid CPS can be regarded as an Internet of things - power generators, distributors, meters, utilities, and customers. It is expected that by employing two way communications, a smart grid CPS can not only enable monitoring and controlling of power delivery in a (near) real-time mode, but also allow customer interactions of electricity usage. The promising benefits of the smart grid CPS include [17]

- improved power reliability and quality
- increased resilience against system faults or natural disasters
- auto-scheduling of power delivery
- predictive maintenance, self-healing, and fast remote repair
- increased capacity and energy efficiency, and reduced carbon emissions
- expanded integration of renewable energy sources, e.g., wind, solar, hydro
- load shedding and lowered electricity tariff
- enhanced customers knowledge of energy usage
- real-time pricing
- remote billing and reduced of manpower costs
- support for smart cities and intelligent transportation systems

Figure 1.4 shows the architecture of a smart grid CPS, with which the above-mentioned promising benefits can be achieved. A smart grid CPS mainly consists of power generation, power transmission, power distribution, power consumption, and control systems.

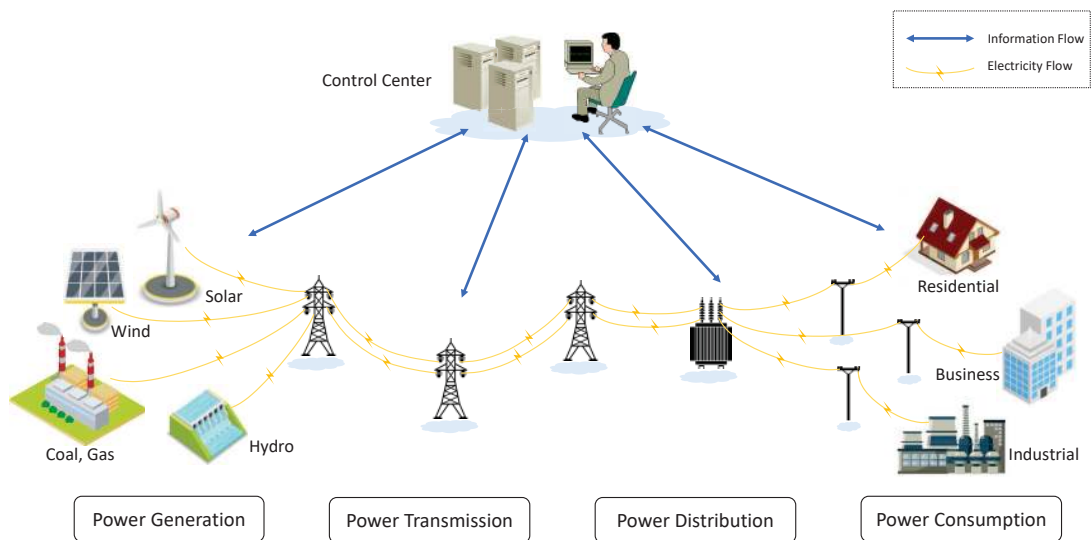


Fig. 1.4 The architecture of a smart grid CPS

- **Power generation:** Energy resources can be classified into renewable (e.g., wind, solar, hydro, biomass, geothermal) and non-renewable resources (e.g., coal, gas, nuclear). Power generation can be classified into centralized generation and distributed generation.
- **Power transmission:** Electricity is transmitted from generating points to substations through power transmission networks, usually at high voltages - 115kV and above - to reduce power loss over long-distance transmission.
- **Power distribution:** Electricity is delivered from the substations to the customer premises through power distribution networks. Transformers are employed to lower the high voltage from transmission networks to medium voltage ranging from 2kV to 35kV, and lower it again to the utilization voltage when approaching to the customer premises.
- **Power consumption:** Electricity consumers include residential houses, business bodies (e.g., schools, hospitals, commercial buildings), and industrial plants. Differential electricity tariffs are usually provided for different purposes of energy usages.

- Control systems: In existing electric grids, there is usually a centralized control center, while there will be more distributed control centers in future smart grid CPSs. The control centers are responsible for power monitoring, management and control to ensure a reliable, secure, and efficient power generation, transmission, distribution, and consumption.

1.2.3 The Architecture of SCADA

SCADA is an industrial computer-based control system employed to gather and analyze the operating status data of the industrial equipment, and to further manage and control the industrial process. It is popular in diverse fields including oil and gas, electric grids, water and waste, agriculture and irrigation, transportation, etc. The architecture of SCADA in electric grids is presented in Fig. 1.5.

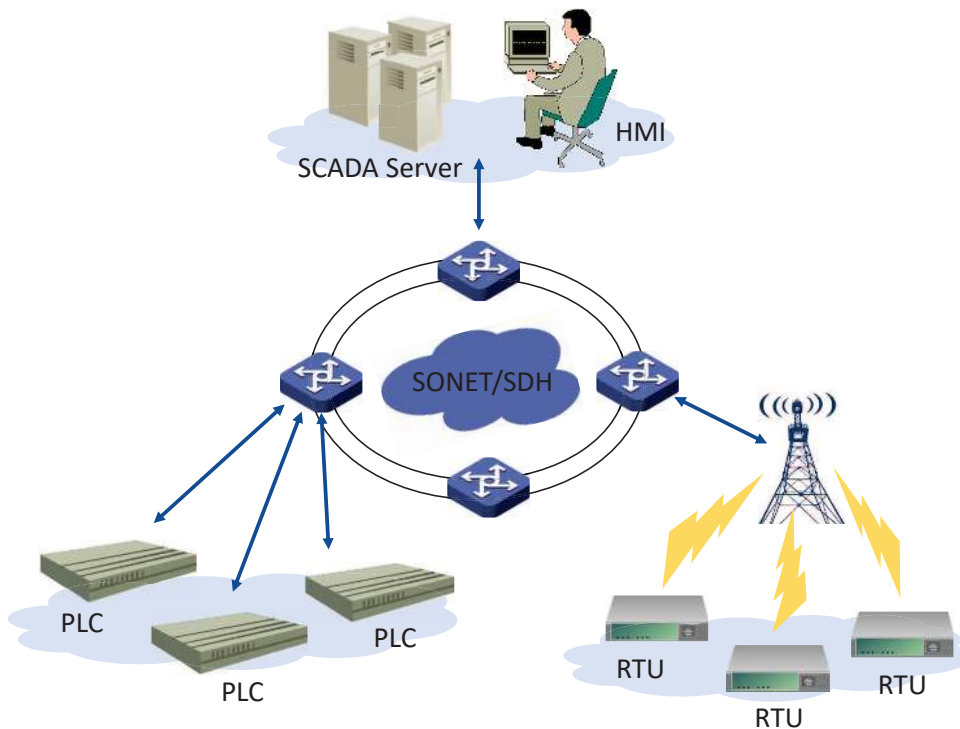


Fig. 1.5 The architecture of SCADA in electric grids

It consists of a SCADA server, a human machine interface (HMI), multiple numbers of remote terminal units (RTUs), programmable logic controllers (PLCs), and communication interfaces. Specifically,

- SCADA server: Serving as a centralized master, the SCADA server monitors and controls the whole system by analyzing the telemetry data reported from the RTUs and generate corresponding feedback orders. These orders will be issued to either RTUs or PLCs.
- HMI: Located in the control room, the HMI presents visualized operating status, scheduled maintenance procedures, logistic information, and diagnostic information of the whole system. In addition, it also provides the interactive interface for a system operator to enable human control and management.
- RTUs: Also termed as remote telemetry units, RTUs are located at the remote substation employed to gather telemetry data from field devices. RTUs also process simple orders from the SCADA server, e.g., orders for controlling the connected physical objects.
- PLCs: As the last-mile controllers, PLCs process the orders from the SCADA server to trigger a set of actions such as turning on/off a line breaker, increasing power generation. PLCs are microcomputer based devices that have advanced data handling, storage, and communication capabilities.
- Communication interfaces: A number of communication interfaces constitute the communication infrastructure across the large-scale SCADA system to deliver message among all entities in a SCADA system. Radio frequency (RF) and directed wired connections are usually used for local communications in a SCADA system, while synchronous optical networking (SONET) and synchronous digital hierarchy (SDH) are frequently used for backbone communications.

1.2.4 The Architecture of WAMS

WAMS, also called WAMCS (wide area measurement and control system) by some researchers, now offers a supplement to existing SCADA system. By incorporating new ICTs, WAMS is able to provide a highly accurate, dynamic, and real-time view of electric grids. WAMS is also featured by phasor synchronization and time stamping of the system operating status data. The architecture of a WAMS is provided in Fig. 1.6. As is shown, a WAMS is comprised of a control center (CC), a set of PMUs, a set of phasor data concentrators (PDCs), global positioning system (GPS), and communication infrastructure - SONET/SDH based backbone and wired local communication networks [18]. Specifically,

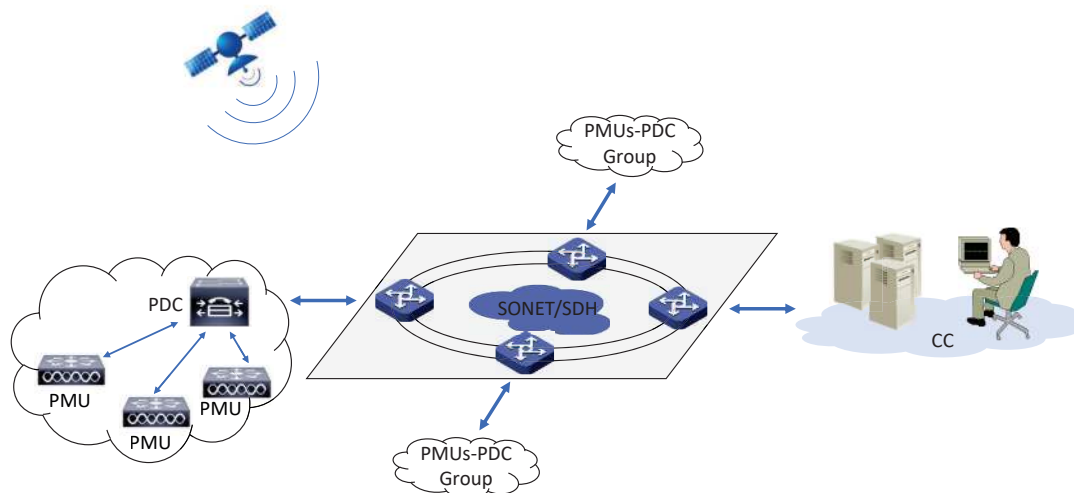


Fig. 1.6 The architecture of a WAMS

- **PMUs:** As major measurement devices, PMUs collect synchronized and time-stamped data of power system operating status such as voltage magnitude and phase, current magnitude and phase, power frequency, and change of frequency. These real-time data are collected at a usual frequency of 50/60Hz and then reported to the regional PDC via local area networks (LANs). The state variables that a PMU can measure are listed in Table 1.1.

Table 1.1 The variables that PMUs can measure

State variable	Description
f	Signal frequency
Δf	Signal frequency variation rate
L_{MW}, L_{Mvar}	Load MW and load Mvar
$P/\delta_P, Q/\delta_Q$	Active and reactive power phasor
$V_A/\Theta_{V_A}, V_B/\Theta_{V_B}, V_C/\Theta_{V_C}$	Phase A, phase B, and phase C voltage phasor
$I_A/\Theta_{I_A}, I_B/\Theta_{I_B}, I_C/\Theta_{I_C}$	Phase A, phase B, and phase C current phasor
$V_1/\Theta_{V_1}, V_2/\Theta_{V_2}, V_0/\Theta_{V_0}$	Positive-, negative-, and zero-sequence voltage phasor
$I_1/\Theta_{I_1}, I_2/\Theta_{I_2}, I_0/\Theta_{I_0}$	Positive-, negative-, and zero-sequence current phasor
On/off	State of breakers

- **PDCs**: The reported measurement data from PMUs are aggregated by regional PDCs and then sent to CC via backbone communication networks.
- **CC**: As a centralized system controller, CC is in charge of management and control of the whole electric grids by analyzing the real-time measurement data, diagnostic data, scheduled information, etc.
- **GPS**: The highly precise synchronous global positioning signals provided by GPS enable the synchronization of measurement data, which also makes it possible for CC to conduct analysis of synchronous data across the whole electric grid.
- **Communication infrastructure**: Serving as courier among all the entities in smart grid CPS, the communication infrastructure plays a significant in delivering the measurement data, commands, and any type of information across the electric grid.

1.3 Security Requirements, Challenges, and Research Motivations

In this section, we brief the security requirements that must be fulfilled to build a secure smart grid CPS, the security challenges that a smart grid CPS is facing with, and the research motivations that stimulates our research studies.

1.3.1 Security Requirements

From a cyber security's perspective, the security requirements can be classified into the following three categories [19]:

- **Integrity:** Protecting against the unauthorized modification or destruction of information, e.g., measurement data from meters or command data from control systems. Modified or destructed information opens the door for mishandling of information, leading to mismanagement of power or malfunction of power applications.
- **Confidentiality:** Protecting privacy and proprietary information, e.g., customer energy consumption data collected by smart meters, by authorizing restrictions on information access, usage, and disclosure.
- **Availability:** Ensuring timely and reliable access to information and services, e.g., command data from control systems and real-time electricity price to customers. Compromised availability may cause delayed power delivery or increased electricity budgets.
- **Authentication:** Protecting against invalid users joining proprietary computer and communication systems, e.g., control systems, to ensure the system users are authentic. Unauthenticated users may deceive or mislead control systems' decisions, or exhaust system resources.

- **Authorization:** Ensuring access to system information and services are legitimate. It is also referred as access control. Unauthenticated user may misuse or jeopardize the system resources.

1.3.2 Security Challenges

The security challenges that the existing electric grids are encountering originate from three domains, including cyber domain, physical domain, and cyber-physical domain.

Cyber Threats

Security threats coming from the cyber domain cause damages or compromise the operating efficiency of electric grids usually by breaching the aforementioned security requirements.

Table 1.2 presents the security breach types with corresponding examples of cyber attacks.

Table 1.2 Some representatives of cyber attacks

Security breach type	Example attacks
Integrity	False data injections (FDIs), man-in-the-middle
Confidentiality	Eavesdropping, theft
Availability	DoS, DDoS
Authentication	Malware, Trojan
Authorization	Malware, Trojan, spoofing

Physical Threats

The physical threats on electric grids include

- cutting fibre optic cables to shut down telecommunication lines
- destroying field equipment, e.g., transformers, cameras, sensors
- destructing documents, installations, and materials

- theft of proprietary information, or equipment
- on-site tempering of electronic devices
- field measurement and investigation

Cyber-Physical Threats

Recent years has been experiencing an increasing trend of cyber-physical threats on electric grids. Cyber-physical threats are usually orchestrated combinations of single cyber threats and physical threats, which are more complicated and threatening than either single threats. For example, cyber-physical attacks can physically compromise and/or destroy some sensors, followed by maliciously reporting falsified sensing data on their behalf.

1.3.3 Research Motivations

The rapid and widespread incorporation of ICTs into the smart grid CPS, as mentioned in the Background section, has been introducing new vulnerabilities and threats even as we are taking actions to prevent, protect against, or minimize the impacts of known threats and hazards. At the same time, the growing dependence of the public, business, government, schools, hospitals, to name a few, on reliable and secure electricity has significantly increased the overall sensitivity to the impacts of any type of power instability. This, including voltage disturbances, momentary power outages, long-term service disruptions, and widespread blackouts with cascading effects, may, regardless of the causes, yield property damages, public health and safety dangers, and financial and life losses. Enhancing the security and resilience of the electric grid against malignant activities is critical to a functioning society.

The challenges of ensuring cybersecurity in a smart grid CPS are diverse in nature, due to the diversity of the components and the contexts where smart grids are deployed. In this thesis, we will mainly address the challenges of protecting data integrity to help build a secure smart grid CPS. Specifically, we will focus on analyzing the system reliability of a smart

grid CPS under data integrity attacks as well as detection techniques of these data integrity attacks.

1.4 Research Contributions

The research focuses in this thesis lie in main topics relating to FDI attacks in smart grid CPSs including modelling and impacts evaluation of FDI attacks, novel detection approaches for FDI attacks - both FmDI and FcDI attacks. Specifically, our main research contributions are summarized as follows:

- In Chapter 3, a stochastic Petri net based analytical model is developed to evaluate and analyze the system reliability of smart grid CPSs, specifically against topology attacks under system countermeasures (i.e., intrusion detection systems and malfunction recovery techniques). Topology attacks are evolved from FDI attacks, where attackers initialize FDI attacks by tempering with both measurement data and grid topology information. This analytical model is featured by bolstering both transient and steady-state analyses of system reliability.
- In Chapter 4, a distributed host-based collaborative detection scheme is proposed to detect FmDI attacks in smart grid CPSs. It is considered in this work that the phasor measurement units (PMUs), deployed to measure the operating states of power grids, can be compromised by FmDI attackers, and the trusted host monitors (HMs) assigned to each PMU are employed to monitor and assess PMUs' behaviors. Neighboring HMs make use of the majority voting algorithm based on a set of predefined normal behavior rules to identify the existence of abnormal measurement data collected by PMUs. In addition, an innovative reputation system with an adaptive reputation updating algorithm is also designed to evaluate the overall operating status of PMUs, by which FmDI attacks as well as the attackers can be distinctly observed.

- In Chapter 5, a Dirichlet-based detection scheme for FcDI attacks in hierarchical smart grid CPSs are proposed. In the future hierarchical paradigm of a smart grid CPS, it is considered that the decentralized local agents (LAs) responsible for local management and control can be compromised by FcDI attackers. By issuing fake or biased commands, the attackers anticipate to manipulate the regional electricity prices with the purpose of illicit financial gains. The proposed scheme builds a Dirichlet-based probabilistic model to assess the reputation levels of LAs. This probabilistic model, used in conjunction with a designed adaptive reputation incentive mechanism, enables quick and efficient detection of FcDI attacks as well as the attackers.
- In Chapter 6, we systematically explore the feasibility and limitations of detecting FmDI attacks in smart grid CPSs using distributed flexible AC transmission system (D-FACTS) devices. Recent studies have investigated the possibilities of proactively detecting FmDI attacks on smart grid CPSs by using D-FACTS devices. We term this approach as proactive false data detection (PFDD). In this work, the feasibility of using PFDD to detect FmDI attacks are investigated by considering single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. It is proved that PFDD can detect all these three types of FmDI attacks targeted on buses or super-buses with degrees larger than 1, as long as the deployment of D-FACTS devices covers branches at least containing a spanning tree of the grid graph. The minimum efforts required for activating D-FACTS devices to detect each type of FmDI attacks are respectively evaluated. In addition, the limitations of this approach are also discussed, and it is strictly proved that PFDD is not able to detect FmDI attacks targeted on buses or super-buses with degrees equalling 1.

1.5 Thesis Outline

The organization of the remainder of this thesis is as follows. Chapter 2 introduces some fundamentals relating to FDI attacks in smart grid CPSs, and provides the state-of-the-art literature reviews. Chapter 3 models the attack strategy of topology attacks using a stochastic Petri net approach and analyzes the system reliability under such attacks. Chapter 4 presents a novel distributed host-based collaborative detection scheme for FmDI attacks. In Chapter 5, a Dirichlet-based detector for FcDI attacks as well as the compromised insiders are introduced, followed by the discussion of feasibility and limitations of detecting FmDI attacks using D-FACTS devices in Chapter 6. Chapter 7 concludes the thesis and briefs some promising research directions for future work.

Chapter 2

Fundamentals and Related Literature

In this chapter, we will introduce some fundamental concepts serving as the building blocks of our research work, which includes the state estimation, bad data detection, as well as FmDI attacks against state estimation. In addition, the state-of-the-art literature reviews on existing intrusion detection systems (IDSs), FDI attacks detection, and insider threats detection will also be provided.

2.1 State Estimation and Bad Data Detection

In this section, we introduce the fundamental concepts of state estimation & bad data detection, one of the most important techniques in electric grids.

2.1.1 State Estimation Formulation

In a power system, state estimation is the heart of control systems to support real-time analysis, contingency analysis, and power management and control. Specifically, state estimation is used to provide estimates of the internal system states given a collection of measurement data. The basic relationship, with an AC power flow model, between the

measurement data and system states are given by [20]

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\eta}, \quad (2.1)$$

where $\mathbf{z} \in \mathbb{R}^{m \times 1}$ is the measurement vector containing information of power generations, power loads, and power flows, $\mathbf{x} \in \mathbb{R}^{n \times 1}$ is the system state vector including bus voltage phase angles, and $\boldsymbol{\eta} \in \mathbb{R}^{m \times 1}$ is the measurement noise vector with zero mean and covariance $\mathbf{W} \in \mathbb{R}^{m \times m}$, a diagonal matrix. Note that m and n are the numbers of measurements and system states, respectively, and $m > n$ indicates that redundant measurements introduced. $\mathbf{h}(\mathbf{x})$ is a non-linear function of \mathbf{x} , which relates the system states to the ideal measurements.

Since state estimation is usually applied over the high-voltage power transmission networks, it is reasonable to approximate AC power flow model to a DC one [21]. In this way, the measurement data and system states are related by

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}, \quad (2.2)$$

where $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the measurement Jacobian matrix, implying the system connection and configuration information. Our research studies in later chapters are all based on DC state estimation. Although AC power flow model is more accurate than DC model, it is computational expensive and too complex to be used in analysis. In contrast, DC power flow model is much faster, more robust, and techno-economic than AC, and it has been widely accepted as a useful simplification of AC model [22–24, 21, 25, 26].

2.1.2 Estimated System States and Measurements

With the relationship shown in Eq. (2.2), the estimated system state vector $\hat{\mathbf{x}}$ using the least squares is given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} (\mathbf{z} - \mathbf{H}\mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{x}). \quad (2.3)$$

The linear DC state estimation has a closed-form solution obtained through a non-iterative procedure by solving Eq. (2.3), which is given by

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{W}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}^{-1} \mathbf{z} \triangleq \mathbf{\Lambda} \mathbf{z}, \quad (2.4)$$

where

$$\mathbf{\Lambda} \triangleq (\mathbf{H}^T \mathbf{W}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}^{-1}. \quad (2.5)$$

Then the estimated measurement data $\hat{\mathbf{z}}$ is given by

$$\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}} = \mathbf{H} \mathbf{\Lambda} \mathbf{z}. \quad (2.6)$$

2.1.3 Bad Data Detection

The existing bad data detection approaches usually use the hypothesis testing, by observing the largest normalized residual (LNR) to detect the bad measurement data. The normalized measurement residual $\mathbf{r} \in \mathbb{R}^{m \times 1}$ is calculated based on the difference between the measurement data \mathbf{z} and the estimated measurement data $\hat{\mathbf{z}}$, i.e.,

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = \mathbf{z} - \mathbf{H} \mathbf{\Lambda} \mathbf{z} = (\mathbf{I} - \mathbf{H} \mathbf{\Lambda}) \mathbf{z}, \quad (2.7)$$

where $\mathbf{I} \in \mathbb{R}^{m \times m}$ is the identity matrix. The hypothesis testing is expressed as

$$\begin{cases} \text{Null hypothesis } \mathbf{H}_0 : \|\bar{\mathbf{r}}\| > \tau \\ \text{Alternative hypothesis } \mathbf{H}_1 : \|\bar{\mathbf{r}}\| \leq \tau, \end{cases} \quad (2.8)$$

where $\bar{\mathbf{r}} = \sqrt{\mathbf{W}^{-1}} \mathbf{r}$ is the normalized measurement residual vector [27]. This testing is to compare the Frobenius norm of the normalized measurement residual $\|\bar{\mathbf{r}}\|$ with a predefined threshold τ . Specifically, if $\|\bar{\mathbf{r}}\| > \tau$, the null hypothesis is accepted indicating the existence of anomalous residuals; hence, bad measurement data presents in \mathbf{z} . Otherwise

(i.e. $\|\bar{\mathbf{r}}\| \leq \tau$), the null hypothesis is rejected, which implies no bad measurement data. The value of τ can be determined by a chi-squared test with a significance level of α , i.e., $\tau = \sqrt{\chi_{m-n, 1-\alpha}^2}$, because $\|\bar{\mathbf{r}}\|^2 = \|\sqrt{\mathbf{W}^{-1}}\mathbf{r}\|^2 = \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}})\|^2$ follows a chi-square distribution χ_{m-n}^2 , where $m - n$ is the degree of freedom [27].

2.2 FmDI Attacks Against State Estimation

In 2011, Liu *et al.* demonstrated that a set of smart attackers can initiate FmDI attacks in electric grids against the existing state estimation & bad data detection technique, as long as they can compromise some meter devices and have some knowledge of electric grid connections and configurations [28]. Specifically, to construct an FmDI attack, the attacker needs to design an attack vector $\mathbf{a} \in \mathbb{R}^{m \times 1}$, and fabricate a malicious measurement vector $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$. If there exists a vector $\mathbf{c} \in \mathbb{R}^{n \times 1}$ that can satisfy $\mathbf{a} = \mathbf{H}\mathbf{c}$, a successful FmDI is constructed and the original estimated system state vector $\hat{\mathbf{x}}$ is injected with an offset by $\hat{\mathbf{x}}_a = \hat{\mathbf{x}} + \mathbf{c}$. This is because that with such false data being injected, the estimated system states vector $\hat{\mathbf{x}}_a$ with reference to Eq. (2.4) is given by

$$\hat{\mathbf{x}}_a = \Lambda \mathbf{z}_a = \Lambda(\mathbf{z} + \mathbf{a}) = \mathbf{x} + \Lambda \mathbf{H} \mathbf{c} = \mathbf{x} + \mathbf{c}, \quad (2.9)$$

where $\Lambda \mathbf{H} = \mathbf{I}$. The physical meaning of \mathbf{c} is the injected offset on the system states (i.e., voltage phase angles in DC power flow model). Then, the Frobenius norm of the normalized measurement residual with false data injected $\|\bar{\mathbf{r}}_a\|$ is given by [28]

$$\begin{aligned} \|\bar{\mathbf{r}}_a\| &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}_a - \mathbf{H}\hat{\mathbf{x}}_a)\| \\ &= \|\sqrt{\mathbf{W}^{-1}}[\mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{x}} + \mathbf{c})]\| \\ &= \|\sqrt{\mathbf{W}^{-1}}[\mathbf{z} - \mathbf{H}\hat{\mathbf{x}} + (\mathbf{a} - \mathbf{H}\mathbf{c})]\| \\ &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}})\| \leq \tau. \end{aligned} \quad (2.10)$$

As we see, in this case, no anomaly can be observed by the existing bad data detection approach, which indicates a success of an FmDI attack.

Note that to construct successful FmDI attacks, the attackers must have the valuable knowledge of the targeted power grid, including branch connection information, system configuration data, as well as the current system operating status. Various channels can be exploited by FmDI attackers to illegally obtain these information, including

- Cyber channels: eavesdropping, intrusion into the control center, insider theft or accidental leaks, and malicious disclosure by disgruntled employees, etc.
- Physical channels: field measurement/investigation acts with specialized tools in areas with insufficient protection, and physical tampering with the hardware components of field devices.
- Cyber & physical channels: coordinated cyber intrusions and physical measurement/investigation acts.

Attackers may have various attack capabilities and, therefore, various knowledge levels of the valuable information of power grids. For example, some of them may have strong attack capabilities and remotely compromise the IED devices of interest through cyber intrusions in a short time, while some others may need a long time to gradually accumulate that information by persistent eavesdropping.

2.3 State-of-the-Art Literature

In this section, we will review the state-of-the-art literature in terms of IDSs, FDI attacks detection techniques, as well as insider threat detection approaches.

2.3.1 A Taxonomy of IDSs

IDSs are one of the primary tools for the protection of computer networks, and they identify and respond to intrusion activities - entities attempting to violate security policies in place - by monitoring and analyzing system behaviors. A typical IDS is usually composed of sensors, analysis engine, and an alerting system. Sensors, deployed at different network places or hosts, are used for collecting network traffic data or host behaviors. The analysis engine is responsible for data analysis with given security models, policies, or signatures. The alerting system reports to the system administrator(s) provided that an intrusion activity is identified by the analysis engine.

Generally, IDSs can be classified as host-based and network-based in terms of the audit sources; signature-based, anomaly-based, specification-based in terms of the detection methods; and passive-based and active-based in terms of the reactions [29]. A taxonomy of IDSs are provided in Fig 2.1. Specifically,

- **Host-based:** this type of IDSs are deployed at specific hosts, e.g., sensors, substations, PCs, routers, to monitor the behaviors of such hosts malicious activities.
- **Network-based:** this type of IDSs usually connect two or more network segments to monitor the traffic over communication links in order to detect malicious intrusions.

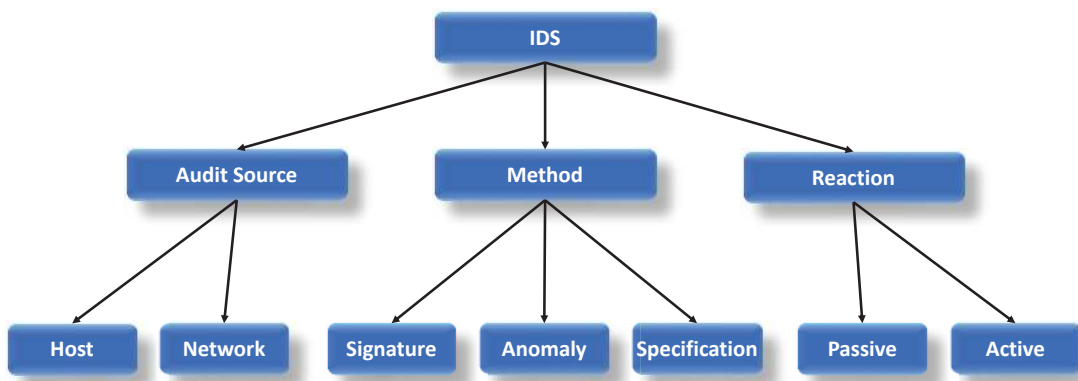


Fig. 2.1 A taxonomy of IDSs

Next, when it comes to detection methods,

- **Signature-based:** this type of IDSs detect attacks when the system or network behaviors match an attack signature/pattern stored in a database. Signature-based IDSs are effective in detecting known-signature/pattern attacks, because a black list of known threats are stored in the database for comparison. However, this approach works ineffectively in detecting new threats or variants of known threats.
- **Anomaly-based:** this types of IDSs detect attacks when the monitored behaviors deviate from normal behaviors. Anomaly-based IDSs are effectively in detecting new threats, because a white list of system or network behaviors are stored in the database for comparison. However, this approach usually suffers from high false positives, because any difference from the given normal behaviors are considered as being anomalous.
- **Specification-based:** this type of IDSs detect attacks when a deviation from pre-defined rule specifications are observed. Unlike anomaly-based approach, in specification-based approach, specifications are defined by domain experts manually according to a set of laws, thresholds, balance requirements, etc. Specification-based approach usually has lower false positives and false negatives, because manually defined specifications can avoid known cases of false positives and false negatives. However, this approach needs to define rule specifications manually to adapt different environments, which may be time-consuming and error-prone to a certain extent.

After review of state-of-the-art literature, we summarize the recent studies relating to IDSs in terms of the detection method into Table 2.1. Specifically, Bao *et al.* in 2016 proposed a specification-based IDS to detect insider threats in smart grid CPS [21]. Thanigaivelan *et al.* in 2016 presented a distributed anomaly-based IDS for detecting routing attacks for Internet-of-Things (IoT) [30]. Pan *et al.* in 2015 developed a hybrid of signature- and specification-based IDS against multiple conventional cyber attacks as well as system

Table 2.1 A summary of example IDSs in recent years

Reference	Security threat	Detection method
Bao <i>et al.</i> [21]	Insider threat	Specification-based
Thanigaivelan <i>et al.</i> [30]	Routing attack	Anomaly-based
Pan <i>et al.</i> [31]	Multiple conventional attacks and system disturbance	Hybrid signature- and specification-based
Faisal <i>et al.</i> [32]	Multiple conventional attacks	Anomaly-based
Lee <i>et al.</i> [33]	DoS	Anomaly-based
Oh <i>et al.</i> [34]	Multiple conventional attacks	Signature-based

disturbances [31]. Faisal *et al.* in 2015 proposed an anomaly-based IDS against multiple conventional attacks for advanced metering architecture in smart grid [32]. Lee *et al.* in 2014 developed an anomaly-based IDS against DoS attacks in 6LoWPAN (IPv6 over Low-Power Wireless Personal Area Networks) [33]. Oh *et al.* in 2014 proposed a signature-based IDS against multiple conventional attacks for the IoT [34].

2.3.2 FDI Attacks Detection

In the following two subsections, we will particularly introduce the state-of-the-art literature relating to FDI attacks detection and insider threat detection - the major focuses in this thesis. The FDI (false data injection) attack is also referred to as or similar to data deception attack, data integrity attack in the existing literature. FDI attacks are crucial security threats to smart grid CPSs, where the attackers attempt to inject false measurement data through compromised meter devices to blind and mislead the control centers. The success of an FDI attack may cause system disturbances, power overloading, power outages, and even system disruptions.

Conventional bad data detection approaches are highly dependent on the power system state estimation. Unfortunately, Liu *et al.* in 2011 demonstrated that new security threats can circumvent the traditional state estimation and bad data detection approach, as long as the

attacker have the knowledge of power system connection and configuration information and a set of compromised meter devices [28]. As a result, bad data detection approaches based on state estimation may no longer be efficient or effective anymore. This type of FDI attacks is also referred to as false measurement data injection (FmDI) attacks. In terms of the data content, another type of FDI attacks is false command data injection (FcDI) attacks. Since FcDI attacks are usually initiated by compromised insiders, we will further review it in the next subsection. In this subsection, we mainly introduce the recent studies focusing on FmDI attacks detection. Yang *et al.* in 2014 studied the optimal FmDI attack strategy to cause maximum damage by identifying the optimal meter set, and developed the spatial-based and temporal-based schemes to detect FmDI attacks [35]. Huang *et al.* in 2013 reviewed the FmDI attacks as well as the defense solutions in smart grid [36]. Esmalifalak *et al.* in 2012 investigated the effect of stealthy FmDI attacks on network congestion in market based power system [37]. Also, Xie *et al.* in 2010 introduced the FmDI attacks in deregulated electricity markets and how such attacks can lead to changes of the locational marginal price and obtain illicit profits [38]. Huang *et al.* in 2011 proposed an adaptive CUSUM algorithm to defend FmDI attacks on smart grid network [39]. Esmalifalak *et al.* in 2011 used independent component analysis to detect stealth FmDI attacks where the attackers are without prior knowledge of the power grid topology [40]. In 2015, Li *et al.* proposed a quickest sequential detector based on the generalized likelihood ratio to detect FmDI attacks with various attack strategies [41]. Also in 2015, Chaojun *et al.* introduced a new detection method against FmDI attacks in smart grid by tracking the dynamics, indicated by the Kullback-Leibler distance, of measurement variations [42].

2.3.3 Insider Threats Detection

Cyber security is vital to the success of a smart grid CPS. The major security threats are coming from the within, as opposed to outside forces. Insider threat detection are significant part of cyber threat mitigation techniques. Insider threats are more challenging compared to

outsider threats, because insiders are usually empowered with legal access and privileges. Each insider has a system role associated with his/her account, such as system administrator, advanced user, normal user. Various levels of access and privileges are provided for different roles. The motivations of insider threats include

- compromised insiders
- misoperations
- emotional-driven (e.g., anger, stress, hostility)
- profit-driven

Motivated insiders either intentionally or unintentionally may cause damages to the system, delay or compromise the services, or stealing or leaking intellectual information. Particularly, FcDI attack is a type of specific insider threats, where, e.g., in a smart grid CPS, the attackers attempts to issue fake commands to the system actuators such as generators, breakers, substations in purpose of causing power outages, overloading, system disturbances, or undesired changes of electricity prices. A myriad of studies focusing on insider threat detection have been observed over the years. Chen *et al.* in 2012 introduced an unsupervised learning based community anomaly detection system against anomalous insiders in collaborative information systems [43]. Ambre and Shekokar in 2015 proposed a probabilistic approach to detect insider threats by using log analysis and event correlation [44]. Legg *et al.* described an automated insider threat detection system by using user tree-structure profiling approach [45]. In 2012, Brdiczka *et al.* proposed a proactive insider threat detection system based on graph learning and psychological context approach [46]. In 2015, Mayhew *et al.* designed an enhanced behavior-based access control technique by integrating machine learning techniques against insider threat detection in big data analytics [47]. Ring *et al.* proposed a new toolset for anomaly-based IDS, particularly for insider threat detection [48].

Chapter 3

SPNTA: Reliability Analysis Under

Topology Attacks: A Stochastic Petri Net

Approach

Building an efficient, smart, and multifunctional power grid while maintaining high reliability and security is an extremely challenging task, particularly in the ever-evolving cyber threat landscape. The challenge is also compounded by the increasing complexity of power grids in both cyber and physical domains. In this chapter, we develop a stochastic Petri net based analytical model to assess and analyze the system reliability of smart grids, specifically against topology attacks and system countermeasures (i.e., IDSs and malfunction recovery techniques). Topology attacks, evolving from FmDI attacks, are growing security threats to smart grids. In our analytical model, we define and consider both conservative and aggressive topology attacks, and two types of unreliable consequences (i.e., system disturbances and failures). The IEEE 14-bus power system is employed as a case study to clearly explain the model construction and parameterization process. The benefit of having this analytical model is the capability to measure the system reliability from both transient and steady-

state analyses. Finally, intensive simulation experiments are conducted to demonstrate the feasibility and effectiveness of our proposed model.

3.1 Introduction

The smart grid is envisioned as a revolutionary alternative of the legacy power grid with the primary expectation to achieve enhanced situational awareness of the enormous, and often dispersed, physical infrastructure [49–51].

Despite those promising benefits mentioned in Section 1.2.2, there are a number of underlying issues and challenges [51–53]. System reliability is one of the critical concerns in smart grids, and can be affected by a wide range of grid components [1, 18]. Specifically, demand-response strategies and peak load shedding techniques play significant roles in balancing (due to load variability) power demands and generations to preserve grid reliability. Other considerations important in preserving grid reliability include performance and lifetime of the substations, transmission lines, and electrical devices. Renewable resources, such as wind, solar, hydro, and tidal, may also impact on system reliability due to their volatile nature. Similar to other consumer technologies, ensuring the integrity and authenticity of measurement data reported by sensing devices (e.g., line meters, circuit breaker monitors, and smart meters) are also vital to ensure grid reliability (i.e., in terms of system state estimation and informing decision-making). For example, biased or fabricated measurements could potentially result in the system control center issuing erroneous feedback commands, and consequently, compromising the system reliability.

With the increasing trend in smart grids being the target of cyber attacks and physical sabotages impacting on the reliability of smart grids [6, 54], it is important to ensure a resilient and reliable system design. A successful attack or compromise can have significant impacts, as illustrated by recent incidents (e.g., Stuxnet [6]). While there has been recent interest in smart grid security research, existing literature generally focus on single-event attacks rather

than coordinated attacks. This is partially because existing mathematical tools for modeling and analyzing coordinated attacks are not well developed to handle sophisticated coordinated attacks [55]. For example, attack trees are popular tools in existing literature used to describe the conceptual diagram of a single attack. However, most existing attack tree models are not suitable for modeling and capturing concurrent and coordinated attacks, especially when there are defenses in place. In addition, there are only a few studies introducing modeling tools that can adequately capture the dynamics between attacks and defenses, as well as capturing the synthetic idiosyncrasies of a smart grid cyber-physical system [56]. This is the gap we seek to address in this chapter.

Specifically, we introduce the topology attacks [57], a typical example of coordinated attacks in the context of smart grids. We then use a stochastic Petri net (SPN) [58, 59] to model the topology attacks and analyze the system reliability in the presence of both IDSs and malfunction recovery techniques. Evolving from bad data injection attacks, topology attacks have been the subject of research in recent years. For example, Liu *et al.* showed in 2011 that by compromising a set of metering devices, attackers are capable of constructing an attack vector that can easily circumvent conventional bad data detector; thus, launching a successful bad data injection attack [28]. A key limitation that may impede a successful implementation of such attacks is the need to compromise a large set of metering devices. This is a significantly strong assumption because ordinary attackers usually have limited time and capabilities to construct FmDI attacks in smart grids. To avoid these limitations associated with bad data injection attacks, topology attacks have quickly emerged recently with reduced requirements for attacks. Ideally, by concurrently compromising only a very small set of sensing devices such as line meters and circuit breaker monitors, the adversary could initiate a successful topology attack. Petri nets are tools that have been widely used for modeling various types of asynchronous and concurrent processes; therefore, they are more suitable for modeling the coordinated topology attacks and capturing the concurrent behaviors of both cyber and physical processes in smart grids.

We regard the contributions of this work to be three-fold:

- First, we pioneer to develop a novel analytical model to assess and analyze the system reliability in the presence of both topology attacks and countermeasures in smart grids (i.e., IDSs and malfunction recovery techniques). Since topology attacks are commonly considered as “undetectable” attacks, understanding their attack behaviors and corresponding potential impacts contribute to the building of a more resilient and reliable smart grid.
- Second, we define and characterize two types of topology attacks, namely conservative topology attacks and aggressive topology attacks. Different attack behaviors and their associated impacts on smart grids are then discussed.
- Third, we propose a scheme to determine whether the undetected compromised sensing devices can launch a successful topology attack. Following this, different types of the impacts, e.g., system disturbances or failures, of successful attacks are discussed. In addition, an algorithm for construction of a maximum spanning tree (MxST) [60] in a power system is proposed.

This is to the best of our best knowledge, the first study to use an SPN to study topology attacks in the context of smart grids. The choice of SPN is due to its capability of incorporating features of both the cyber domain (e.g., cyber intrusion process and corresponding state transition process) and the physical domain (e.g., physical measurement data and possible impacts and outages).

We will introduce some preliminaries in Section 3.2, before presenting the system model, threat model, and our design goals in Section 3.3. Our proposed analytical model is elaborated in Section 3.4, and the performance evaluation is presented in Section 3.5. Section 3.6 concludes the chapter.

3.2 Preliminaries

In this section, we introduce some preliminaries relating to Petri net modeling as well as coordinated attacks.

3.2.1 Petri Net Modeling

The Petri net modeling techniques are increasingly popular, partly due to the rapid advancements of networked and distributed systems [59]. A basic Petri net can be described as a 4-tuple $(\mathcal{P}, \mathcal{T}, \mathcal{F}, \mathcal{M})$, where \mathcal{P} is a finite set of places (or states), \mathcal{T} is a finite set of transitions (or actions, behaviors), $\mathcal{F} \subset (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P})$ is a finite set of input and output arcs, and \mathcal{M} is a finite set of markings that denote the number of tokens in each place. A token represents an object that holds a specific condition or the occurrence of a specific event. Tokens can be transferred from place to place when a specific condition changes or event occurs. Since a basic Petri net can only model fairly simple processes, a number of extended Petri nets have been proposed in the literature to support a broader range of applications [61].

The Petri net technique was first used for modeling cyber attacks by, presumably, McDermott as an alternative tool to traditional attack trees [62]. It was demonstrated that Petri nets were more effective than traditional attack trees in describing the concurrent processes. The generalized stochastic Petri net (SPN) technique was subsequently introduced by Bause *et al.* to model cyber attacks [63]. An SPN is a timed Petri net, where the firing time between the transitions is assumed to be exponentially distributed. With the SPN, the state transition process can be easily transformed to a reachability graph, and then a continuous time Markov chain. In this way, it facilitates system administrators in performing steady-state analysis. SPNs are increasingly accepted by the research community, and have been used to support diverse applications [64–68]. Another extension is the colored Petri net, where tokens are represented by different colors. Different from the basic Petri net, a colored Petri net can be used to model more complex systems or processes. For example, Jensen suggested that

colored Petri nets can be used for a wide range of practical applications, such as ATM networks, ISDN networks, and naval vessel systems [65].

Petri net modeling techniques have also been used in power systems to describe the state transition processes of physical systems and the communication infrastructures [66–68]. For instance, Laprie *et al.* used a Petri net to model the interdependence related failures of both electric infrastructure and connected information system. In addition, Zeng *et al.* analyzed the dependability of control center networks in smart grids using SPN by considering that the servers in control center networks can suffer from Byzantine failures and, thus compromise the network dependability. However, there has been no effective modeling technique developed for modeling topology attacks in smart grids, especially in the presence of system countermeasures.

3.2.2 Coordinated Attacks

In the increasingly complex and large-scale cyber-physical infrastructures, it is usually far beyond the capability of a single attacker to disrupt such an infrastructure. It is more likely that well-resourced attackers (e.g. state-sponsored actors and organized cyber crime groups) will attempt to launch a coordinated attack collectively, an observation echoed in the report from CERT [69].

In the context of smart grid CPS, coordinated attacks include false data injection attacks [28, 52, 51], topology attacks [70, 57], DoS, and DDoS attacks [71, 72]. Since Liu demonstrated that a set of coordinated attackers can successfully circumvent the traditional bad data detection mechanisms in power systems [28], researchers have started studying such attacks (e.g., see the survey of attack strategies, potential impacts on power systems, and potential countermeasures in [52]). False data injection attackers can construct an attack vector containing injected false measurement data by compromising an ideal set of data meters. The injected attack vector, if well designed, will easily circumvent bad data detector at the data center without triggering an alarm. Hug *et al.* introduced a new analytical technique

for vulnerability analysis of state estimation, designed to detect hidden false data injection attacks [73]. A distributed host-based collaborative detection scheme for false data injection attacks in smart grid cyber-physical system was also recently proposed [51].

The focus of this work is on the topology attacks that have not yet been widely investigated. Topology attacks are generally considered evolved false data injection attacks, where meter data and breaker status data used for determining the current system topology need to be manipulated. Similar to false data injection attackers, topology attackers also try to blind the bad data detector by constructing matched meter data and breaker status data. A small number of existing studies have discussed such attacks. For example, Weimer *et al.* proposed a distributed detection and isolation method for topology attacks in power networks [70]. Kim and Tong proposed a graph theory based scheme to counter topology attacks by placing the phasor measurement units across the power grid in an optimal way [57]. Apart from the above-mentioned few studies that focus on topology attacks, one particular relevant area that is understudied is topology attack modeling and system reliability analysis when subject to such attacks. Thus, in this chapter, we provide an SPN-based analytical model for smart grids to model the attack behaviors of topology attacks and analyze the system reliability in the presence of such attacks.

3.3 Models and Design Goals

In this section, we formalize both the system and threat models, as well as describing the design goals.

3.3.1 System Model

In this work, we use the power system state estimation workflow as our system model (see Fig. 3.1). The basic concepts of state estimation can be found in Section 2.1. In this section,

we provide additional introduction of how the measurement data can be processed and fed into the state estimator, topology processor, as well as the bad data detector.

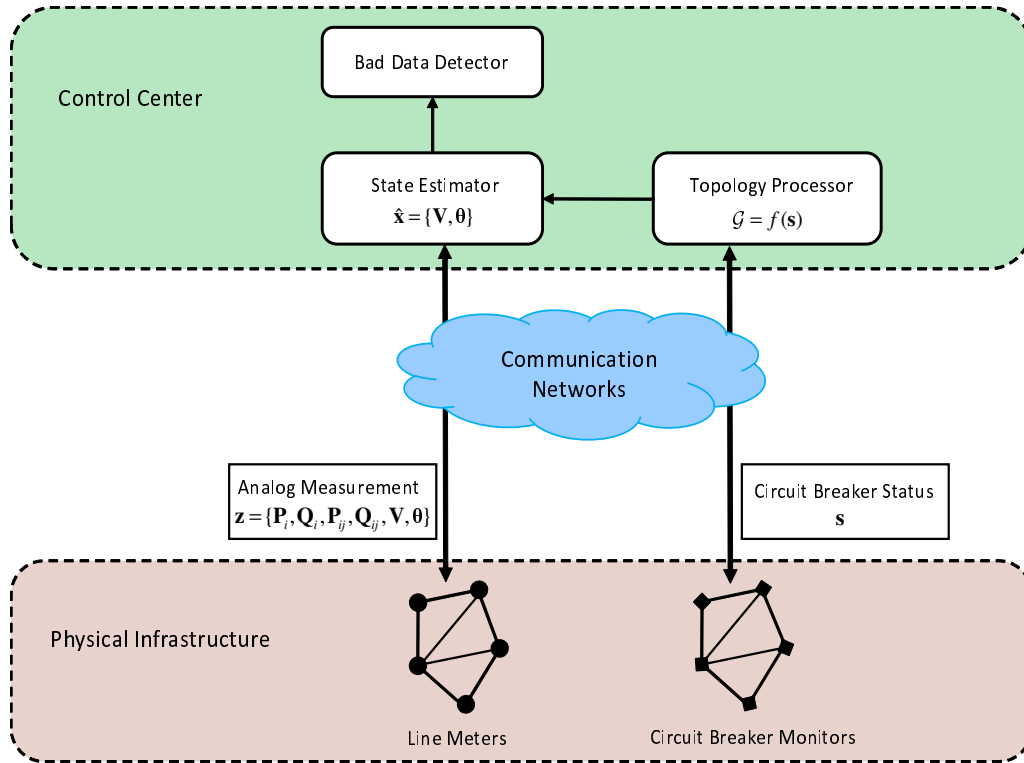


Fig. 3.1 System model: power system state estimation

In state estimation, the system control center collects two types of data from the sensing devices throughout the grid. One type of data is the line flow and nodal injection analog measurement data $\mathbf{z} = \{P_i, Q_i, P_{ij}, Q_{ij}, V, \theta\}$ provided by line meters, where $P_i, Q_i, P_{ij}, Q_{ij}, V, \theta$ denote real power injection, reactive power injection, real power flow, reactive power flow, bus voltage magnitude, and bus voltage angle, respectively. Another type of data is the circuit breaker on/off status data $\mathbf{s} = s_i^N$, where $s_i \in \{0, 1\}$ and N is the total number of branches in a power grid. The status data \mathbf{s} is provided by circuit breaker monitors [74] and then analyzed by the topology processor to determine the current grid topology \mathcal{G} , that is $\mathcal{G} = f(\mathbf{s})$. After that, both measurement data \mathbf{z} and grid topology \mathcal{G} are fed into the state estimator for further data processing. Using an AC or DC power flow model, the state estimator produces the estimated real system status data $\mathbf{x} = \{\hat{V}, \hat{\theta}\}$. At the last

step, through residual checking, the bad data detector determines whether any bad data is collected by the sensing devices.

As mentioned in Section 2.1, the relationship between the measurement data \mathbf{z} and real system status data \mathbf{x} based on the DC power flow model is given by [52]:

$$\mathbf{z} = \mathbf{H}_{\mathcal{G}}\mathbf{x} + \boldsymbol{\eta}, \quad (3.1)$$

where, particularly, $\mathbf{H}_{\mathcal{G}} \in \mathbb{R}^{m \times n}$ is the measurement Jacobian matrix associated with the current system topology \mathcal{G} . Then, the estimated $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} (\mathbf{z} - \mathbf{H}_{\mathcal{G}}\mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}_{\mathcal{G}}\mathbf{x}) = (\mathbf{H}_{\mathcal{G}}^T \mathbf{W}^{-1} \mathbf{H}_{\mathcal{G}})^{-1} \mathbf{H}_{\mathcal{G}}^T \mathbf{W}^{-1} \mathbf{z} \triangleq \boldsymbol{\Lambda} \mathbf{z}, \quad (3.2)$$

where

$$\boldsymbol{\Lambda} \triangleq (\mathbf{H}_{\mathcal{G}}^T \mathbf{W}^{-1} \mathbf{H}_{\mathcal{G}})^{-1} \mathbf{H}_{\mathcal{G}}^T \mathbf{W}^{-1}. \quad (3.3)$$

Then, the estimated measurement data $\hat{\mathbf{z}}$ can then be given by

$$\hat{\mathbf{z}} = \mathbf{H}_{\mathcal{G}} \hat{\mathbf{x}} = \mathbf{H}_{\mathcal{G}} \boldsymbol{\Lambda} \mathbf{z}, \quad (3.4)$$

and the measurement residual \mathbf{r} is calculated by

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = \mathbf{z} - \mathbf{H}_{\mathcal{G}} \boldsymbol{\Lambda} \mathbf{z} = (\mathbf{I} - \mathbf{H}_{\mathcal{G}} \boldsymbol{\Lambda}) \mathbf{z}. \quad (3.5)$$

3.3.2 Adversary Model

In this work, we consider the case where the adversaries are topology attackers whose objective is to compromise the system reliability and survivability. As previously discussed, topology attacks can be regarded as evolved FmDI attacks. With regard to false data injection attacks, adversaries generally tamper with only the measurement data \mathbf{z} . In practice, these false data can often be detected by the bad data detector due to mismatches with the

current system topology \mathcal{G} . However, topology attackers may also attempt to falsify both measurement data \mathbf{z} and circuit breaker status data \mathbf{s} associated with the system topology \mathcal{G} (i.e., $\mathcal{G} = f(\mathbf{s})$ [57]). In other words, they attempt to construct a pair of matched measurement data and the grid topology. Such an attack strategy can be more effective to blind the bad data detector.

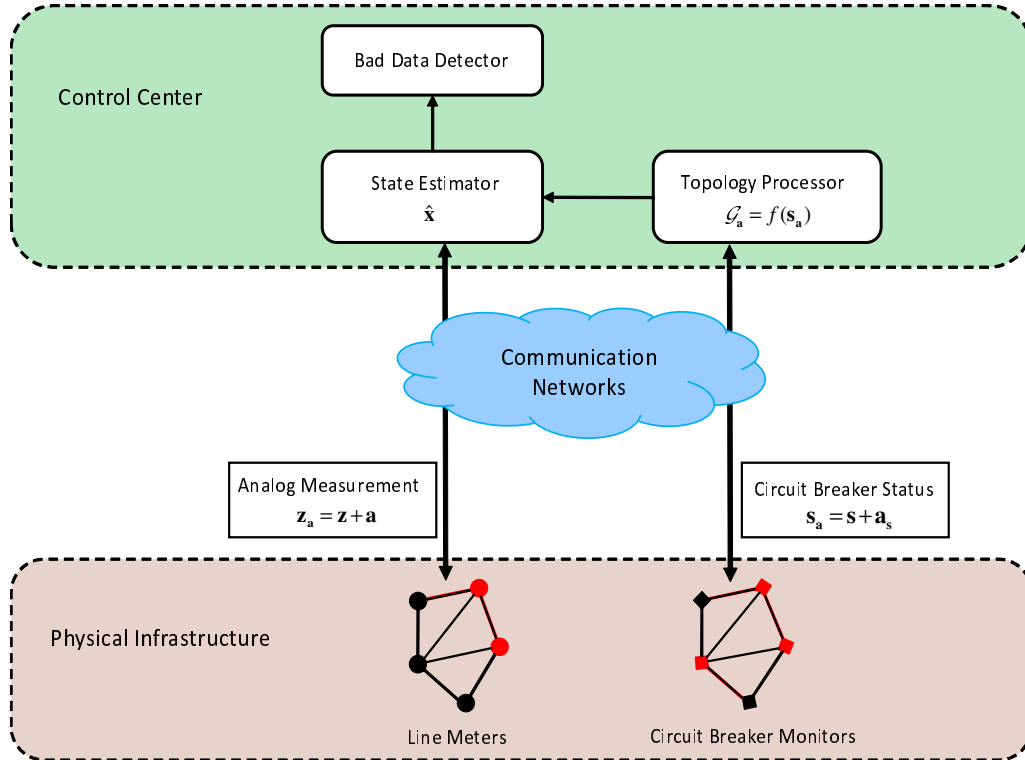


Fig. 3.2 Adversary model

In our adversary model as shown in Fig. 3.2, we assume that the sensing devices (i.e., line meters and circuit breaker monitors) deployed throughout the power grid can be compromised by malicious attackers (including malicious insiders and external attackers). These attackers are capable of controlling the sensing devices to report false measurement data. In Fig. 3.2, compromised line meters and circuit breaker monitors, represented by red circles and squares, construct attack vectors \mathbf{a} and \mathbf{a}_s . Then, the reported measurement data \mathbf{z}_a and circuit breaker

status data \mathbf{s}_a are respectively expressed as the followings:

$$\mathbf{z}_a = \mathbf{z} + \mathbf{a}, \quad (3.6)$$

and

$$\mathbf{s}_a = \mathbf{s} + \mathbf{a}_s. \quad (3.7)$$

Correspondingly, the processed grid topology is given by

$$\mathcal{G}_a = f(\mathbf{s}_a) = f(\mathbf{s} + \mathbf{a}_s), \quad (3.8)$$

where $f(\cdot)$ is the system function of the topology processor.

Likewise, according to the DC power flow model, the relationship between the measurement data \mathbf{z}_a and real system status data \mathbf{x} is described as follows:

$$\mathbf{z}_a = \mathbf{H}_{\mathcal{G}_a} \mathbf{x} + \boldsymbol{\eta}. \quad (3.9)$$

Then, the estimated $\hat{\mathbf{x}}_a$ is given by:

$$\hat{\mathbf{x}}_a = \arg \min_{\mathbf{x}} (\mathbf{z}_a - \mathbf{H}_{\mathcal{G}_a} \mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z}_a - \mathbf{H}_{\mathcal{G}_a} \mathbf{x}) = (\mathbf{H}_{\mathcal{G}_a}^T \mathbf{W}^{-1} \mathbf{H}_{\mathcal{G}_a})^{-1} \mathbf{H}_{\mathcal{G}_a}^T \mathbf{W}^{-1} \mathbf{z}_a \triangleq \boldsymbol{\Lambda}_a \mathbf{z}_a, \quad (3.10)$$

where

$$\boldsymbol{\Lambda}_a = (\mathbf{H}_{\mathcal{G}_a}^T \mathbf{W}^{-1} \mathbf{H}_{\mathcal{G}_a})^{-1} \mathbf{H}_{\mathcal{G}_a}^T \mathbf{W}^{-1}. \quad (3.11)$$

Thus, the estimated measurement data $\hat{\mathbf{z}}_a$ is calculated by

$$\hat{\mathbf{z}}_a = \mathbf{H}_{\mathcal{G}_a} \hat{\mathbf{x}} = \mathbf{H}_{\mathcal{G}_a} \boldsymbol{\Lambda}_a \mathbf{z}_a. \quad (3.12)$$

The normalized residual $\bar{\mathbf{r}}_{\mathbf{a}}$ is then given by

$$\bar{\mathbf{r}}_{\mathbf{a}} = \sqrt{\mathbf{W}^{-1}}(\mathbf{z}_{\mathbf{a}} - \hat{\mathbf{z}}_{\mathbf{a}}) = \sqrt{\mathbf{W}^{-1}}(\mathbf{I} - \mathbf{H}_{\mathcal{G}_{\mathbf{a}}}\Lambda_{\mathbf{a}})\mathbf{z}_{\mathbf{a}}. \quad (3.13)$$

One last and critical step is to detect the bad data. The Frobenius norm $\|\bar{\mathbf{r}}_{\mathbf{a}}\| = \|\sqrt{\mathbf{W}^{-1}}(\mathbf{I} - \mathbf{H}_{\mathcal{G}_{\mathbf{a}}}\Lambda_{\mathbf{a}})\mathbf{z}_{\mathbf{a}}\|$ can be seen as a function of \mathbf{a} and $\mathbf{s}_{\mathbf{a}}$ (recall that $\mathcal{G} = f(\mathbf{s})$). In this way, as long as the constructed vectors \mathbf{a} and $\mathbf{s}_{\mathbf{a}}$ can lead to

$$\|\bar{\mathbf{r}}_{\mathbf{a}}\| = \|\sqrt{\mathbf{W}^{-1}}(\mathbf{I} - \mathbf{H}_{\mathcal{G}_{\mathbf{a}}}\Lambda_{\mathbf{a}})\mathbf{z}_{\mathbf{a}}\| < \tau, \quad (3.14)$$

and the adversaries can launch successful topology attacks.

In this work, we define two types of topology attacks in terms of the attack strategies, which are shown as follows:

- *Conservative topology attacks*: such attacks aim to manipulate a single or a few transmission lines or buses by compromising a small number of sensing devices. Accordingly, manipulation of these limited resources results in minor impact on the power system (e.g. disturbances).
- *Aggressive topology attacks*: such attacks attempt to manipulate as large an area of power grid as possible (e.g. by compromising as many sensing devices as possible). These attacks usually result in devastating damages to the power system (e.g. system failures), if successful.

Note that in the adversary/threat model of each piece of work presented in this thesis, we do not consider those data integrity attacks targeting on the communication links (e.g., man-in-the-middle attacks). We propose solutions to detect FDI attacks on the basis that the data integrity while data transmission networks are ensured. If in need, the BLS short signature [75] can be employed to ensure data integrity during transmission.

3.3.3 Design Goals

The key objective of our work is to provide an analytical model for studying topology attacks in smart grids, as well as analyzing the system reliability in the presence of such attacks. Specifically, our design goals are as follows:

1. Carrying out in-depth analyses on the attack strategies of different types of topology attacks and potential impacts on the power system they may cause.
2. Establishing an SPN-based state transition model for smart grids to describe the system behaviors in the presence of topology attacks.
3. Defining credible metrics to accurately assess the system reliability of smart grids.

3.4 Proposed Analytical Model

In this section, we present our SPN-based analytical model (see Fig. 3.3) to describe the system behaviors in the presence of both topology attacks and system security countermeasures (e.g. IDSs and malfunction recovery techniques).

3.4.1 Construction of the Proposed SPN Model

We now present the construction of the proposed SPN model. Tables 3.1 and 3.2 annotate the physical meanings of places and transitions in the SPN model, respectively. Cyber transitions are denoted using blank bars, while physical transitions are shown as filled bars. Note that, in particular, immediate transitions also appear in our model, which belong to cyber transitions and they are presented by slim vertical bars. In this SPN model, we mainly consider two types of sensing devices, namely line meters and circuit breaker monitors. Small filled circles in red (tokens) are used to represent the sensing devices holding specific conditions. In terms of countermeasures, we use filled black stars ★ to denote the presence of IDSs. The IDSs are deployed for periodical detection of sensing device malfunctions. In addition, the

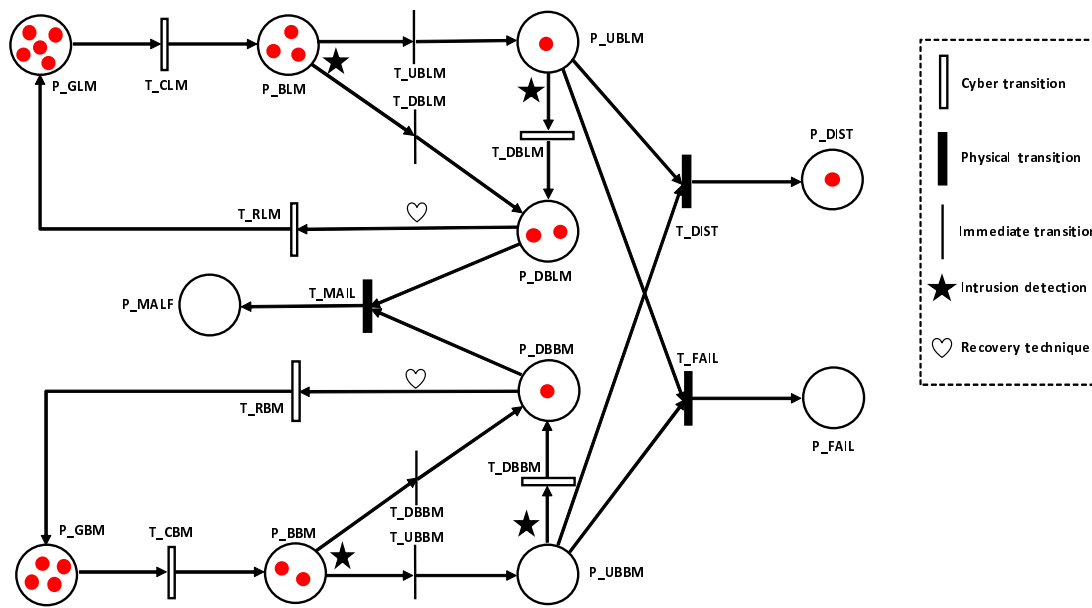


Fig. 3.3 Analytical SPN model

malfunction recovery techniques, represented by blank hearts ♡, are designed for recovering the malfunction devices identified by the IDSs.

The SPN model has an 11-element set of place $\mathcal{P} = \{P_GLM, P_BLM, P_GBM, P_BBM, P_DBLM, P_UBLM, P_DBBM, P_UBBM, P_DIST, P_FAIL, P_MALF\}$. Specifically, places $P_GLM, P_BLM, P_GBM, P_BBM$ hold the counts for good line meters, bad line meters, good breaker monitors, and bad breaker monitors, respectively. Likewise, places $P_DBLM, P_UBLM, P_DBBM, P_UBBM$ hold the counts for detected bad line meters, undetected bad line meters, detected bad breaker monitors, and undetected bad breaker monitors, respectively. Place P_DIST , if holding a token, represents a system disturbance event resulting from places P_UBLM and P_UBBM when transition T_DIST is enabled. Similarly, place P_FAIL , if holding a token, represents a system failure event resulting from places P_UBLM and P_UBBM when transition T_FAIL is enabled. In particular, if place P_MALF holds a token, the whole power system is encountered with a system malfunction transferred from places P_DBLM and P_DBBM when detected bad line meters and breaker monitors are unable to be recovered timely.

Table 3.1 Places in the SPN model

Place	Meaning
P_GLM	Place of good line meters
P_BLM	Place of bad line meters
P_GBM	Place of good breaker monitors
P_BBM	Place of bad breaker monitors
P_DBLM	Place of detected bad line meters
P_UBLM	Place of undetected bad line meters
P_DBBM	Place of detected bad breaker monitors
P_UBBM	Place of undetected bad breaker monitors
P_DIST	Place of system disturbance: 0 before and 1 after
P_FAIL	Place of system failure: 0 before and 1 after
P_MALF	Place of system malfunction: 0 before and 1 after

Table 3.2 Transitions in the SPN model

Transition	Meaning
T_CLM	Transition that the attacker compromises a line meter
T_CBM	Transition that the attacker compromises a breaker monitor
T_DBLM	Transition that the IDS detects a bad line meter
T_UBLM	Transition that the IDS fails to detect a bad line meter
T_DBBM	Transition that the IDS detects a bad breaker monitor
T_UBBM	Transition that the IDS fails to detect a bad breaker monitor
T_RLM	Transition that the system operator recovers a line meter
T_RBM	Transition that the system operator recovers a breaker monitor
T_DIST	Transition that the power grid encounters a system disturbance
T_FAIL	Transition that the power grid encounters a system failure
T_MALF	Transition that the power grid encounters a system malfunction

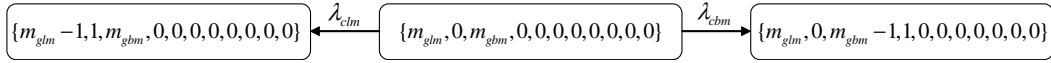


Fig. 3.4 System state transitions triggered by the second event

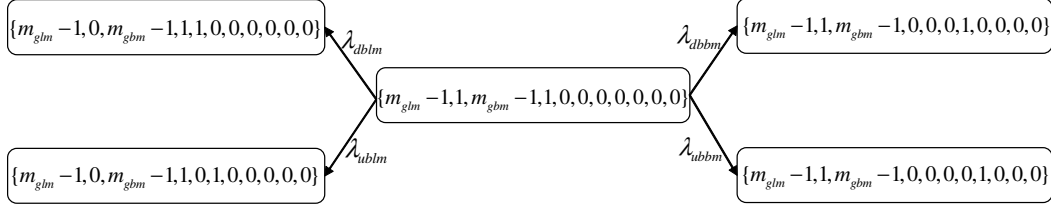


Fig. 3.5 System state transitions triggered by the third event

We use the following events to show how this SPN model is constructed, and how the system behaves under various event triggers.

- The first event is model initialization. We use tokens in a place to represent the sensing devices that meet the conditions specified by this place. Particularly, as for places P_DIST, P_FAIL, and P_MALF, holding a token represents the occurrence of this event; otherwise, empty places denote no occurrence of such events. A marking is a sequence of token states in all the places, which is denoted by $\mathcal{M} = \{m_{glm}, m_{blm}, m_{gbm}, m_{bbm}, m_{dblm}, m_{ublm}, m_{dbbm}, m_{ubbm}, m_{dist}, m_{fail}, m_{malf}\}$, where, particularly as abovementioned, m_{dist} , m_{fail} , and m_{malf} can only take values of either 0 or 1. Initially, all the devices are uncompromised/good, thereby the marking can be initialized as $\mathcal{M}_0 = \{m_{glm}, 0, m_{gbm}, 0, 0, 0, 0, 0, 0, 0, 0\}$.
- The second event is an attacker compromising a line meter or a breaker monitor. We use the compromising rates λ_{clm} and λ_{cbm} to denote, for each token in good places, the average number of tokens per unit time that can be transferred to bad places. Such transitions can be seen from Fig. 3.4. Firing a transition will move one token from the input place to the output place. For example, firing T_CLM in state $\{m_{glm}, 0, m_{gbm}, 0, 0, 0, 0, 0, 0, 0, 0\}$ will move one token from P_GLM to P_BLM, transforming to state $\{m_{glm} - 1, 1, m_{gbm}, 0, 0, 0, 0, 0, 0, 0, 0\}$.

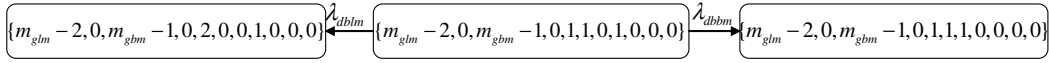


Fig. 3.6 System state transitions triggered by the fourth event

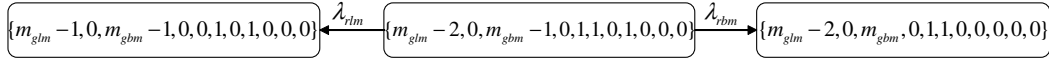


Fig. 3.7 System state transitions triggered by the fifth event

- The third event is concerned with detecting or failing to detect a compromised bad sensing device from place P_BLM or P_BB M using the IDS. For a newly compromised device within a detection interval, the IDS may fire two kinds of transitions. For example, as shown in Fig. 3.5, if the IDS successfully detects a bad line meter in state $\{m_{glm} - 1, 1, m_{gbm} - 1, 1, 0, 0, 0, 0, 0, 0, 0\}$, T_DBLM will be fired transforming the system into state $\{m_{glm} - 1, 0, m_{gbm} - 1, 1, 1, 0, 0, 0, 0, 0, 0\}$ with a rate $\lambda_{dbl m}$; otherwise, T_UBLM will be fired transforming the system into state $\{m_{glm} - 1, 0, m_{gbm} - 1, 1, 0, 1, 0, 0, 0, 0, 0\}$ with a rate $\lambda_{ubl m}$. Similar transitions for bad breaker monitors transitions are also shown in Fig. 3.5.
- The fourth event is concerned with detecting a compromised bad device from place P_UBLM or P_UBBM using the IDS. Devices fell in place P_UBLM or P_UBBM are compromised bad devices that, heretofore, have not been detected yet. The IDS runs periodically to check all the devices, perhaps by trust reputation [49]; thus, the compromised devices may be identified at any detection interval or even undetected for a significantly long period. As shown in Fig. 3.6, if the IDS successfully detects a bad line meter in state $\{m_{glm} - 2, 0, m_{gbm} - 1, 0, 1, 1, 0, 1, 0, 0, 0\}$, T_DBLM will be fired transforming the system into state $\{m_{glm} - 2, 0, m_{gbm} - 1, 0, 2, 0, 0, 1, 0, 0, 0\}$ with a rate $\lambda_{dbl m}$. Similarly, if transition T_DBBM is fired, a bad breaker monitor will be detected transforming the system into state $\{m_{glm} - 2, 0, m_{gbm} - 1, 0, 1, 1, 1, 0, 0, 0, 0\}$ with a rate $\lambda_{dbb m}$.

- The fifth event is concerned with recovering a detected bad device using malfunction recovery techniques. When a bad device is successfully detected by an IDS, the system administrator will carry out the malfunction recovery techniques to record and reset the compromised bad device. As shown in Fig. 3.7, if transition T_RLM is fired with a rate λ_{rlm} in state $\{m_{glm} - 2, 0, m_{gbm} - 1, 0, 1, 1, 0, 1, 0, 0, 0\}$, a detected bad line meter is recovered to a good line meter, transforming the system state into $\{m_{glm} - 1, 0, m_{gbm} - 1, 0, 0, 1, 0, 1, 0, 0, 0\}$. Likewise, if transition T_RBM is fired with a rate λ_{rbm} in state $\{m_{glm} - 2, 0, m_{gbm} - 1, 0, 1, 1, 0, 1, 0, 0, 0\}$, a detected bad breaker monitor is recovered to a good breaker monitor, transforming the system state into $\{m_{glm} - 2, 0, m_{gbm}, 0, 1, 1, 0, 0, 0, 0, 0\}$.
- The sixth event considers a successful topology attack, usually conservative, causing a system disturbance. A small number of undetected bad line meters and breaker monitors can construct a conservative topology attack to fire transition T_DIST. An example is shown in Fig. 3.8, where a system disturbance occurs with state $\{m_{glm} - 3, 0, m_{gbm} - 1, 0, 1, 2, 0, 1, 1, 0, 0\}$ resulting from state $\{m_{glm} - 3, 0, m_{gbm} - 1, 0, 1, 2, 0, 1, 0, 0, 0\}$ when T_DIST is enabled. The enabling function is a complex process based on the spanning tree of the power grid topology, which is detailed in the unreliability enabling scheme that will be introduced in the next subsection.
- The seventh event considers a successful topology attack, usually aggressive, causing a system failure. A multitude of undetected bad line meters and breaker monitors can collectively construct an aggressive topology attack to fire transition T_FAIL. An example is shown in Fig. 3.9, where a system failure occurs with state $\{m_{glm} - 5, 0, m_{gbm} - 2, 0, 1, 4, 0, 2, 0, 1, 0\}$ resulting from state $\{m_{glm} - 5, 0, m_{gbm} - 2, 0, 1, 4, 0, 2, 0, 0, 0\}$ when T_FAIL is enabled. Likewise, the enabling function is a complex process based on the spanning tree of the power grid topology, which is detailed in the unreliability enabling scheme that will be introduced in the next subsection.

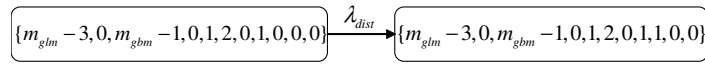


Fig. 3.8 System state transitions triggered by the sixth event

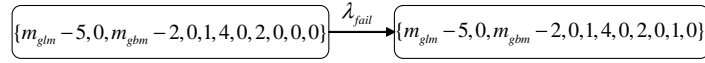


Fig. 3.9 System state transitions triggered by the seventh event

- The last event considers a system malfunction caused by insufficient good sensing devices. If the system has a low recovery rate (i.e., λ_{rlm} and λ_{rbm} are significantly small values), the detected bad sensing devices cannot be recovered in time leaving many detected bad devices remained in places P_DBLM and P_DBBM. In this case, insufficient good sensing devices can operate normally to support the wide area monitoring functionality. Then, the power system malfunctions because the system states are no longer fully observable to the system control center [76]. As we see, a number of unrecovered bad line meters and breaker monitors can collectively fire transition T_MALF to cause a system malfunction. As shown in Fig. 3.10, a system malfunction with state $\{m_{glm} - 9, 0, m_{gbm} - 5, 0, 8, 1, 5, 0, 0, 0, 1\}$ may occur from state $\{m_{glm} - 9, 0, m_{gbm} - 5, 0, 8, 1, 5, 0, 0, 0, 0\}$ with a rate λ_{mal} when T_MALF is enabled. The enabling function is straightforward and integrated in the unreliability enabling scheme that will be introduced in the next subsection.

3.4.2 Maximum Spanning Tree Based Unreliability Enabling Scheme

We will now present the proposed scheme composed of two algorithms to determine under what conditions the power system will fall into the unreliability status (i.e., disturbance, failure, or malfunction).



Fig. 3.10 System state transitions triggered by the eighth event

Table 3.3 Weights assigned to each bus

Bus type	Description	Weight assigned
Type 1	Bus with line(s) only but no generator or load	1 unit
Type 2	Bus with line(s) and load(s) but no generator	2 units
Type 3	Bus with line(s) and generator but no load	3 units
Type 4	Bus with line(s), generator and load(s)	4 units

MxST Construction Algorithm

In our scheme, we use the spanning tree in graph theory to determine the most critical measurements. According to the contraction-deletion theorem [77], there may be multiple spanning trees for a graph \mathcal{G} . To obtain the best results, we use the MxST [60]. MxST is a spanning tree of a weighted graph \mathcal{G} , where the weight sum of all edges is the maximum over all \mathcal{G} 's spanning trees. In our scheme, we allocate weights towards the branches to indicate the different levels of significance to the power grid, in terms of its observability and reliability. This is how we use the MxST of a grid topology to determine the most critical branches of a power grid.

There may be various methods to determine the weights assigned to each line. In this work, we use a straightforward approach to achieve this goal as shown in Table 3.3. Buses can be classified into four types, namely: (1) buses with line(s) only but no generator or load; (2) buses with line(s) and load(s) but no generator; (3) buses with line(s) and generator but no load; and (4) buses with line(s), generator and load(s) [79]. From the system administrator's perspective, the system administrator may be more interested in buses with generators and/or loads than those with only transmission lines; and in buses with generators than those with loads, due to cost savings and reliability concerns. As such, we simply assign branches connected to a bus with a total weight of 1, 2, 3, and 4 units respectively for the four types of buses. Then, the total weight is equally divided among all connected branches. For example,

Algorithm 3.1 MxST Construction

Input: Initial graph $\mathcal{G} = \{\mathcal{V}_{\mathcal{G}}, \mathcal{E}_{\mathcal{G}}\}$ of a power grid topology; set of weights \mathcal{W} assigned for all the branches

Output: A MxST $\mathcal{S} = \{\mathcal{V}_{\mathcal{S}}, \mathcal{E}_{\mathcal{S}}\}$

- 1: *Initialization:* $\mathcal{V}_{\mathcal{S}} = \emptyset, \mathcal{E}_{\mathcal{S}} = \emptyset$
- 2: Step 1: Weight Assignment for each branch.
- 3: (1.1). Assign a total weight to each bus according to Table 3.3.
- 4: (1.2). Equally divide the weight assigned for each bus into k parts, where k is the number of branches connected to this bus.
- 5: (1.3). Assign the divided weights to each connected branch.
- 6: (1.4). Add the weights for each branch assigned from the two end buses.
- 7: Step 2: Arrange all branches in a decreasing order of their weights using, for example, the quicksort algorithm [78].
- 8: Step 3: Add to $\mathcal{E}_{\mathcal{S}}$ with ϵ_{ij} that has the maximum weight $\omega_{ij} \in \mathcal{W}$; add to $\mathcal{V}_{\mathcal{S}}$ with ν_i and ν_j that is connected by ϵ_{ij} .
- 9: Step 4: Remove ν_i and ν_j from $\mathcal{V}_{\mathcal{G}}$, and ϵ_{ij} from $\mathcal{E}_{\mathcal{G}}$.
- 10: Step 5: Loop over all the remaining edges $\epsilon_{kt} \in \mathcal{E}_{\mathcal{G}}$ connecting to vertices $\nu_k \in \mathcal{V}_{\mathcal{S}}$. Add the edge ϵ_{kt} has currently the maximum weight $\omega_{kt} \in \mathcal{W}$ to $\mathcal{E}_{\mathcal{S}}$; add $\nu_t \in \mathcal{V}_{\mathcal{G}}$ but $\nu_t \notin \mathcal{V}_{\mathcal{S}}$ to $\mathcal{V}_{\mathcal{S}}$.
- 11: Step 6: Remove the edge ϵ_{kt} from $\mathcal{E}_{\mathcal{G}}$ and ν_t from $\mathcal{V}_{\mathcal{G}}$ concerned in the last step.
- 12: Step 7: Repeat Steps 5 and 6 until $\mathcal{V}_{\mathcal{G}} = \emptyset$.
- 13: Step 8: Add all generators to $\mathcal{V}_{\mathcal{S}}$, and generator and load edges to $\mathcal{E}_{\mathcal{S}}$.
- 14: **Return:** $\mathcal{S} = \{\mathcal{V}_{\mathcal{S}}, \mathcal{E}_{\mathcal{S}}\}$

if a type 2 bus has 4 branches, then this bus is assigned a total weight of $\omega = 2$ units, and each of its 4 branches is allocated a weight of $\omega_i = 2/4 = 0.5$ unit, where $i \in \{1, 2, 3, 4\}$.

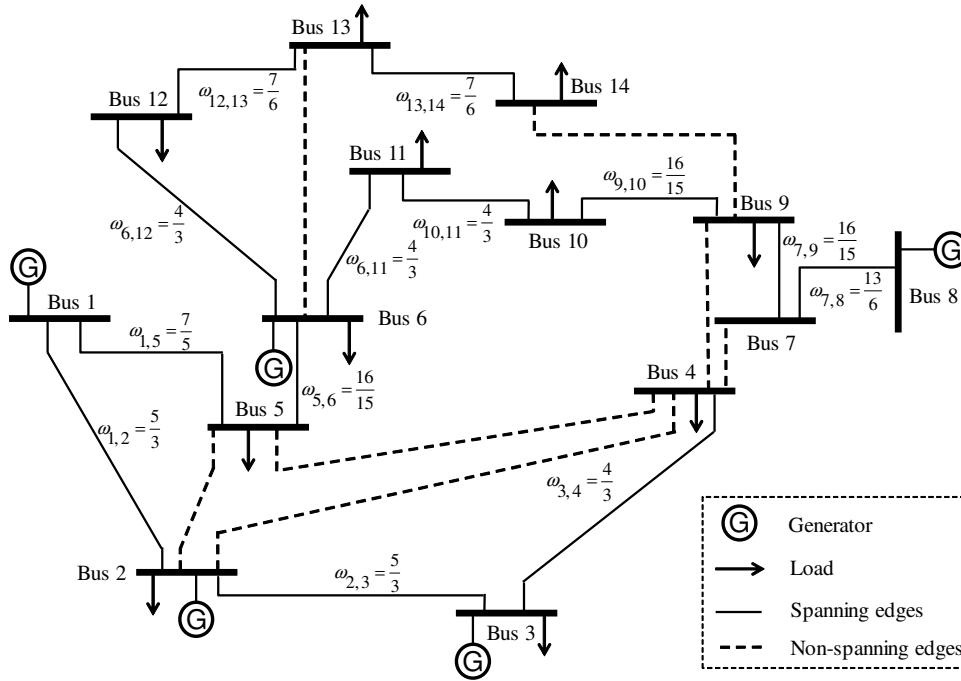


Fig. 3.11 The MxST of the IEEE 14-bus system

Based on this method, we develop an algorithm to show the construction of an MxST in a power grid (see Algorithm 3.1), with reference to the existing algorithm for constructing a minimum spanning tree [80, 81]. In Algorithm 3.1, the weights for all branches are calculated by equally dividing the total weights of each bus and adding the two component weights for each branch. Then, MxST is constructed for the weighted graph step by step. Starting from a branch with the highest weight, branches and buses of interest are added to MxST following a decreasing order of weights, until all bus nodes are added to MxST but avoid adding them repeatedly. Lastly, put in all generators, and generator and load edges to MxST as well, because generators and loads are always important to a power grid. Note that the time complexity of Algorithm 3.1, which is mainly determined by the time complexity of the sorting algorithm used in Step 2, is $\mathcal{O}(|\mathcal{E}_G| \log |\mathcal{E}_G|)$ if using the quicksort algorithm [78].

In order to prove the correctness of this algorithm, we give the following theorem with its proof.

Theorem 1. *After running Algorithm 3.1 on a connected weighted graph \mathcal{G} , its output \mathcal{S} is an MxST.*

Proof. First, \mathcal{S} is a spanning tree. This is because:

- \mathcal{S} is a forest. No cycles are ever created.
- \mathcal{S} is spanning. Suppose that there is a vertex ν_k that is not incident to the edges of \mathcal{S} . Then the incident edges of ν_k must have been considered in the algorithm at some step. The first edge (in edge order) would have been included because it could not have created a cycle, which contradicts the definition of \mathcal{S} .
- \mathcal{S} is connected. Suppose that \mathcal{S} is not connected. Then \mathcal{S} has two or more connected components. Since \mathcal{G} is connected, then these components must be connected by some edges in \mathcal{G} , not in \mathcal{S} . The first of these edges (in edge order) would have been included in \mathcal{S} because it could not have created a cycle, which contradicts the definition of \mathcal{S} .

Second, \mathcal{S} is a spanning tree of maximum weight. We will prove this using induction. Let \mathcal{S}^* be an MxST. If $\mathcal{S} = \mathcal{S}^*$, then \mathcal{S} is an MxST. If $\mathcal{S} \neq \mathcal{S}^*$, then there exists an edge $e \in \mathcal{S}^*$ of maximum weight that is not in \mathcal{S} . Further, $\mathcal{S} \cup e$ contains a cycle C such that:

1. Every edge in C has weight larger than $wt(e)$, where $wt(\cdot)$ presents the weight of an edge. (This follows from how the algorithm constructed \mathcal{S} .)
2. There is some edge f in C that is not in \mathcal{S}^* . (Because \mathcal{S}^* does not contain the cycle C .)

Consider the tree $\mathcal{S}_2 = \mathcal{S} \setminus \{e\} \cup \{f\}$:

1. \mathcal{S}_2 is a spanning tree.

2. \mathcal{S}_2 has more edges in common with \mathcal{S}^* than \mathcal{S} did.
3. And $wt(\mathcal{S}_2) \leq wt(\mathcal{S})$. (We exchanged an edge for one that is no more expensive.)

We can redo the same process with \mathcal{S}_2 to find a spanning tree \mathcal{S}_3 with more edges in common with \mathcal{S}^* . By induction, we can continue this process until we reach \mathcal{S}^* , from which we see $wt(\mathcal{S}) \geq wt(\mathcal{S}_2) \geq wt(\mathcal{S}_3) \geq \dots \geq wt(\mathcal{S}^*)$. Since \mathcal{S}^* is a MxST, then these inequalities must be equalities and we conclude that \mathcal{S} is an MxST.

□

Let us take the IEEE 14-bus power system (as shown in Fig. 3.11) as an example to introduce the construction of an MxST in a power system. Based on our proposed scheme, the bus types of this power system and the total weights assigned to each bus are summarized in Table 3.4. Then, the weights of all branches in the power system are calculated and listed in Table 3.5. According to Algorithm 3.1, we construct the MxST of the IEEE 14-bus system as shown in Fig. 3.11, wherein all the MxST branches are denoted by solid red lines and the weights of all the MxST branches are annotated.

Table 3.4 The bus type and total weight assigned in IEEE 14-bus system

Bus	Bus type	Weight	Bus	Bus type	Weight
#1	Type 3	3 units	#8	Type 3	3 units
#2	Type 4	4 units	#9	Type 2	2 units
#3	Type 4	4 units	#10	Type 2	2 units
#4	Type 2	2 units	#11	Type 2	2 units
#5	Type 2	2 units	#12	Type 2	2 units
#6	Type 4	4 units	#13	Type 2	2 units
#7	Type 1	1 unit	#14	Type 2	2 units

Table 3.5 Weights assigned for each branch in IEEE 14-bus system (g denotes generator and l denotes load)

Branch	Weight	Branch	Weight	Branch	Weight
$\epsilon_{1,2}$	$3/3+2/3=5/3$	$\epsilon_{4,7}$	$2/6+2/3=1$	$\epsilon_{8,g}$	$3/2$
$\epsilon_{1,5}$	$3/3+2/5=7/5$	$\epsilon_{4,9}$	$2/6+2/5=11/15$	$\epsilon_{9,10}$	$2/5+2/3=16/15$
$\epsilon_{1,g}$	$3/3=1$	$\epsilon_{4,l}$	$2/6=1/3$	$\epsilon_{9,14}$	$2/5+2/3=16/15$
$\epsilon_{2,3}$	$4/6+4/4=5/3$	$\epsilon_{5,6}$	$2/5+4/6=16/15$	$\epsilon_{9,l}$	$2/5$
$\epsilon_{2,4}$	$4/6+2/6=1$	$\epsilon_{5,l}$	$2/5$	$\epsilon_{10,11}$	$2/3+2/3=4/3$
$\epsilon_{2,5}$	$4/6+2/5=16/15$	$\epsilon_{6,11}$	$4/6+2/3=4/3$	$\epsilon_{10,l}$	$2/3$
$\epsilon_{2,g}$	$4/6=2/3$	$\epsilon_{6,12}$	$4/6+2/3=4/3$	$\epsilon_{11,l}$	$2/3$
$\epsilon_{2,l}$	$4/6=2/3$	$\epsilon_{6,13}$	$4/6+2/4=7/6$	$\epsilon_{12,13}$	$2/3+2/4=7/6$
$\epsilon_{3,4}$	$4/4+2/6=4/3$	$\epsilon_{6,g}$	$4/6=2/3$	$\epsilon_{12,l}$	$2/3$
$\epsilon_{3,g}$	$4/4=1$	$\epsilon_{6,l}$	$4/6=2/3$	$\epsilon_{13,14}$	$2/4+2/3=7/6$
$\epsilon_{3,l}$	$4/4=1$	$\epsilon_{7,8}$	$2/3+3/2=13/6$	$\epsilon_{13,l}$	$2/4=1/2$
$\epsilon_{4,5}$	$2/6+2/5=11/15$	$\epsilon_{7,9}$	$2/3+2/5=16/15$	$\epsilon_{14,l}$	$2/3$

Unreliability Judgement Scheme

We employ MxST in our analytical model with the expectation to find the most critical branches, which (together with all the buses) provide the most useful information of power system operation status. As such, we can define the *critical devices* and *critical data*.

Definition 1. Given an MxST \mathcal{S} of a power grid topology \mathcal{G} , sensing devices are termed as *critical devices* if they host on branches ϵ_{ij} that are involved in \mathcal{S} , that is

$$\epsilon_{ij} \in \mathcal{G} \cap \mathcal{S}, i, j \in \{1, 2, \dots, N_S\} \text{ and } i \neq j, \quad (3.15)$$

where N_S is the number of edges in \mathcal{S} ; otherwise, they are termed as *non-critical devices* if

$$\epsilon_{ij} \in \mathcal{G} \setminus \mathcal{S}, i, j \in \{1, 2, \dots, N_S\} \text{ and } i \neq j. \quad (3.16)$$

Algorithm 3.2 Unreliability Judgement Scheme

Input: Set of compromised line meter \mathcal{L} ; set of compromised breaker monitor \mathcal{B} ; grid topology \mathcal{G} ; MxST \mathcal{S} of \mathcal{G} ; number of unrecovered detected bad devices N_L and N_B

Output: Decision outcome O

- 1: *Initialization:* threshold N_{th} , the maximum number of unrecovered detected bad devices a system can tolerate prior to a system malfunction
 - 2: **if** $N_L > N_{th}$ or $N_B > N_{th}$ **then**
 - 3: $O \leftarrow$ system malfunction
 - 4: **else if** at least one pair of the compromised line meters and breaker monitors are located at the same branch **then**
 - 5: The attacker is capable of launching a topology attack.
 - 6: **if** at least one of the compromised line meters and breaker monitors are *critical devices* **then**
 - 7: $O \leftarrow$ system failure
 - 8: **else**
 - 9: $O \leftarrow$ system disturbance
 - 10: **end if**
 - 11: **else**
 - 12: $O \leftarrow$ bad data detected with no system unreliability
 - 13: **end if**
- return** O
-

Accordingly, the measurement data or status data generated by these *critical devices* are *critical data*, and data generated by the *non-critical devices* are termed as *non-critical data*.

With the constructed MxST and the above definitions, we design an unreliability judgement scheme to show under what circumstances can transitions T_DIST, T_FAIL and T_MALF be fired to cause system unreliability. As described in Algorithm 3.2, the first step is to check whether the compromised devices have the capability to collectively construct a topology attack. We consider a most optimistic condition from the attackers' perspective that, if the line meter and the circuit breaker monitor on the same branch are unfortunately compromised by the adversary and undetected by the IDS, the attackers are considered to have the capability to launch a topology attack. Otherwise, the injected false meter data can be easily detected by the bad data detector. Then, for cases where the attackers have the capability to launch a topology attack, further classification is conducted to determine whether a system failure or disturbance happens. Note that, this classification procedure also, to a

large extent, successfully differentiates between conservative topology attacks and aggressive topology attacks. The reason is that, according to the definition provided in Section 3.3.2, conservative topology attacks usually cause system disturbances while aggressive topology attacks usually cause system failures. The time complexity of Algorithm 3.2, which is mainly determined by the operation as Line 4 shows, is $\mathcal{O}(N_L \times N_B)$.

3.5 Performance Evaluation

3.5.1 Performance Metrics

In this chapter, the reliability performance of the smart grids using our analytical model is analyzed using both transient analysis and steady-state analysis.

Transient Analysis

The metrics for transient analysis are the mean time to disturbance (MTTD) and mean time to failure (MTTF). Specifically, MTTD is the average time before the power system functions into a system disturbance. Likewise, MTTF is the average time before the power system functions into a system failure. They are given by

$$\text{MTTD} = \int_0^{\infty} t[1 - Q_D(t)]dt, \quad (3.17)$$

and

$$\text{MTTF} = \int_0^{\infty} t[1 - Q_F(t)]dt, \quad (3.18)$$

where $Q_D(t)$ is the probability of the first visit to a system disturbance, and $Q_F(t)$ is the probability of the first visit to a system failure. Note that in the transient analysis, we ignore the mean time to malfunction (MTTM). The reason is that compared to MTTD and MTTF, MTTM is considerably larger due to the negligible probability of a system malfunction occurrence.

Steady-State Analysis

The steady-state analysis is presented by a self-defined metric: *reliability* R , which is defined by

$$R = (1 - p_{malf}) * \left(1 - \frac{\alpha * p_{dist} + \beta * p_{fail}}{\alpha + \beta}\right)^k, \quad (3.19)$$

where

$$p_{malf} = 1 - \sum_{i=0}^{N_{th}} \sum_{j=0}^{N_{th}} p_{dblm}(i) p_{dbbm}(j), \quad (3.20)$$

denoting the steady-state probability of a system malfunction occurrence when the number of unrecovered detected bad devices in either place P_DBLM or P_DBBM exceeds the acceptable threshold N_{th} . p_{dist} and p_{fail} denote the steady-state probabilities of system disturbance and failure, respectively. In addition, α and β represent the negative impacts of p_{dist} and p_{fail} posed to the system reliability. k is the average number of pairs of compromised line meter and breaker monitor that host on the same transmission line. k describes the average number of attack events in the power system under the optimistic condition. In practice, the power system usually has a sufficient recovery rate, wherein a really small value can be satisfied; therefore, $p_{dblm}(i)$ and $p_{dbbm}(j)$ always hold large probabilities when i and j are small values (e.g., 0 or 1). According to Eq. (3.20), p_{malf} approaches to zero in steady state, which is negligible. In this case, Eq. (3.19) can be reduced to

$$R = \left(1 - \frac{\alpha * p_{dist} + \beta * p_{fail}}{\alpha + \beta}\right)^k. \quad (3.21)$$

Since the steady states of P_DIST and P_FAIL are absorbing states, it is hard to find the corresponding steady-state probabilities. Therefore, we transform this problem into several sub-problems. A corresponding workflow is shown in Fig. 3.12. In this workflow, the SPN model is reduced by temporarily removing places P_DIST and P_FAIL (P_MALF as well) and corresponding transitions T_DIST and T_FAIL (T_MALF as well). Then, the steady-state probabilities of $p_{ublm}(i)$ and $p_{ubbm}(j)$, where $i, j \in \{1, 2, \dots, N\}$, can be easily

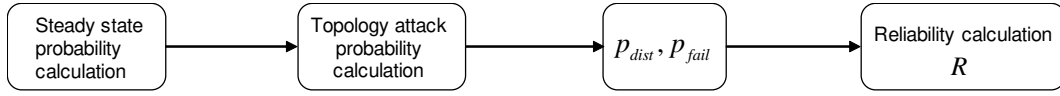


Fig. 3.12 The workflow of calculating reliability

obtained using steady-state analysis of the remaining SPN. Next, we determine the expected probability of constructing a topology attack p_{att} , which is given by

$$p_{att} = \sum_{i=1}^N \sum_{j=1}^N p_{ublm}(i) p_{ubbm}(j) \times \begin{cases} \frac{\sum_{l=1}^{\bar{i}} C_l^{\bar{i}} C_{\bar{j}-l}^{N-\bar{i}}}{C_{\bar{j}}^N}, & i + j \leq N \\ 1, & i + j > N, \end{cases} \quad (3.22)$$

where $\bar{i} = \min\{i, j\}$ and $\bar{j} = \max\{i, j\}$. $\frac{C_l^{\bar{i}} C_{\bar{j}-l}^{N-\bar{i}}}{C_{\bar{j}}^N}$ calculates the probability of l (out of \bar{i}) pairs of undetected compromised line meters and breaker monitors with a same host location. The dual summations then calculate the average probability of at least one pair of undetected compromised line meters and breaker monitors with a same host location, which is the optimistic condition as stated in the above section. After that, the steady-state probabilities p_{dist} and p_{fail} can then be determined by

$$p_{dist} = p_{att} \frac{\bar{N}_S}{N}, \quad (3.23)$$

and

$$p_{fail} = p_{att} \frac{N_S}{N}, \quad (3.24)$$

where \bar{N}_S and N_S are the number of non-spanning tree branches and spanning tree branches, respectively. Particularly, when the compromising rate is sufficiently large, there are always enough compromised bad devices in places P_UBLM and P_UBBM. As a result, $i + j > N$ in Eq. (3.22) is always satisfied so that we always have $p_{att} = 1$. In this case, the reliability

R is reduced as

$$R = \left(1 - \frac{\alpha * p_{att} * \bar{N}_S/N + \beta * p_{att} * N_S/N}{\alpha + \beta}\right)^k = \left(1 - \frac{\alpha * \bar{N}_S/N + \beta * N_S/N}{\alpha + \beta}\right)^k. \quad (3.25)$$

Note that k , describing the average number of attack events in the power system, is defined by

$$k = \sum_{x=1}^N x * p(x), \quad (3.26)$$

where

$$p(x) = \sum_{i=x}^N \sum_{j=x}^N p_{ublm}(i) p_{ubbm}(j) \times \begin{cases} \frac{C_x^i C_j^{N-i}}{C_j^{N-x}}, & i + j \leq N \\ 1, & i + j > N, \end{cases} \quad (3.27)$$

calculating the probability of x pairs of undetected compromised line meters and breaker monitors with a same host location. Then, k is determined by the expectation of the probability distribution.

3.5.2 Numerical Results

In the simulation experiments, we use MATLAB 2015a and PIPE 2 [82] as our simulators for transient and steady-state analysis, respectively. To facilitate comparison, we set the same levels of compromising rates, detection rates, and recovery rates for the two sensing devices, i.e., $\lambda_{clm} = \lambda_{cbm}$, $\lambda_{dblm} = \lambda_{dbbm}$, and $\lambda_{rlm} = \lambda_{rbm}$. The detection interval is set as 10 hours for the IDS. Note that in this section, we do not explicitly differentiate between conservative topology attacks from aggressive topology attacks, because the various compromising levels actually simulate all levels of the attack capability ranging from the conservative topology attacks all the way to aggressive topology attacks. The numerical results of all the simulations are shown as follows.

In Fig. 3.13, we plot the MTTD and MTTF of the IEEE 14-bus test system versus the compromising rate $\lambda_{clm} = \lambda_{cbm}$, for different detection rates $\lambda_{dblm} = \lambda_{dbbm}$. Note that,

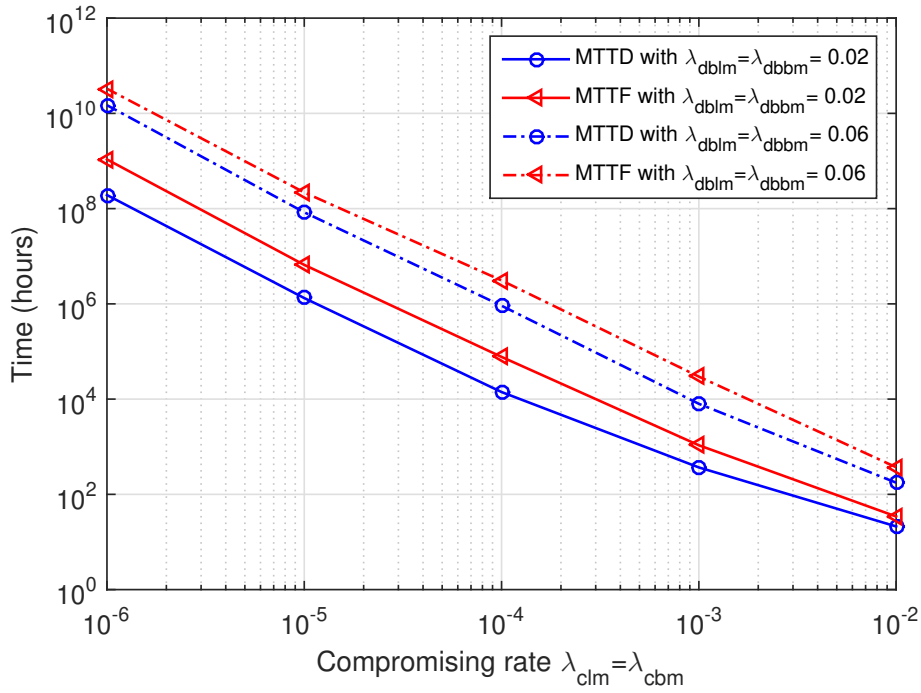


Fig. 3.13 The MTTD and MTTF vs. the compromising rate for IEEE 14-bus system ($\lambda_{rlm} = 0.08$)

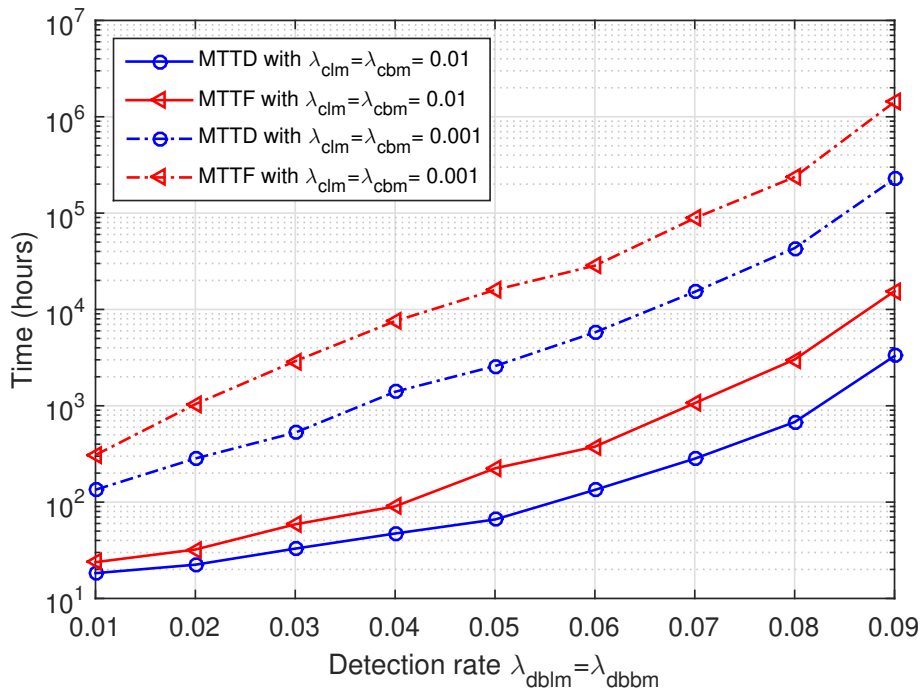


Fig. 3.14 The MTTD and MTTF vs. the detection rate for IEEE 14-bus system ($\lambda_{rlm} = 0.08$)

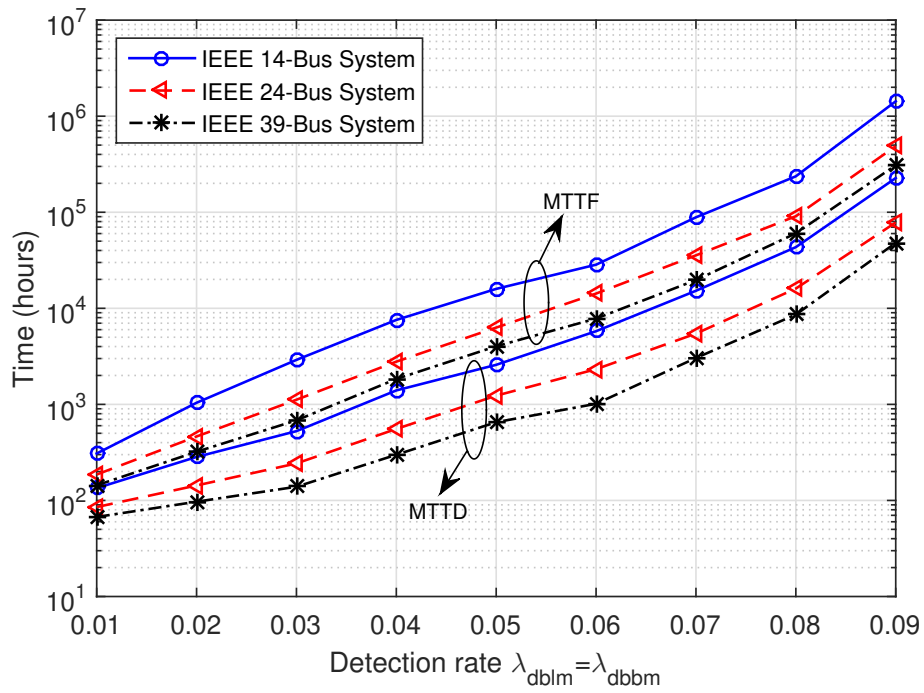


Fig. 3.15 The MTTD and MTTF vs. the compromising rate for various IEEE testing power systems ($\lambda_{dbl m} = 0.02$ and $\lambda_{rl m} = 0.08$)

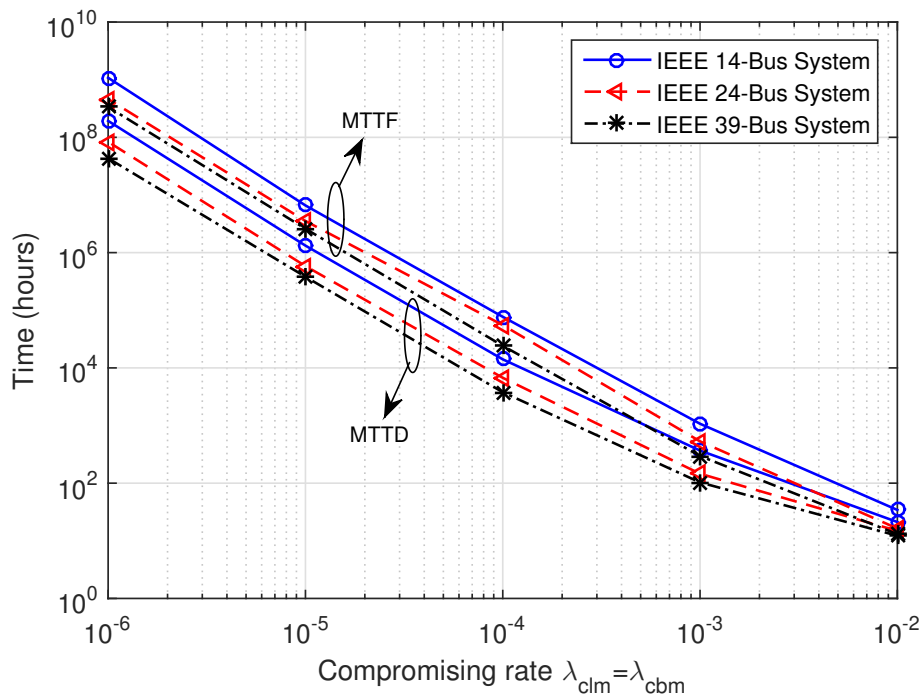


Fig. 3.16 The MTTD and MTTF vs. the detection rate for various IEEE testing power systems ($\lambda_{cl m} = 0.001$ and $\lambda_{rl m} = 0.08$)

a sufficient recovery rate of $\lambda_{rlm} = \lambda_{rbm} = 0.08$ is used as a constant parameter while analyzing the compromising rate and the detection rate. Figure 3.13 shows that when the compromising rate is relatively small, the power system has good operating conditions that both the MTDD and MTTF levels are significantly high. Larger compromising rates result in lower MTDD and MTTF levels as more sensing devices can be compromised by the adversary, increasing the probability to initiate a topology attack. In addition, we observe that MTTF is usually larger than MTDD. This is because when we have constrained knowledge and capability, it is much easier for an adversary to construct a relatively weak attack that causes system disturbances than construct a complicated strong attack to cause system failures. Also, higher levels of MTDD and MTTF can be obtained by increasing the detection rate, for example, from 0.02 to 0.06.

Figure 3.14 shows the MTDD and MTTF of the IEEE 14-bus test system versus the detection rate $\lambda_{dbl} = \lambda_{dbb}$, for different compromising rates $\lambda_{clm} = \lambda_{cbm}$. Clearly, when the detection rate increases, the MTDD and MTTF improve quickly because high detection rates will be more likely to detect compromised bad sensing devices and prevent them from launching topology attacks; thus leading to higher MTDD and MTTF levels. Similar observation is being made in Fig. 3.14, where lower compromising rates can also enhance the MTDD and MTTF levels.

In comparison to IEEE 14-bus test system, similar experiments pertaining to the MTDD and MTTF are also conducted in IEEE 24-bus and 39-bus test systems. The results are plotted in Figs. 3.15 and 3.16 for the compromising rate and detection rate, respectively. As is shown in Fig. 3.15, similar to all the three test systems, larger compromising rates correspond to lower MTDD and MTTF levels, while smaller compromising rates correspond to higher MTDD and MTTF levels. Most importantly, under the same level of compromising rate, detection rate, and the recovery rate, the IEEE 14-bus system has the highest levels of MTDD and MTTF, followed by the IEEE 24-bus system, and IEEE 39-bus system. This is because when the total number of sensing devices increases, the average number of sensing devices

that can be compromised per unit time also increases; thus, the probability of constructing a topology attack will increase, resulting in relatively lower levels of MTTD and MTTF. Figure 3.16 presents the parallel results that for all three test systems, MTTD and MTTF grow exponentially as the system detection rate increases, and the IEEE 14-bus system has the highest levels of MTTD and MTTF while the IEEE 39-bus system has the lowest.

After presenting the numerical results of transient analysis, we now present the steady-state analysis. Using PIPE 2 simulator, the steady-state probability distribution of the number of tokens in each place can be obtained. Figures 3.17 and 3.18 present the steady-state probability distribution of the number of tokens in place P_UBLM (also P_UBBM) under different compromising rates and detection rates, respectively, for IEEE 14-bus system. As observed from Fig. 3.17, when the compromising rate is relatively low where only a few good devices may be transferred to bad ones and most devices remain in good status, $\Pr\{\#P_UBLM = 0\}$ is significantly high with narrow probabilities for other number of tokens. $\Pr\{\#P_UBLM > 0\}$ can be increased by increasing the compromising rate. In contrast, as shown in Fig. 3.18, a smaller detection rate value of, for example, $\lambda_{dblm} = \lambda_{dbbm} = 0.015$ ($\lambda_{dblm} = \lambda_{dbbm} = 0.04$ in Fig. 3.17), may result in more bad compromised devices being undetected, and result in relatively larger $\Pr\{\#P_UBLM > 0\}$. Increasing the detection rate can improve $\Pr\{\#P_UBLM = 0\}$; thus, reducing the potential for an adversary to construct a topology attack.

With the obtained steady-state probability distribution, system *reliability* can be determined by Eq. (3.21). Figure 3.19 plots the system reliability of IEEE 14-bus system versus the compromising rate for various values of α/β . We used the α to β ratio in our experiments because based on Eq. (3.21), the definition of system reliability R can be written as $R = (1 - \frac{\alpha * p_{dist} + \beta * p_{fail}}{\alpha + \beta})^k = (1 - \frac{p_{dist}}{1 + \beta/\alpha} - \frac{p_{fail}}{\alpha/\beta + 1})^k$. Thus, it is more convenient to use the ratio α/β for analyzing the system reliability. As shown in Fig. 3.19, under the same level of compromising rate, detection rate, and recovery rate, higher values of α/β result in higher reliability. The reason is that, according to Eqs. (3.23) and (3.24), p_{fail} is usually greater than

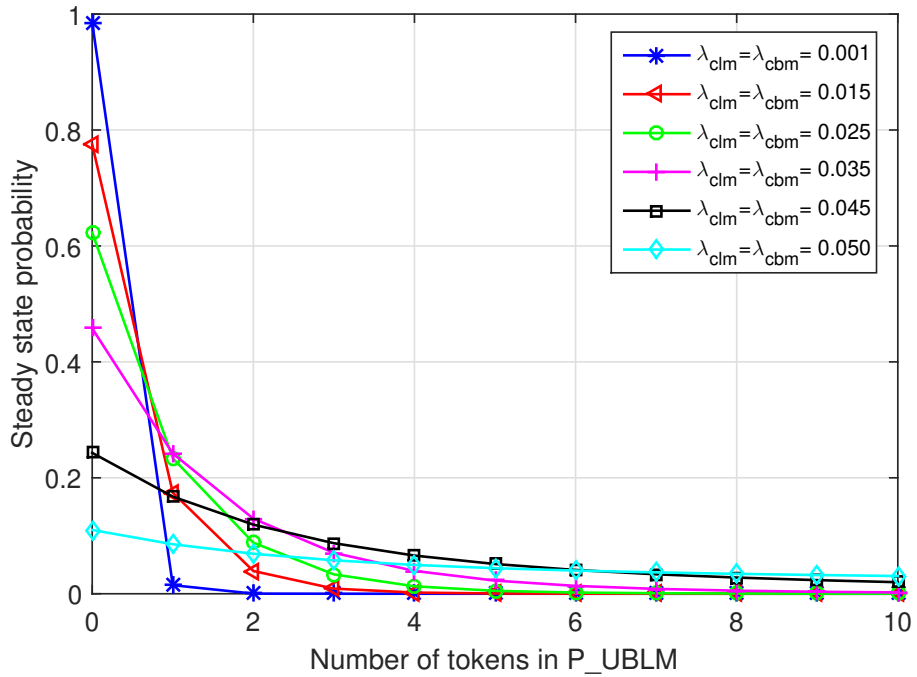


Fig. 3.17 Steady-state probability distribution of the number of tokens in place P_UBLM under different compromising rates for IEEE 14-bus system ($\lambda_{dblm} = 0.04$ and $\lambda_{rlm} = 0.08$)

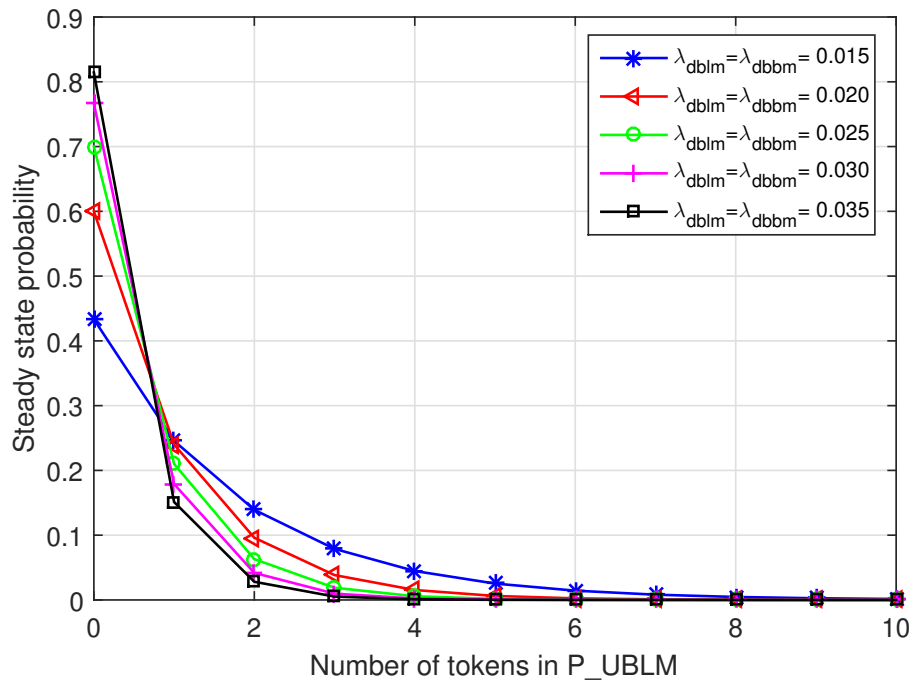


Fig. 3.18 Steady-state probability distribution of the number of tokens in place P_UBLM under different detection rates for IEEE 14-bus system ($\lambda_{clm} = 0.01$ and $\lambda_{rlm} = 0.08$)

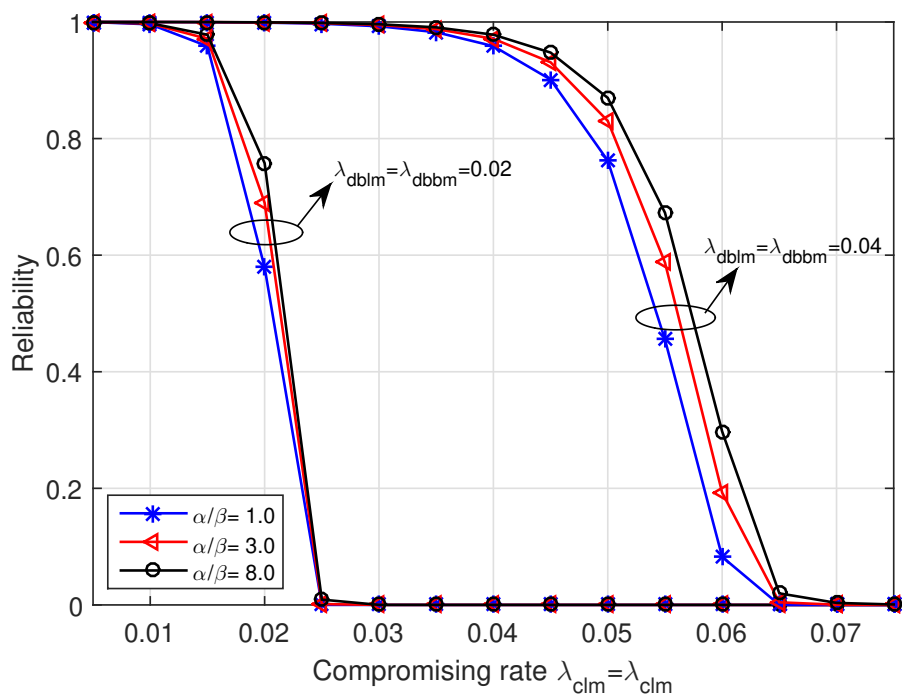


Fig. 3.19 System reliability of IEEE 14-bus system vs. the compromising rate ($\lambda_{rlm} = 0.08$)

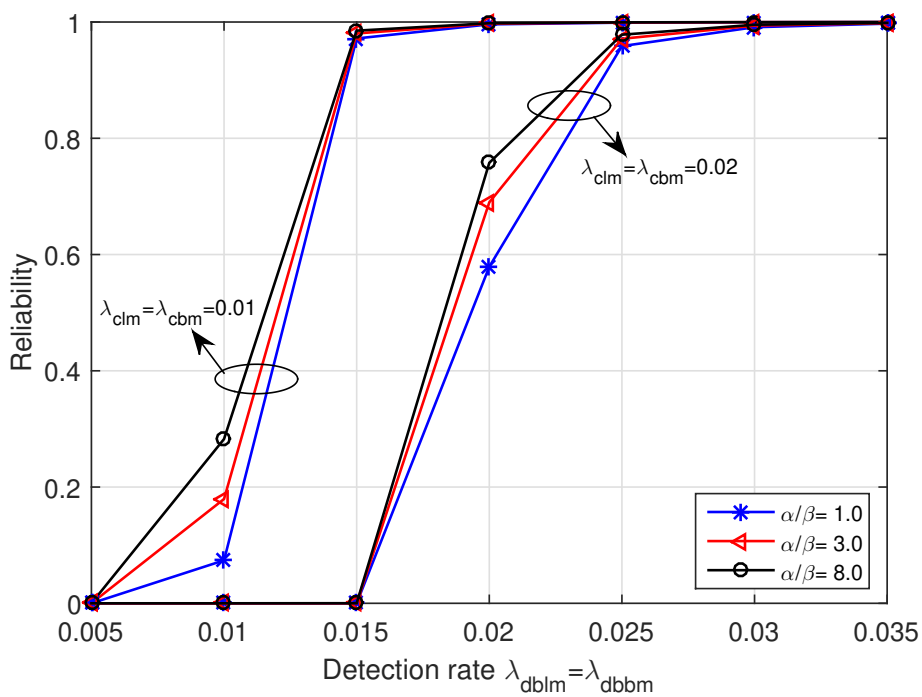


Fig. 3.20 System reliability of IEEE 14-bus system vs. the detection rate ($\lambda_{rlm} = 0.08$)

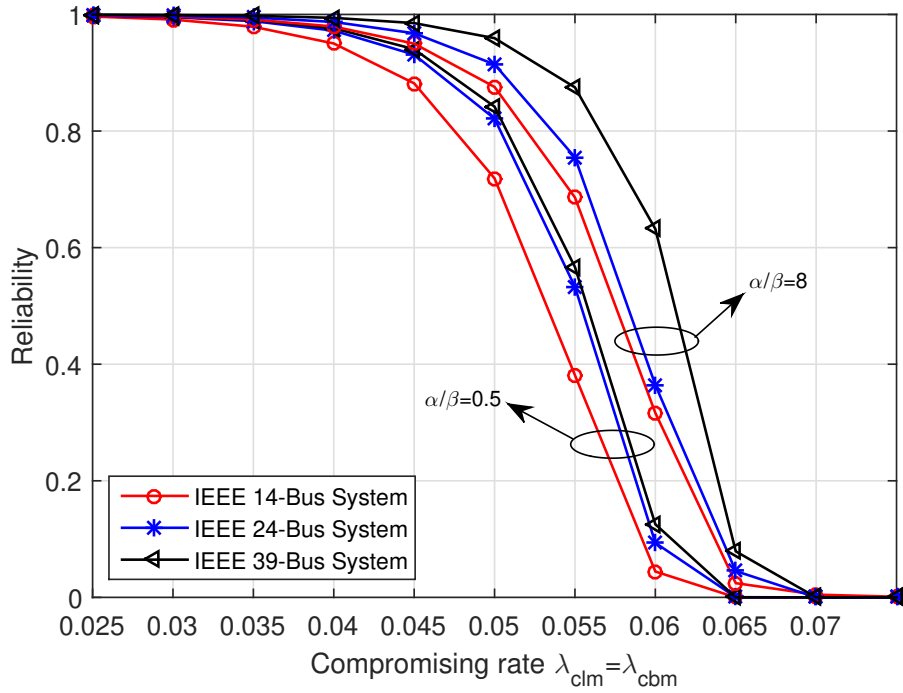


Fig. 3.21 System reliability vs. the compromising rate for various IEEE testing power systems ($\lambda_{dblm} = 0.04$ and $\lambda_{rlm} = 0.08$)

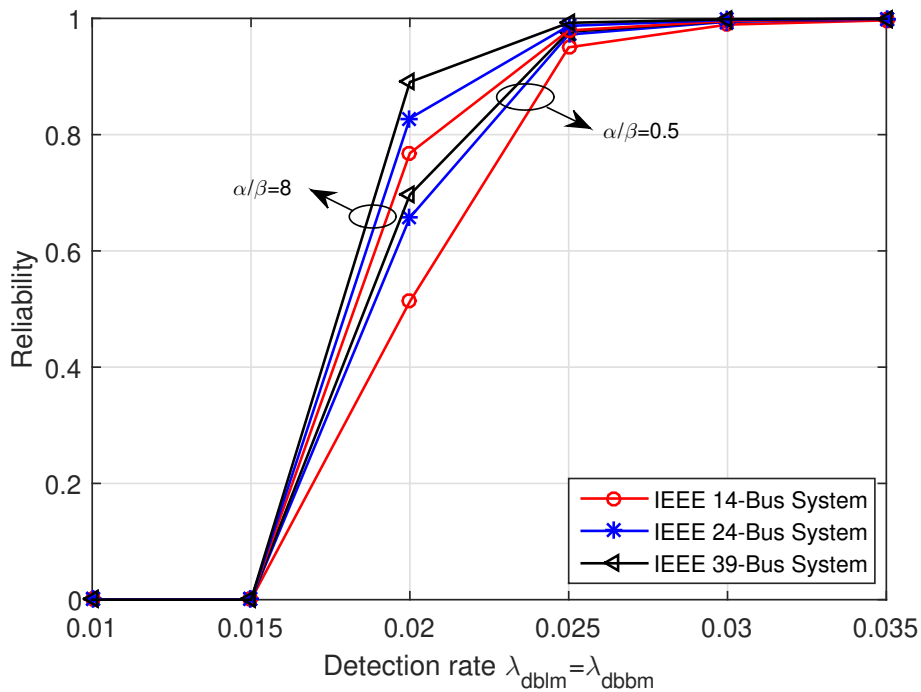


Fig. 3.22 System reliability vs. the compromising rate for various IEEE testing power systems ($\lambda_{cdm} = 0.02$ and $\lambda_{rlm} = 0.08$)

p_{dist} due to $N_S > \bar{N}_S$. Thus, increasing α/β assigns more weight to p_{dist} and less weight to p_{fail} , which as a result decreases the value of $\frac{\alpha*p_{dist}+\beta*p_{fail}}{\alpha+\beta}$. Therefore, the resulting value of reliability will be increased, and vice versa. More interestingly, the reliability decreases gradually from $R = 1$ at the origin and drops quickly to nearly zero as the growth of the compromising rate. This is because when the compromising rate is less than the detection rate, bad devices can be usually detected so that high reliability can be obtained, but when the compromising rate is large enough that exceeds the detection rate, there are a multitude of bad devices that cannot be successfully detected. In this case, there are always sufficient bad devices together in places P_UBLM and P_UBBM that can easily launch topology attacks, i.e., $p_{att} = 1$ holds all the time. In addition, the higher the compromising rate is, the larger the k is, which indicates the presence of multiple topology attacks and the large value of k will decrease the reliability in an exponential manner.

The relationship between the system reliability of IEEE 14-bus system and the detection rate for various values of α/β is plotted in Fig. 3.20. Likewise, this figure shows that under the same conditions, increasing the value of α/β can lead to higher system reliability, while decreasing it can lead to lower system reliability. In addition, the full simulation trace shows that system reliability experiences a slight growth from the beginning and eventually reaches a plateau at around $R = 1$ when the detection rate increases. This indicates that rise of the detection rate can slowly mitigate the number of undetected compromised devices, reduce the probability of initiating an attack, and further improve the system reliability.

In addition to the IEEE 14-bus system, simulation experiments concerning steady-state analysis for IEEE 24-bus and 39-bus systems are conducted as well. Figure 3.21 presents the system reliability of the three test systems against the compromising rate under different values of α/β . Similar results to the IEEE 14-bus system can be obtained for the IEEE 24-bus and 39-bus systems. Specifically, for all the three test systems, the reliability decreases gradually from $R = 1$ as the compromising rate increases, and drop quickly when the compromising rate exceeds the detection rate. In addition, we can also observe that the power

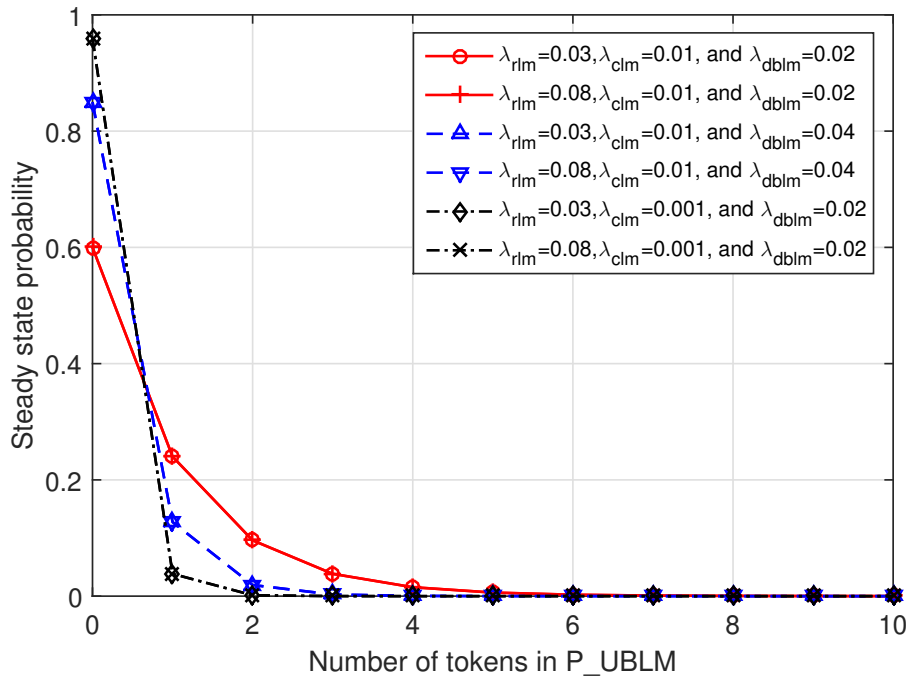


Fig. 3.23 Steady-state probability distribution of the number of tokens in place P_UBLM under different recovery rates for IEEE 14-bus system

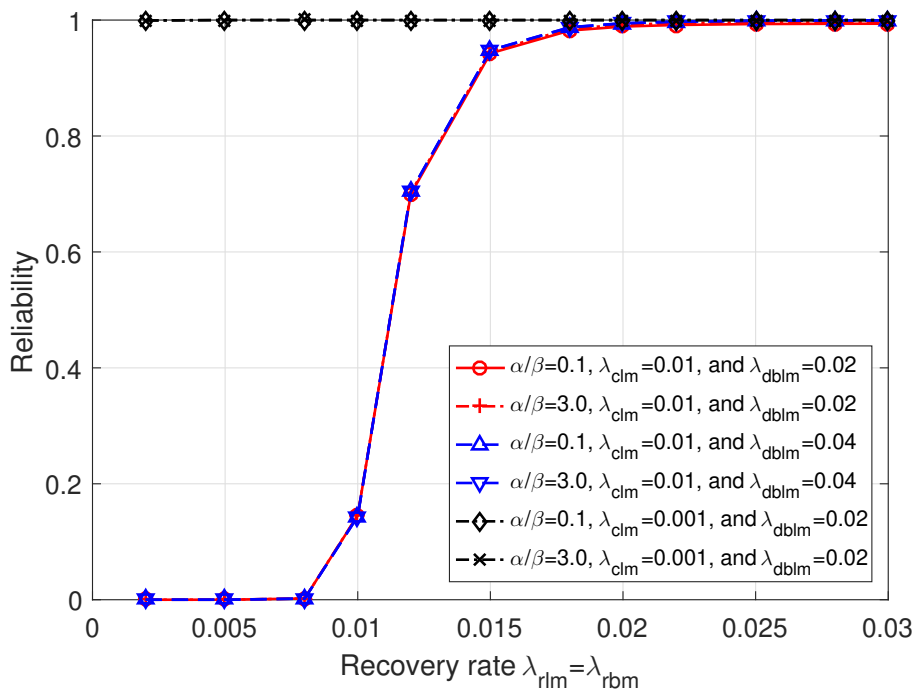


Fig. 3.24 System reliability vs. the recovery rate for IEEE 14-bus system

systems can have a high system reliability when α/β is set to be high. Furthermore, the numerical results show that the IEEE 39-bus system has generally the highest reliability, followed by the 24-bus system and 14-bus system the last. The reason is that a power system with more redundant branches and higher connection complexity may be more resilient to the attacks.

In Fig. 3.22, the reliability of different power systems against the detection rate is plotted. As we can see, the reliability stabilizes at nearly zero at the origin of each test system, due to insufficiency of the detection rate to identify compromised bad devices. When the detection rate increases gradually, the reliability begins to increase drastically and finally approaches to $R = 1$ when the detection rate is sufficient.

Finally, our simulation experiments focus on steady-state analysis against the recovery rate. The corresponding numerical results are summarized in Figs. 3.23 and 3.24. As observed in Fig. 3.23, three groups of comparative experiments show that under the same compromising rate and detection rate, the steady-state probability distribution of the number of tokens in place P_UBLM are the same for different recovery rates. This is because the number of detected bad devices arriving in place P_UBLM is determined by the compromising rate and detection rate collectively; thus, the same compromising rate and detection rate will result in the same distribution of tokens in steady states. Meanwhile, it can be seen that the recovery rate is of little impact on this distribution, i.e., different values of the recovery rate do not make any significant differences to the probability of topology attacks. The other curves once again show that the number of compromised devices in place P_UBLM can be further mitigated by increasing the detection rate or reducing the compromising rate.

In Fig. 3.24, we observe that from the two groups of red and blue curves, the system reliability experiences a sharp increase from 0 to 1 as the slight growth of the recovery rate. In such cases, system malfunctions occur when the recovery rate is not sufficient (e.g., $\lambda_{rlm} = \lambda_{rbm} < 0.02$ here) to recover the detected bad devices under a compromising rate of $\lambda_{clm} = \lambda_{cbm} = 0.01$ and a detection rate of either $\lambda_{dblm} = \lambda_{dbbm} = 0.02$ or 0.04. In

contrast, as the two black curves show that under a rather small compromising rate (i.e., $\lambda_{clm} = \lambda_{cbm} = 0.001$), a recovery rate of $\lambda_{rlm} = \lambda_{rbm} = 0.002$ is reasonably sufficient to recover all the detected bad devices. This leads to a full system reliability (i.e., $R = 1$) with no system malfunction occurs. In previous simulations related to the compromising rate and detection rate, we use a sufficient recovery rate of $\lambda_{rlm} = \lambda_{rbm} = 0.08$ to exclude the impacts on system reliability that insufficient recovery rate may cause. Compared to the compromising rate and detection rate, we observe that, in this figure, the recovery rate has relatively marginal impact on the system reliability as long as it can reach a basic acceptable level. More importantly, the recovery rate highly relies on the compromising rate and, then, the detection rate.

Such observations help inform future design of the system, so that system designers can devote more efforts to significant aspects, such as mitigating the compromising rate and improving the detection rate, rather than focusing too much on the recovery rate.

3.6 Summary

Smart grid cyber-physical systems will be increasingly deployed in the foreseeable future, and there are a number of research challenges that need to be addressed.

In this chapter, we developed an SPN-based analytical model for smart grid cyber-physical systems to assess and analyze the system reliability in the presence of both topology attacks and system countermeasures. We also demonstrated how to construct successful topology attacks in a smart grid. In our analytical SPN model, we took into account two types of sensing devices involving line meters and circuit breaker monitors, and two kinds of typical system countermeasures (i.e., IDSs and malfunction recovery techniques), we demonstrated how we can use several events to describe the system behaviors under these event triggers. Moreover, using the IEEE 14-bus as example, we proposed two algorithms pertaining to the construction of an MxST and identification of system disturbances, failures,

70 SPNTA: Reliability Analysis Under Topology Attacks: A Stochastic Petri Net Approach

and malfunctions. Finally, simulation experiments on the IEEE 14-bus, 24-bus, and 39-bus test systems and, correspondingly, both transient- and steady-state analysis demonstrated the utility and efficiency of our proposed analytical model. The findings (e.g., confirming that compromising rate and detection rate are of paramount significance to system reliability) will inform future system design.

Chapter 4

DHCD: Distributed Host-Based Collaborative Detection for FmDI Attacks

FmDI attacks are a crucial security threat to smart grid CPS, and could result in cataclysmic consequences to the entire power system. However, due to the high dependence on open information networking, countering FmDI attacks is challenging in smart grid CPS. Most existing solutions are based on state estimation at the highly centralized control center; thus, computationally expensive. In addition, these solutions generally do not provide a high level of security assurance, as evidenced by recent work that smart FmDI attackers with knowledge of system configurations can easily circumvent conventional state estimation-based false data detection mechanisms. In this paper, in order to address these challenges, a novel distributed host-based collaborative detection method is proposed. Specifically, in our approach, we use a conjunctive rule based majority voting algorithm to collaboratively detect false measurement data inserted by compromised PMUs. In addition, an innovative reputation system with an adaptive reputation updating algorithm is also designed to evaluate the overall running status of PMUs, by which FmDI attacks can be distinctly observed. Extensive simulation

experiments are conducted with real-time measurement data obtained from the PowerWorld simulator, and the numerical results fully demonstrate the effectiveness of our proposal.

4.1 Introduction

In recent times, a number of high-profile incidents targeting smart grid as well as other CPSs have been reported, e.g., Stuxnet [6], Conficker [83], and US drones hack [84]. Malicious attackers may attempt to falsify sensor measurements, embed fake control commands, delay or drop sensor readings or control commands [49, 21, 85–87]. FmDI attacks are increasingly recognized as a serious threat to smart grid CPS, and unsurprisingly, have been the focus of computer security researchers and industry practitioners. FmDI attacks and mitigation strategies on smart grid CPS have been also evolved over the years. Conventional FDD approaches are generally based on system state estimation [88–90]. For example, Merrill and Schweppe presented a bad data suppression estimator based on a non-quadratic cost function to improve the performance of static state estimation [88]. Handschin *et al.* presented a method to detect and identify the bad data and structural error problems, and improved bad data analysis (detection probability, and effects of bad data) [87]. Cutsem *et al.* also proposed an identification method attempting to alleviate some existing difficulties, such as multiple and interacting bad data [91].

However, Liu *et al.* in [28] showed that smart FmDI attackers armed with the knowledge of system configurations could easily bypass the traditional state estimation-based FDD schemes without detection. Consequently, existing FDD approaches may be ineffective against newer or emerging FmDI attacks. The major limitation of legacy FDD schemes is that they mainly focus on the inter-correlations among the measurement data (e.g., residuals and errors), rather than the malicious behaviors of meter devices, such as PMUs and smart meters. Furthermore, in existing literature FDD is generally performed by the power system's centralized CC, due to the demanding computational requirements [89, 90]. Although

a small number of hierarchical or distributed FDD schemes are designed to reduce the computation requirements at the CC [92, 93], most of them are still based on state estimation; thus, vulnerable to smart attackers. Another limitation of legacy FDD methods is that some prevailing countermeasures against cyber intrusion only aim to detect the “bad” data without further evaluating the true running status of the meter devices that might already be compromised by malicious attackers [88, 92, 87]. These undetected hidden attackers can continue to launch or improve their attacks subsequently. Therefore, countering against FmDI attacks in smart grid CPS remains a research challenge, and one that we seek to address in this paper.

Thus, we propose a distributed host-based collaborative detection (DHCD) method based on rule specifications, rather than state estimation. DHCD can not only reduce the computational burden of the CC, but also achieve fast FDD and the capability to evaluate the running status of meter devices. Specifically, in our method, each PMU is assigned a trusted host monitor (HM) serving as the distributed local false data detector. Based on a set of pre-defined rule specifications, the monitors determine the anomalous levels of measurement data collected by their supervised PMUs. Then, by sharing and comparing the anomalous levels of the measurement data collected by the neighboring interconnected PMUs, these interconnected monitors collaboratively make a decision based on the majority voting algorithm to determine whether their own measurement data is falsified. To evaluate the overall running status of the PMUs, a reputation system with an adaptive reputation updating (ARU) algorithm is designed, where a malfunction PMU can be easily identified. The contributions of our work are summarized as follows:

- We develop a DHCD method to detect FmDI attacks in smart grid CPS based on rule specifications, which can be used to effectively mitigate smart FmDI attacks.
- Our method can not only achieve fast and high accuracy of FDD, but also allow the identification of compromised PMUs using our designed reputation system.

- Our distributed detection method will “displace” the computational burden of the CC by delegating FDD tasks to the local monitors.

The remainder of this paper is organized as follows. Section 4.2 presents the system model, the threat model, and our design goals. The DHCD method is detailed in Section 4.3, followed by the performance evaluation in Section 4.4. Section 4.5 concludes the paper with future research directions.

4.2 Models and Design Goals

In this section, we introduce the system model, the threat model, and our design goals.

4.2.1 System Model

A smart grid CPS is a fully automated system capable of achieving self-healing, cost reduction, improved reliability and efficiency. These promising benefits are intensively grounded on the WAMS, as it can provide high-level observability and controllability in power system operations [18, 94, 95]. Thus, in this paper, we consider the hierarchical WAMS as our system model.

As shown in Fig. 4.1, WAMS is an integrated system consisting of PMUs, PDCs, heterogeneous communication networks, and a CC. Specifically, PMUs, located at the substations of the power generation and transmission system, are capable of measuring the real-time status of the power system. For example, the real-time amplitude and phase angle of voltage at the bus, of current on the transmission line, and of the power at each branch, can be measured by the PMUs. These measurement data are then periodically transmitted to the PDCs, usually in 50Hz, through the LAN. Then, the aggregated data at the PDCs are delivered to the CC via the WAN for further data analysis, such as state estimation, event diagnostics, and contingency analysis.

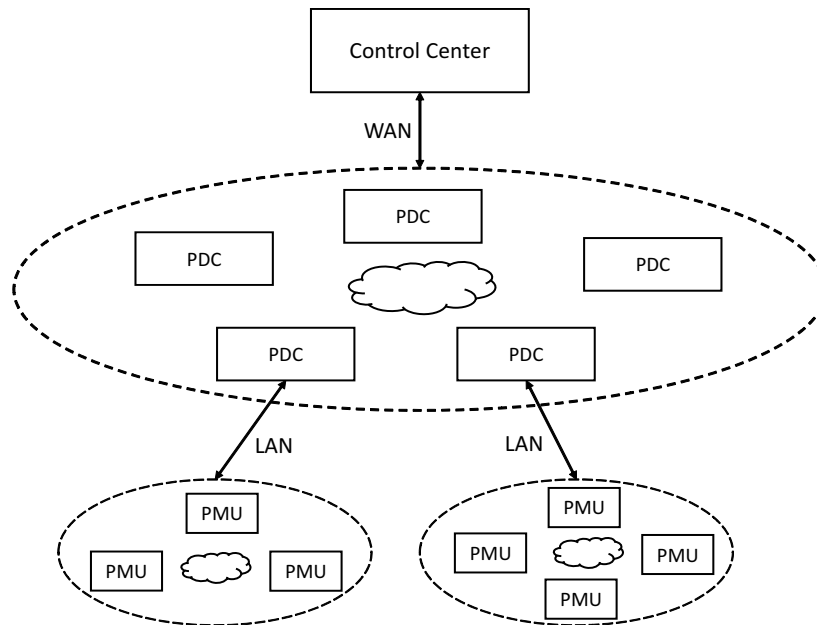


Fig. 4.1 The hierarchical architecture of WAMS

4.2.2 Threat Model

The real-time data provided by PMUs serve as the basis for automated, efficient, and reliable system control. However, adversaries seeking to intervene or manipulate system operations can attempt to inject false measurement data through compromised PMUs. Successful FmDI attack may compromise the above-mentioned promising functionalities or even jeopardize the system operations.

In our threat model, we consider that PMUs in the WAMS can be compromised by FmDI attackers (e.g., rewriting the program settings, or stealing the secret information for data communication). Note that, in smart grid CPS, a single piece of false measurement data may not have significant impact on system operations, because the system is capable of correcting trivial faults or mistakes. However, the system may not be able to auto-correct in the event that consecutive false measurement data are received; consequently, resulting in system failures. As such, to successfully launch an FmDI attack in practice, attackers usually recklessly and persistently inject false measurement data once they have an opportunity. This is the behavior pattern of FmDI attackers we consider in the threat model.

4.2.3 Design Goals

Based on the aforementioned system model and threat model, our design goals are to develop an accurate, efficient, and scalable FDD method in smart grid CPS. Specifically, the following specific objectives should be achieved.

Accuracy: The devised method is able to effectively detect smart FmDI attacks, achieving both high detection rate and low false alarm rate.

Efficiency: The detection method should not introduce extensive computational burden to the system, particularly to the CC inherent in traditional FDD schemes. In other words, a light-weight detection scheme is expected.

Scalability: The smart grid CPS needs to be scalable (similar to a cloud system) by allowing new devices to be added, etc, without incurring expensive (financial) costs.

4.3 Proposed DHCD Method

In this section, we present the proposed DHCD method, which is composed of two steps (subsections): collaborative FDD and determination of compromised PMU. In the first step, we employ a set of rule specifications to identify anomalous measurement data reported by the PMU. Then, in the second step, we devise a reputation system with an ARU algorithm to monitor and assess PMUs' overall behaviors in order to further detect compromised PMU.

4.3.1 Collaborative FDD

In normal operational circumstances, the power grid operates in a stable status. In other words, all state variables vary in a mutual balanced manner according to Kirchhoff's law, demand-response constraints, etc. As such, any change of a variable state on one bus or transmission line, resulting from either the normal demand variation or system faults, would lead to corresponding state changes of the same and/or other variables on interconnected buses or transmission lines. For example, as shown in Fig. 4.2, the contouring maps with

comparison are plotted, which describe the distribution of the current amplitude on each transmission line (a) before and (b) after an open circuit event on transmission line from Bus 16 to Bus 17. As shown in Fig. 4.2(b), after the occurrence of this open circuit event, the current amplitude values near Line 16 to 17 shift. The closer to this line, the more the value changes.

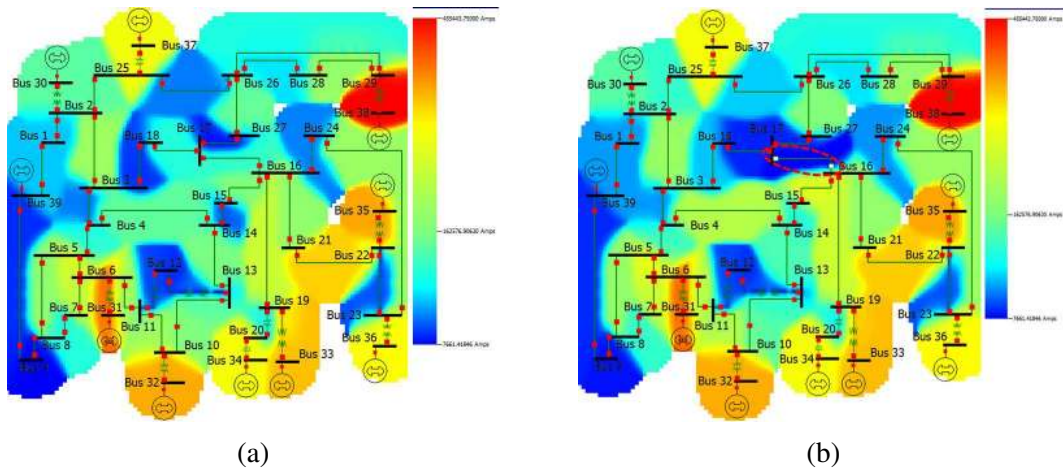


Fig. 4.2 Comparison of contouring maps describing the distribution of current amplitude on transmission lines: (a) before open circuit and (b) after open circuit on line from Bus 16 to 17 (marked by a red dashed elliptical circle) in IEEE-39 bus system.

In contrast, if only some changes of variable states occur on one bus, without a corresponding shift in the parallel variables of interconnected buses, such changes can be regarded as anomalous. These anomalies may originate from either malfunction PMU devices or malicious activities due to compromised PMUs. In this paper, we only consider possible malicious activities rather than device malfunction, as there are many existing approaches to address issues relating to device malfunction. Based on the inter-correlations of power systems, we design a collaborative detection method to detect anomalous measurement data reported by PMUs [96, 97].

Normal Rule Specifications

When power system is under normal operation, all state variables must naturally follow some constraints and hold some properties. Let us take active power P as an example, which should obey the following rules:

- $P_{min} < P^t < P_{max}$: P at any time under stable status must vary within an experienced range $[P_{min}, P_{max}]$.
- $|P^t - P^{t-1}| < P_{\Delta}$: The variation of P within one time interval should be less than an experienced threshold P_{Δ} .
- $|P_{in}^t - P_{out}^t| < P_{loss}$: The difference of P flowing into a bus and flowing out the bus ought to be less than an experienced power loss threshold P_{loss} .
- Other more complicated rules.

As such, we pre-define some useful but example rule specifications as listed in Table 4.1 that PMUs have to coincide with in the stable status. These rule specifications serve as the basis of our method to identify the anomalous measurement data (for convenience, the superscript t is omitted).

Table 4.1 Rules specifications for PMUs in stable status

Index	Variable	Rule description
1	Active Power Angle	$\Delta\delta < \delta_{\Delta}$
2	(Phase A) Voltage Amplitude	$\Delta V < V_{\Delta}$
3	Load Mvar	$\Delta L_{Mvar} < L_{Mvar\Delta}$
4	Load MW	$\Delta L_{MW} < L_{MW\Delta}$

To represent the results of whether the rule specifications have been violated, we employ a binary system, where “0” denotes that the measurement data of one variable follows the relevant rule specification and “1” indicates a violation. A binary sequence with length E (E is the number of rule specifications, and here E is 4) is utilized to represent the conjunctive

results pertaining to the entire measurement data. For instance, “1001” denotes that both rules 1 and 4 are violated. A non violation of the conjunctive four rule specifications is represented by “0000”, which is our *baseline* of PMUs’ behaviors.

In order to assess to what extent each piece of measurement data is anomalous, we introduce a normalized Euclidean distance strategy to determine the *anomalous level* l^t , which is shown as follows:

$$l^t = D_0(seq^t, seq_0), \quad (4.1)$$

where seq^t is the binary sequence representing the conjunctive results of measurement data at time t , while $seq_0 = “0000”$ is the baseline. D_0 is the normalized Euclidean distance of the two sequences seq^t and seq_0 . Euclidean distance is the square root of the sum of results that are different between two sequences. For example, the Euclidean distance between sequence “1001” and the *baseline* “0000” is $\sqrt{1^2 + 0 + 0 + 1^2} \approx 1.414$. Then, the *anomalous level* l is computed by the normalized distance, i.e., $1.414/\sqrt{1^2 + 1^2 + 1^2 + 1^2} \approx 0.707$.

FDD algorithm with Iterative Majority Voting

Figure 4.3 shows the distributed host-based collaborative FDD system, where each HM (host monitor) is responsible for monitoring and assessing the behaviors of its administrated PMU. Let $\mathcal{M} = \{M_1, M_2, \dots, M_N\}$ denote the set of monitors and $\mathcal{U} = \{U_1, U_2, \dots, U_N\}$ the set of PMUs, where N is the total number of HMs or PMUs. HMs communicate among each other following the connection pattern of the PMUs, which means each HM only communicates with HMs that their monitored PMUs have interconnection relations. Note that these host monitors are trusted entities for monitoring PMUs. It is designed that host monitors and the networks between them are equipped with high-level security mechanisms to ensure their trustworthiness.

As stated above, we utilize the inter-correlations between the state variables to build our detection method. Algorithm 4.1 outlines the FDD algorithm with iterative majority voting

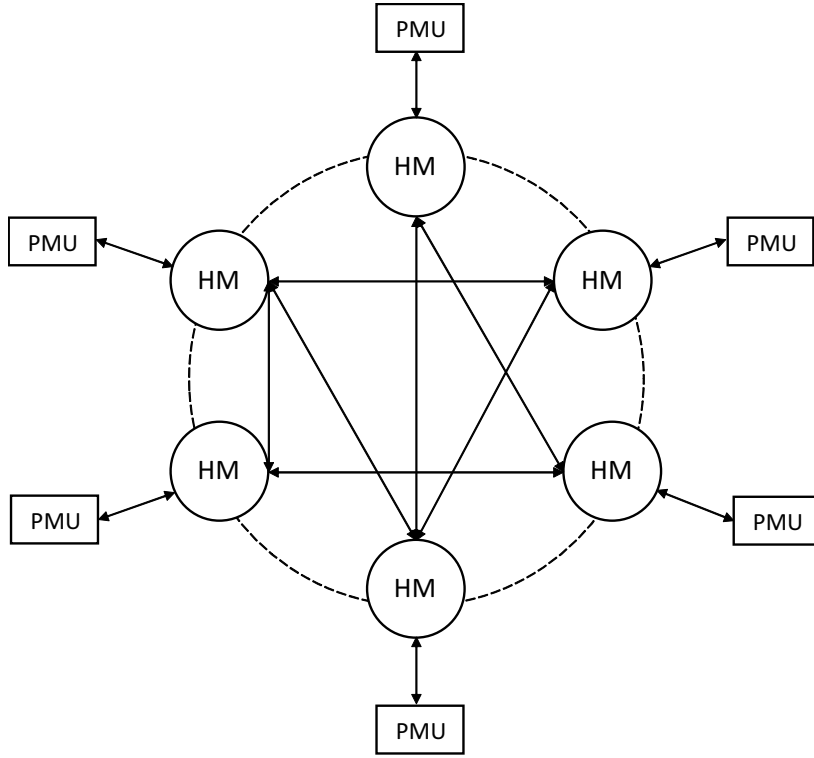


Fig. 4.3 The distributed host-based collaborative FDD system

process. Concretely, set \mathcal{M} is initialized as $\mathcal{M} = \{M_1, M_2, \dots, M_N\}$, and a flag variable *repeat_flag* as '0'. Note that *repeat_flag* = '0' indicates that the procedure does not need to be repeated, while *repeat_flag* = '1' indicates the need to repeat the procedure. Next, each monitor $M_i \in \mathcal{M}$ determines the conjunctive result R_i^t of current piece of measurement data, and broadcasts the result to neighbouring connected monitors $\mathcal{M}_i = \{M_j | M_j \sim M_i\}$. An example is shown in Fig. 4.4.

Then, M_i launches the false data identification process. If there is no bit "1" in the result R_i^t , then no false data is detected. Otherwise, M_i needs to determine how many of its connected monitors have a bit "1" in their conjunctive results R_j^t . If more than or equal to half of the connected monitors have a bit "1" at the same position in R_j^t , M_i concludes that U_i has reported a piece of false measurement data; otherwise, R_i^t is tentatively considered suspicious. After all $M_i \in \mathcal{M}$ have concluded the first procedure, the termination criterion is determined. If *repeat_flag* == '1', this procedure is repeated to further identify the false

$M_1 :$	Rule_Index	R1	R2	R3	R4
	Rule_Result	0	0	0	0

$M_2 :$	Rule_Index	R1	R2	R3	R4
	Rule_Result	0	0	0	1

•
•
•

$M_N :$	Rule_Index	R1	R2	R3	R4
	Rule_Result	0	1	0	0

Fig. 4.4 An example of the conjunctive results transmitted between HMs

data; otherwise, the procedure goes to the end. Note that the time complexity of Algorithm 4.1 is $\mathcal{O}(N \times L)$, where $L = \frac{1}{2} \sum_{i=1}^N |\mathcal{M}_i|$, the number of branches in a power grid.

4.3.2 Determination of Compromised PMU

FDD step is a critical process to detect false data, but it is not sufficient to identify compromised PMUs. Therefore, in the second step, we employ a reputation-based algorithm to monitor and assess the PMUs' overall behaviors over a period of time, which allows us to identify compromised PMUs if their reputation level drops below an acceptable threshold [98, 99].

Specifically, in this subsection, we first model the probability distribution of the anomalous level of measurement data with a Beta distribution. Then, we estimate its two shape parameters α and β using maximum likelihood estimation (MLE) and Newton-Raphson method. Then, a detailed description of an adaptive reputation updating (ARU) algorithm is presented.

Algorithm 4.1 FDD Algorithm

```

1: initialization:  $\mathcal{M} = \{M_1, M_2, \dots, M_N\}$ ,  $Upperbound = 5$ ,  $Iteration = 0$ ,  $repeat\_flag =$ 
   '0'
2: procedure
3:   for each monitor  $M_i \in \mathcal{M}$  do
4:     (1). determines the conjunctive result  $R_i^t$  of current piece of measurement data.
5:     (2). broadcasts the result  $R_i^t$  to the neighbouring connected monitors  $\mathcal{M}_i = \{M_j | M_j \sim$ 
       $M_i\}$ .
6:     (3). identifies false data:
7:     if there is no bit "1" in the result  $R_i^t$  then
8:       output: no false data detected.
9:     else if more than or equal to half of the monitors in  $\mathcal{M}_i$  hold bit "0" at the same position
      in the result  $R_j^t$  then
10:      (a). output: false data detected.
11:      (b). removes  $M_i$  from  $\mathcal{M}$  and its connections with other monitors.
12:     else
13:      (a). keeps  $R_i^t$  as suspicious result.
14:      (b).  $repeat\_flag = '1'$ .
15:     end if
16:   end for
17:   (4). judges the termination criteria:
18:   if  $repeat\_flag == '1'$  and  $Iteration < Upperbound$  then
19:     (a). repeats procedure.
20:     (b).  $Iteration = Iteration + 1$ .
21:   else
22:     ends the procedure.
23:   end if
24: end procedure

```

Probability Distribution of Anomalous Level

Let random variable X be the anomalous level of a piece of measurement data, where X can either be 0 or 1 and it is determined by the normalized Euclidean distance (see section 4.3.1). Particularly, $X = 0$ represents compliance of the rule specifications, while $X = 1$ represents a violation. Here, to determine the exact distribution of the probabilities of different anomalous level and its future values, we model the random variable X using a $Beta(\alpha, \beta)$ distribution. Beta distribution family can represent a collection of probability distributions, and can be used to depict a prior distribution of an unknown distribution with only a series of collected observations.

The probability density function of a Beta distribution is

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad (4.2)$$

where α and β are the two shape parameters. The mean value of a Beta distribution is

$$\mu = E[X] = \int_0^1 x \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\alpha}{\alpha + \beta}. \quad (4.3)$$

To obtain the exact distribution of X , we estimate the parameters α and β using a well-known method MLE. We suppose that the n independent and identically distributed observations $\{x_1, x_2, \dots, x_n\}$ are from an unknown distribution with pdf $f_0(\cdot|\vec{\theta})$, $\vec{\theta}$ is a vector of parameters. As for our model, the Beta distribution, $\vec{\theta} = [\alpha \ \beta]$. By using MLE, we formulate the joint density probability function of these n independent and identically distributed observations $\{x_1, x_2, \dots, x_n\}$ as

$$f(x_1, x_2, \dots, x_n | \alpha, \beta) = \prod_{i=1}^n f(x_i | \alpha, \beta). \quad (4.4)$$

Now we look at this equation from a different perspective by fixing the observed samples $\{x_1, x_2, \dots, x_n\}$ of this function, then α, β are the variables of the function that we call the likelihood:

$$\mathcal{L}(\alpha, \beta | x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i | \alpha, \beta). \quad (4.5)$$

In most cases, it is easier to work with the natural logarithm of the likelihood function. We rewrite it as

$$\begin{aligned}
\ln \mathcal{L}(\alpha, \beta \mid x_1, x_2, \dots, x_n) &= \ln \prod_{i=1}^n f(x_i \mid \alpha, \beta) \\
&= \sum_{i=1}^n \ln \left\{ \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1 - x_i)^{\beta-1} \right\} \\
&= n \ln \Gamma(\alpha + \beta) - n[\ln \Gamma(\alpha) + \ln \Gamma(\beta)] + (\alpha - 1) \sum_{i=1}^n \ln x_i + (\beta - 1) \sum_{i=1}^n \ln(1 - x_i).
\end{aligned} \tag{4.6}$$

Then, we have to find the optimal values of α and β that maximize $\ln \mathcal{L}(\alpha, \beta \mid x_1, \dots, x_n)$. Since logarithm is a strictly monotonically increasing function, the maximum value, if it exists, could be calculated by

$$\begin{cases} \frac{\partial \ln \mathcal{L}}{\partial \alpha} = 0, \\ \frac{\partial \ln \mathcal{L}}{\partial \beta} = 0. \end{cases} \tag{4.7}$$

That is

$$g_1(\alpha, \beta) = \psi(\alpha) - \psi(\alpha + \beta) - \frac{1}{n} \sum_{i=1}^n \ln x_i = 0, \tag{4.8}$$

$$g_2(\alpha, \beta) = \psi(\beta) - \psi(\alpha + \beta) - \frac{1}{n} \sum_{i=1}^n \ln(1 - x_i) = 0. \tag{4.9}$$

where $\psi(x)$ is the digamma function defined as

$$\psi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}. \tag{4.10}$$

There is no closed-form solution to Equations (4.8) and (4.9), so we use the Newton-Raphson method to find the approximate roots. The parameters $\vec{\theta} = [\hat{\alpha} \ \hat{\beta}]$ can be iteratively estimated by [100]

$$\vec{\theta}_{i+1} = \vec{\theta}_i - \frac{g(\vec{\theta}_i)}{\vec{J}_g(\vec{\theta}_i)}, \tag{4.11}$$

where $\vec{g} = [g_1 \ g_2]$, and $J_g(\hat{\theta}_i)$ is a 2×2 Jacobian matrix defined over the function vector $g(\hat{\theta}_i)$ defined as

$$\begin{bmatrix} \frac{dg_1}{d\alpha} & \frac{dg_1}{d\beta} \\ \frac{dg_2}{d\alpha} & \frac{dg_2}{d\beta} \end{bmatrix}, \quad (4.12)$$

with

$$\frac{dg_1}{d\alpha} = \psi'(\alpha) - \psi'(\alpha + \beta), \quad (4.13)$$

$$\frac{dg_1}{d\beta} = \frac{dg_2}{d\alpha} = -\psi'(\alpha + \beta), \quad (4.14)$$

$$\frac{dg_2}{d\beta} = \psi'(\beta) - \psi'(\alpha + \beta). \quad (4.15)$$

This Newton-Raphson method converges when the estimates of $\hat{\theta}$ and $\hat{\beta}$ change by less than an acceptable threshold with each successive iteration.

ARU Algorithm

With the exact probability distribution of the anomalous level, we can obtain its expectation value μ , which is the best indicator of the overall performance of the PMUs over the observation period. Here, we define the history reputation level of a PMU as

$$T = 1 - \mu = \frac{\beta}{\alpha + \beta}. \quad (4.16)$$

While, a dependable reputation system should be able to adaptively adjust the reputation values according to dynamic behavioral changes [101]. Thus, in this paper, we incorporate the history reputation level and the subsequent behavior fluctuations of PMUs to assess their real-time reputation levels. In addition, adaptive parameters are used to allow different impacts due to the reputation levels with different behavior observations. The real-time

reputation level of a PMU is then defined as

$$\begin{aligned} T^t &= \omega \cdot T_h + (1 - \omega) \cdot T_u^t \\ &= \omega \cdot \frac{\beta}{\alpha + \beta} + (1 - \omega) \cdot \frac{\lambda_g \cdot N_g^t + 1}{\lambda_g \cdot N_g^t + \lambda_b^t \cdot N_b^t + 1}, \end{aligned} \quad (4.17)$$

where T_h is the history reputation level of a PMU, and T_u^t is the updating reputation level at time instant t . ω is the weight assigned for the history reputation level to evaluate the importance of history experience to the real-time reputation level, while $1 - \omega$ is for the updating reputation level to evaluate the impacts of recent performance to the real-time reputation level [102]. N_g^t and N_b^t denote the cumulative number of observations regarding “good” data (not false data) and “bad” data (false data) of a PMU, respectively. Correspondingly, λ_g and λ_b^t are designed as the impact factors for “good” data and “bad” data. It is natural that, from the social perspective, one needs to spend a longer period of time performing successive good behaviors to establish a high reputation level, yet only a few bad behaviors would adversely affect the reputation built over time [103]. As such, we penalize the PMUs when “bad” data are observed. In our algorithm, λ_b^t is designed relatively larger than λ_g , and λ_b^t will be increased if successive “bad” data are observed to amplify the impacts.

Algorithm 4.2 presents the ARU procedure, where S_b^t denotes the number of successive observations of “bad” data. They increment by 1 when corresponding behavior occurs. If successive “bad” data is observed, the corresponding impact factor λ_b^t will be increased by $\lambda_b^{t-1} \cdot (e^\tau - 1)$, otherwise, the counter for successive “bad” observations S_b^t will be reset to 0 and the impact factor λ_b^t remains unchanged. Here, τ is initialized as a small value (e.g., 0.0001) in our experiments, and can be adjusted according to different application environments. Note that the time complexity of Algorithm 4.2 is $\mathcal{O}(1)$.

Algorithm 4.2 Adaptive Reputation Updating Algorithm

```

1: procedure
2:   Input:  $N_g^{t-1}, N_b^{t-1}, \lambda_g, \lambda_b^{t-1}, S_b^{t-1}, \tau$ 
3:   if the judgement result of current data is "good" then
4:      $N_g^t \leftarrow N_g^{t-1} + 1;$ 
5:      $S_b^t \leftarrow 0;$ 
6:   else
7:      $N_b^t \leftarrow N_b^{t-1} + 1;$ 
8:      $S_b^t \leftarrow S_b^{t-1} + 1;$ 
9:     if  $S_b^t > 1$  then
10:       $\lambda_b^t = \lambda_b^{t-1} \cdot e^\tau;$ 
11:     end if
12:   end if
13:   Compute updating reputation level by:
14:    $T_u^t = \frac{\lambda_g \cdot N_g^t + 1}{\lambda_g \cdot N_g^t + \lambda_b^t \cdot N_b^t + 1},$ 
15:   and the overall reputation level by:
16:    $T^t = \omega \cdot \frac{\beta}{\alpha + \beta} + (1 - \omega) \cdot \frac{\lambda_g \cdot N_g^t + 1}{\lambda_g \cdot N_g^t + \lambda_b^t \cdot N_b^t + 1}.$ 
17:   output:  $T^t.$ 
18: end procedure

```

With the real-time reputation level of each PMU, it is easy to identify the compromised PMU by testing the following binary hypothesis:

$$\begin{cases} \mathbf{H}_0: \text{PMU } U_j \text{ is compromised,} & \text{if } T_j^t < D_{th} \\ \mathbf{H}_1: \text{PMU } U_j \text{ is not compromised,} & \text{otherwise.} \end{cases} \quad (4.18)$$

where D_{th} is an acceptable detection threshold. This hypothesis is tested once the reputation level is updated in order to ensure real-time detection.

4.4 Performance Evaluation

In this section, we present a set of simulation experiments and the results to demonstrate the efficacy of our proposed DHCD method, including the collaborative FDD process and determination of compromised PMU process. Figure 4.5 shows the IEEE 39-bus power system that is used as a benchmark system in our simulation experiments. IEEE 39-bus

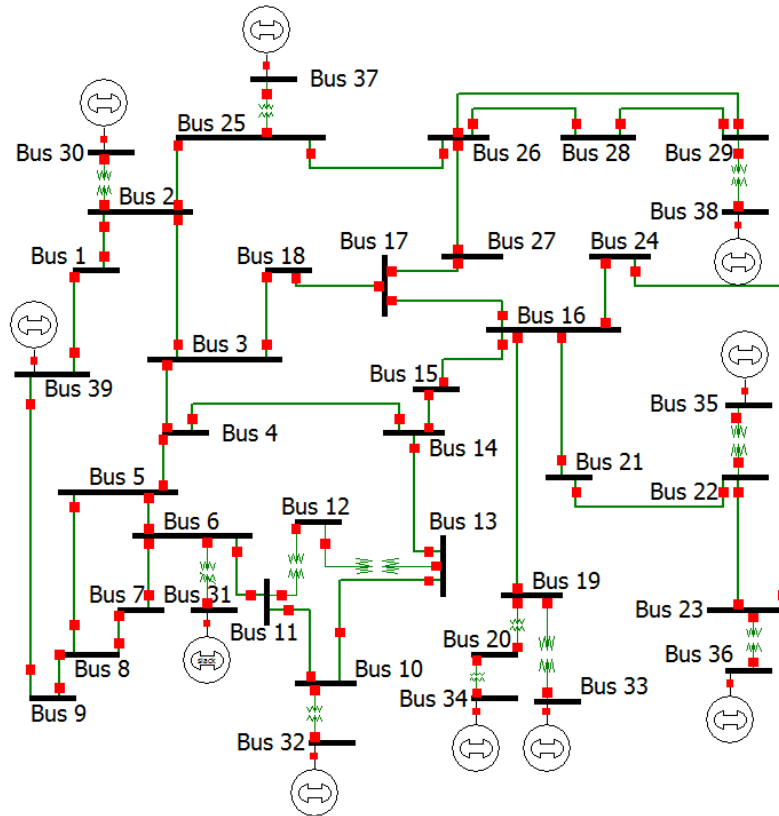


Fig. 4.5 IEEE 39-bus power system

power system is a well-known New England power system with 10 generators, 39 buses, and 46 transmission lines, which is commonly used as a benchmark system to test and verify new schemes [49, 21, 104]. Combined with the PowerWorld simulator [105], the power system can provide real-time, accurate and precise state information of the power system. Our experiments are conducted using the PowerWorld simulator on an IEEE standard 39-bus power system, where a number of scenarios are simulated and corresponding real-time measurement data from PMUs are collected. These data are then used to evaluate our proposed DHCD method in MATLAB. The key parameters are summarized in Table 4.2.

4.4.1 Efficacy of FDD Algorithm

In this section, we simulate two groups of simulation experiments. The first group shows that only one piece of the four rule specifications is violated (with a single “1” in R_j^t). In contrast,

Table 4.2 Parameter settings

Parameter	Default setting
T_h	0.8
S_b	10
D_{th}	0.6
ω, τ	$\omega = 0.4, \tau = 0.001$
λ_g, λ_b^0	$\lambda_g = 0.1, \lambda_b^0 = 0.5$
Number of PMUs: N	39
Number of samples each test: K	1000
State variables that collected	$\delta, V, L_{Mvar}, L_{MW}$

the second group shows that multiple pieces of the four rule specifications are violated (with multiple “1”s in R_j^t). Further, as shown in Fig. 4.6, each group is divided into four different cases: (a) single, (b) sparse, (c) random, and (d) dense, representing four distribution types of false measurement data. To be specific, case (a) describes that only single PMU is inserted with false measurement data; case (b) describes that multiple sparsely distributed PMUs are inserted with false measurement data; case (c) describes that multiple randomly distributed PMUs are inserted with false measurement data; and case (d) describes that multiple densely distributed PMUs are inserted with false measurement data.

Table 4.3 The detection rate and the average iterations of FDD algorithm with single rule violated false measurement data under four different distribution types. The number of PMUs with false measurement data is 6

Distribution type	Detection rate	Average iterations
Single	100.0%	1.000
Sparse	100.0%	1.000
Random	97.10%	1.173
Dense	80.40%	2.071

Tables 4.3 and 4.4 show the simulation results in terms of the detection rate and the average iterations of the FDD algorithm for detecting false measurement data with single violated rule and multiple violated rules, respectively. We observe from both Tables 4.3 and 4.4 that, either singly or sparsely distributed PMU(s) with inserted false measurement

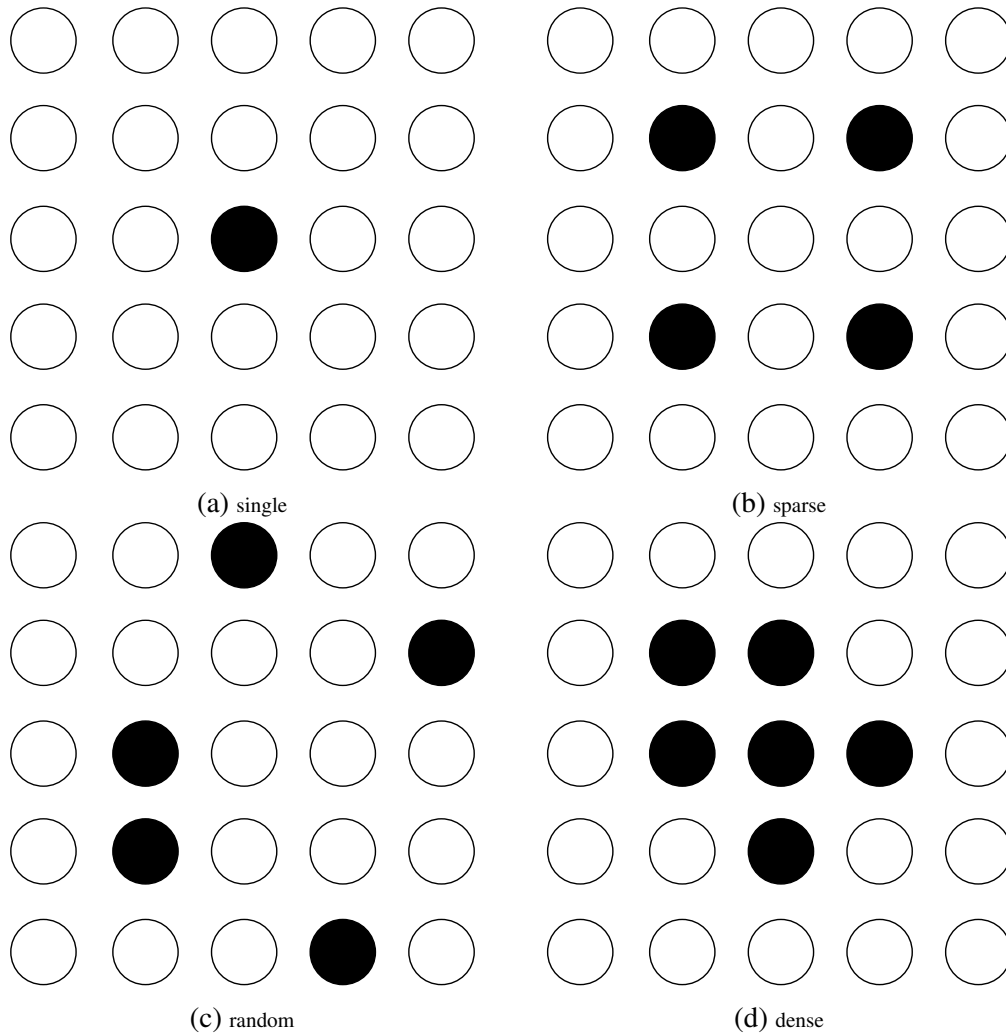


Fig. 4.6 Four different cases of the distribution of PMUs with inserted false measurement data: single, sparse, random, and dense

Table 4.4 The detection rate and the average iterations of FDD algorithm with multiple rules violated false measurement data under four different distribution types. The number of PMUs with false measurement data is 6

Distribution type	Detection rate	Average iterations
Single	100.0%	1.000
Sparse	100.0%	1.000
Random	97.90%	1.107
Dense	93.70%	1.520

data can be easily detected by our FDD algorithm with a 100% detection rate. As for either randomly or densely distributed PMUs with inserted false measurement data, FDD has a

high detection rate but not 100%. The reason is that, in most cases, the collaborative FDD performs well for detecting anomalous data when these corresponding PMUs are located near the inner regions of the grid. The anomalies can be identified by starting from the peripheral PMUs at the first iteration to the inner PMUs at the subsequent iterations. While, in some extreme and rare cases, if these anomalous PMUs are concentrated at the marginal regions of the grid, only peripheral PMUs in the vicinity of the inner regions can be identified. After the first or two iterations, the peripheral anomalous PMUs can be identified and their connections to other PMUs removed. Therefore, other anomalous PMUs in marginal regions may be isolated with only anomalous neighbouring PMUs. They can collude with each other to mutually protect each other by showing the same results R_i^t . Such extreme cases may occur in dense distribution type simulation experiments, so the dense type holds relatively lower detection rate in both group one and group two.

The average iterations for either singly or sparsely distributed PMU(s) with inserted false measurement data in both group one and group two are 1.000, as the inserted anomalous data of these two types can be easily identified by collaborative detection with only one iteration. In random distribution type, the average iterations are 1.173 and 1.107 for the two groups, respectively. This means that one round FDD can successfully detect the inserted false data, but in some situations, it requires another one to two rounds to detect the false data. Note that, in our simulation experiments, for undetected false data, the number of iterations is set as 5, the upper bound of FDD algorithm. As for the densely distribution type, the average iterations are 2.071 and 1.0520 respectively. This shows that, compared with random distribution type, more cases require additional FDD iterations to detect the inner false data.

Interestingly, the simulation results also show that, group two simulations can achieve a higher or equal detection rate with fewer average iterations than group one. This is because our FDD algorithm detects the false data when at least one rule is violated, so in group two it is much easier for FDD to detect the anomalous data.

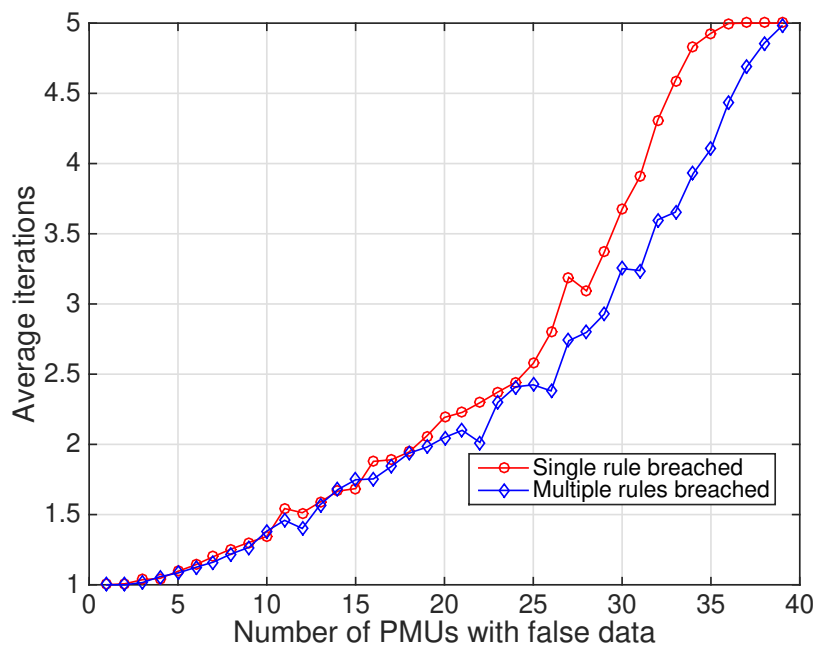


Fig. 4.7 The average iterations needed for FDD algorithm vs. different numbers of PMUs with false measurement data

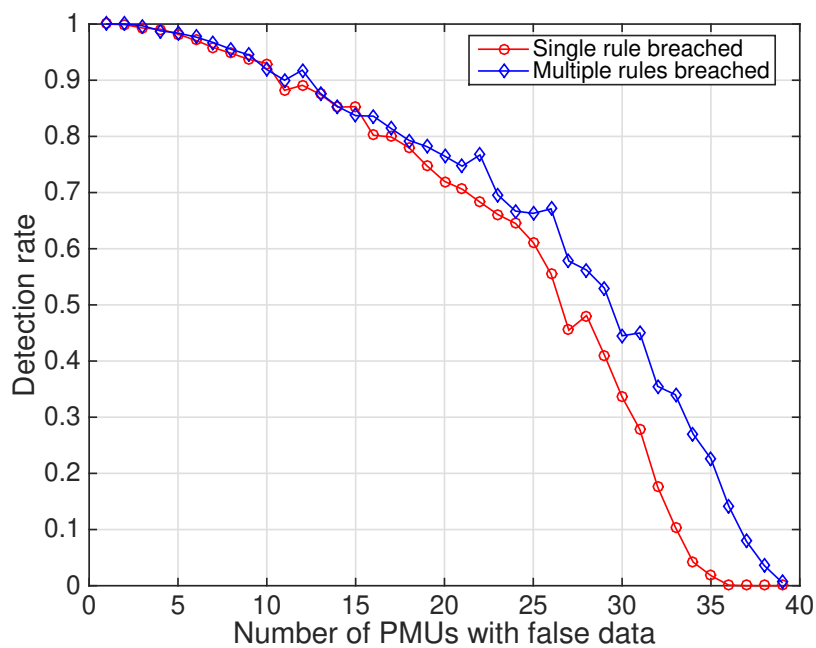


Fig. 4.8 The detection rate of FDD algorithm vs. different numbers of PMUs with false measurement data

In addition to the above results, we studied the relationship between the average iterations and the number of PMUs with false data under random distribution type as shown in Fig. 4.7, and the corresponding detection rate as well in Fig. 4.8. Clearly, the value of the average iterations increases, and eventually up to 5, the upper bound, as the increase in the number of PMUs with false data. Correspondingly, the value of the detection rate drops from 1 to 0 while the number of PMUs with false data increases. We also observe similar results in the sense that both values of the average iterations and the detection rate of multiple rules violation outperformed the single rule violated data.

4.4.2 Identification of Compromised PMUs with Our Reputation System

The performance of our reputation system can be affected by the following critical parameters: (1) ω , the weight assigned for weight assigned for the history reputation level; (2) D_{th} , the detection threshold; (3) λ_b , the impact factor; and (4) S_b^t , the number of successive observations of “bad” data.

Figure 4.9 shows the fluctuations of a PMU’s reputation level under various values of ω . Three FmDI events, each lasting 10 samples, are inserted into the PMU’s measurement data. This figure shows that, the higher the ω is, the more the current reputation level T^t relies on its history value T_h . Particularly, $\omega = 0.0$ indicates that $T^t = T_h$, and $\omega = 1.0$ indicates that $T^t = T_u$.

Figure 4.10 shows the fluctuations of a PMU’s reputation level under various values of D_{th} . Six FmDI events, each lasting 10 samples, are inserted into the PMU’s measurement data. We observe from this figure that a higher D_{th} s hold a lower tolerance to PMUs’ “bad” behaviors, while lower D_{th} s have higher tolerance to PMUs’ “bad” behaviors. In other words, higher D_{th} s are more sensitive than lower D_{th} s. For example, when $D_{th} = 0.65$, our reputation system raises an alarm when the first FmDI event is inserted.

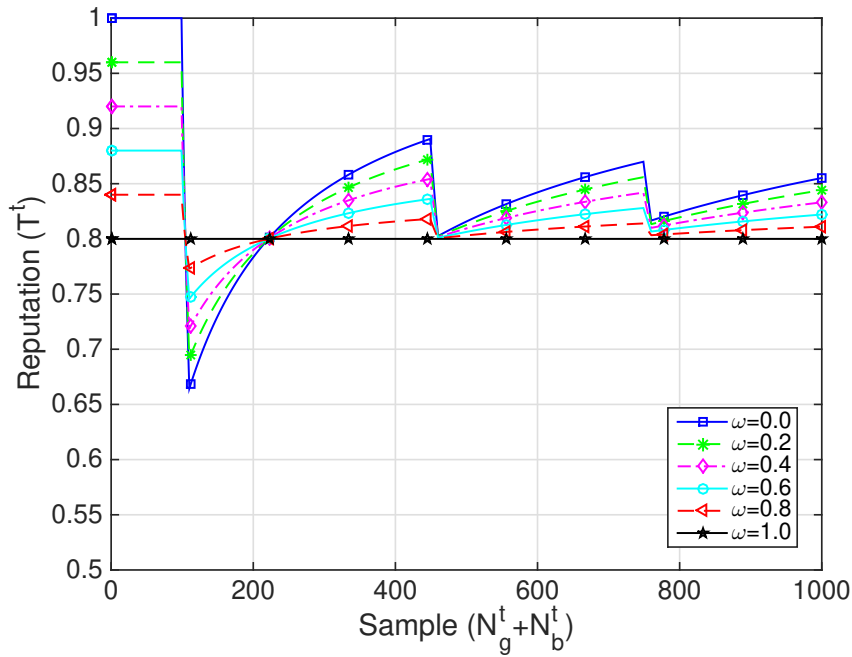


Fig. 4.9 The reputation level of a PMU under various values of ω ($T_h = 0.8, D_{th} = 0.6, S_b = 10, \lambda_b^0 = 0.5$)

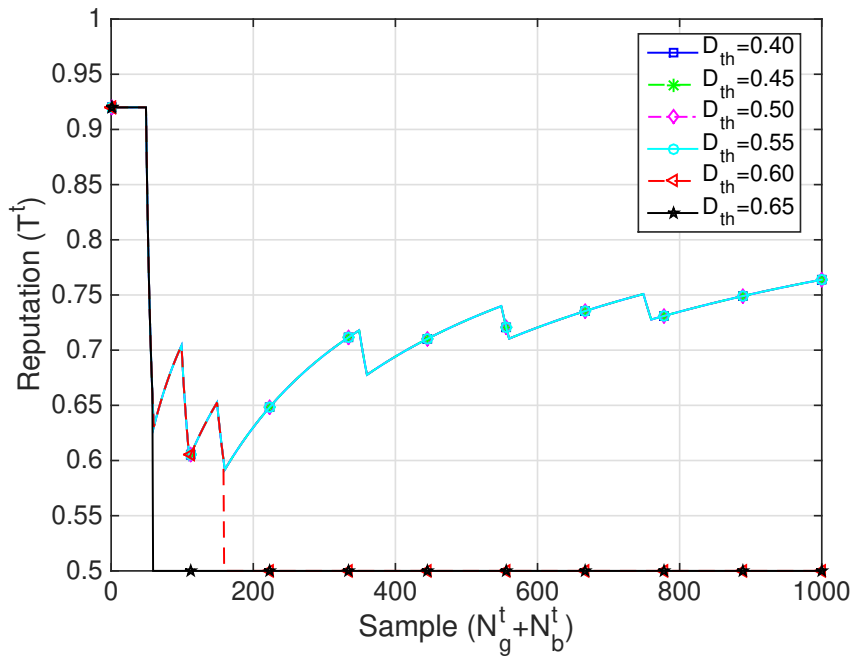


Fig. 4.10 The reputation level of a PMU under different under various values of D_{th} ($T_h = 0.8, \omega = 0.4, S_b = 10, \lambda_b^0 = 0.5$)

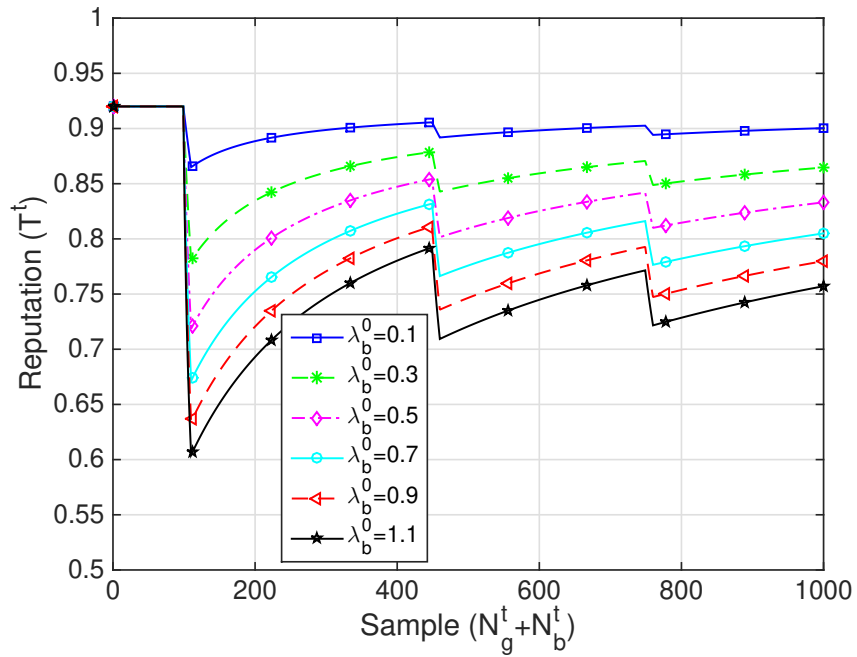


Fig. 4.11 The reputation level of a PMU under various values of λ_b^0 ($T_h = 0.8, \omega = 0.4, D_{th} = 0.6, S_b = 10$)

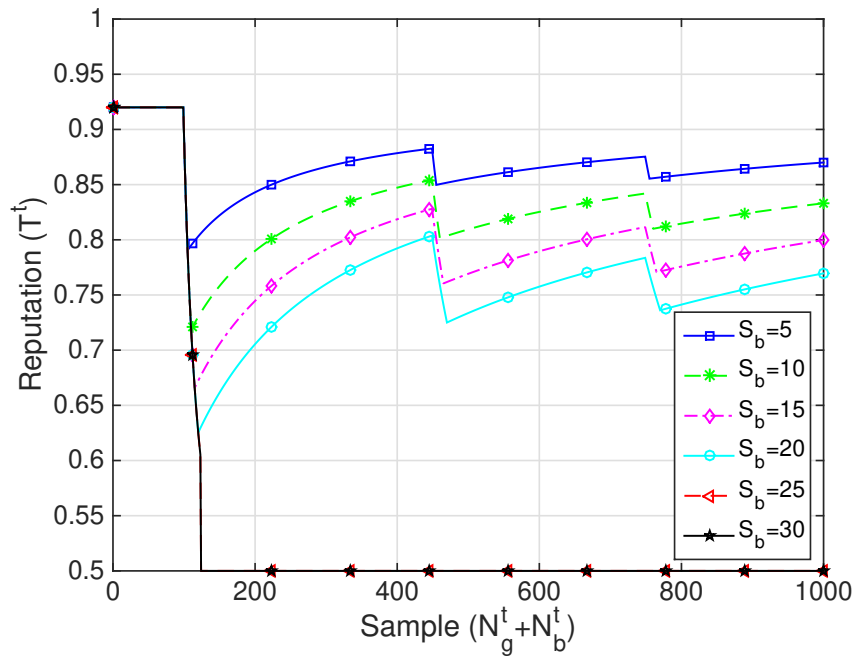


Fig. 4.12 The reputation level of a PMU under various values of S_b ($T_h = 0.8, \omega = 0.4, D_{th} = 0.6, \lambda_b^0 = 0.5$)

The relationship between the reputation level and the λ_b^0 is plotted in Fig. 4.11. Three FmDI events, each lasting 10 samples, are inserted into the PMU's measurement data. Clearly, the higher the λ_b^0 , the more adverse the consequence of penalty to the reputation level, which means that the reputation level decreases significantly.

A similar relationship between the reputation level and the S_b is plotted in Fig. 4.12. Also, three FmDI events but different lengths are inserted into the PMU's measurement data. Similar to Fig. 4.11, this figure shows that the larger the S_b , the more significance the penalty has on the reputation level, as large S_b results in more times of λ_b^t adjustment, i.e., $\lambda_b^t = \lambda_b^{t-1} * e^\tau$. For instance, with $D_{th} = 0.6$, the reputation level drops quickly below D_{th} if $S_b = 30$.

4.5 Summary

In this paper, we proposed a novel DHCD method to identify and mitigate FmDI attacks in smart grid CPS. Specifically, a rule specification based real-time collaborative detection system was designed to identify the anomalies of measurement data. In addition, a new reputation system with an ARU algorithm was presented to evaluate the overall running status of the PMUs, which can be used to identify compromised PMUs. We then demonstrated the utility of the proposed approach using simulations of the IEEE 39-bus power system.

As previously discussed, our method is designed to detect the malicious activities resulting in the anomaly of measurement data. Future work would include extending the proposed approach to capture power system faults (e.g., voltage disturbance, open circuit, and short circuit).

Chapter 5

DDOA: A Dirichlet-Based Detection

Scheme for Opportunistic Attacks

In the hierarchical control paradigm of a smart grid cyber-physical system, decentralized LAs (local agents) can potentially be compromised by opportunistic attackers to manipulate electricity prices for illicit financial gains. In this chapter, to address such opportunistic attacks, an example of FcDI attacks, we propose a Dirichlet-based detection scheme (DDOA), where a Dirichlet-based probabilistic model is built to assess the reputation levels of LAs. Initial reputation levels of the LAs are first trained using the proposed model, based on their historical operating observations. An adaptive detection algorithm with reputation incentive mechanism is then employed to detect opportunistic attackers. We demonstrate the utility of our proposed scheme using data collected from the IEEE 39-bus power system with the PowerWorld simulator.

5.1 Introduction

With the increasing connectivity of society and advancement of ICT, smart grid cyber-physical system is increasingly a commonplace. Smart grid cyber-physical system is a large-scale interconnected power infrastructure spanning across one or more jurisdictions.

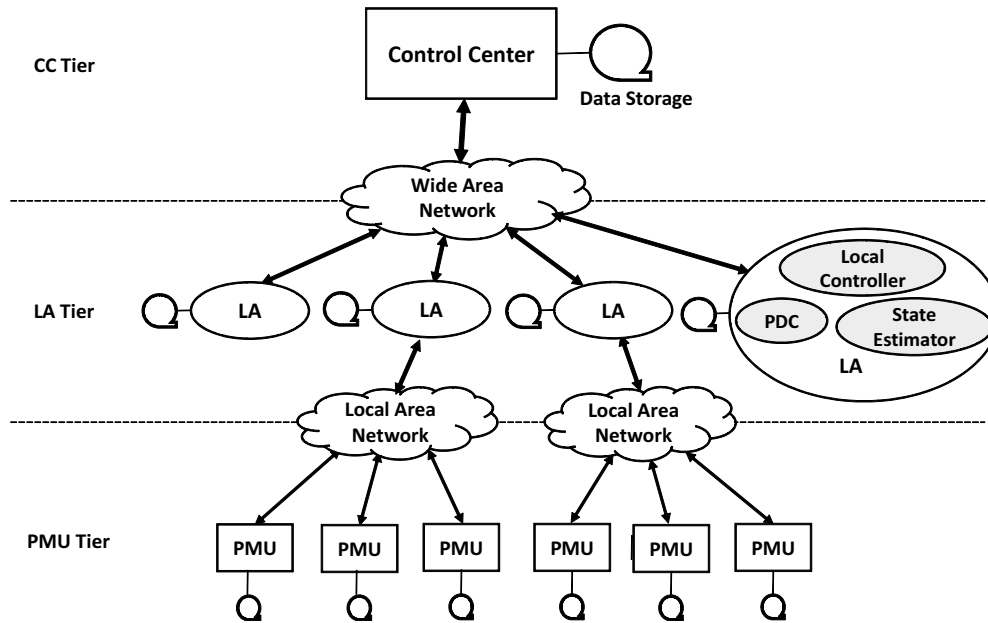


Fig. 5.1 Three-tier hierarchical flocking-based framework for future smart grids

To guarantee high reliability and robustness of the underlying critical infrastructure, real-time monitoring, data analytics, and control are highly critical. Empirically, data analytics is generally performed by the state estimator at the system control center [21, 106, 107]. However, with the increasing number of interconnections, nonlinearity, and dynamics, real-time data analytics will inevitably impose significant computational burden and complexity on control center [108]. If this is not well-managed, control center's operating efficiency will be adversely affected, resulting in cascading effects - e.g. affecting the reliability and the robustness of the power grid and eventually crippling the power grid. One of the potential solutions to address the exacting computational requirements on the control center identified in the literature is the hierarchical control framework. In such a framework, decentralized LAs perform real-time data analytics activities in their local region [109, 110].

While hierarchical framework can effectively reduce the computational burden of the control center, it may result in unintended security consequences [108]. For example, in the current centralized power system, it is easier to devote efforts and resources to secure a central entity (i.e. control center); thus, control center is generally regarded as a fully trusted

party. In a hierarchical framework, however, it is not realistic to expect that all decentralized LAs can be secured to the same level as the control center.

The upward trend in Internet-of-Things and integration of power grids with ICT have also resulted in an increased attack vector. For example, vulnerabilities in existing power system, or connected devices and/or entry points can be exploited by cybercriminals. According to the monitor newsletter of Industrial Control Systems Cyber Emergency Response Team (ICS-CERT), in Fiscal Year 2015 (i.e. 1 October 2014 to 30 September 2015), ICS-CERT of the U.S. Department of Homeland Security has reportedly responded to 295 cybersecurity incidents involving critical infrastructures, and the energy sector is the second most targeted critical infrastructure sector [111]. The dangers of threats to cyber-physical systems are evidenced by recent attacks (e.g. on a German steel mill that destroyed a blast furnace [112]) and attempts (e.g. ISIS attempted to hack U.S. electric power utilities to steal confidential grid information and launch terrorist attacks [54]). Successful attacks could potentially overwhelm and paralyse the country's interconnected critical infrastructure sectors and, consequently, cause severe social unrest.

Unsurprisingly, security of smart electricity markets has attracted the attention of security researchers [113, 38, 114]. However, we observe existing efforts appear to focus on mitigating data integrity attacks (i.e. attackers falsify measurement data to “blind” the system in order to manipulate electricity prices [115]). Generally, it is assumed that attackers have access to the system configuration, and are able to simultaneously falsify a set of measurement data at several PMUs at will. Kosut investigated the various attack strategies and their countermeasures for malicious data integrity attackers in smart grids [116]. Xie and Esmalifalak *et al.* also examined FDI attacks in deregulated electricity markets, which could be used to manipulate nodal electricity prices [38, 117, 37]. These studies focused on the centralized power system model. With the increasing demands on interconnectivity between systems in future smart grids, recent research focus have shifted to security in hierarchical smart grids (see [108, 118, 103, 41]). For example, Li [41] proposed a distributed quick

detection scheme for FDI attacks in smart grids. Vukovic [108] analyzed the security issues in distributed power system and proposed a methodology to detect and mitigate data integrity attacks.

Unfortunately, most existing efforts were directed to the insider data integrity attacks. In addition to criminally-, politically-, and ideologically-motivated attacks, cybercriminals may be interested in compromising smart grids by manipulating smart electricity markets for illicit financial gains [117, 119, 115]. Opportunistic attacks [21, 120] are one such example, which usually are also initiated by inside attackers. Specifically, rather than seeking to falsify measurement data by compromising a set of PMUs, opportunistic attackers attempt to manipulate electricity prices by only compromising the intelligent electronic device which is responsible for determining the real-time electricity prices, say the LA. The compromised LA can issue fake commands to the local generators, distributors, and transformers to shift the normal demand-supply relations, which will further influence the electricity price at each local bus. If colluded with other participants in smart electricity markets (e.g. power suppliers and utilities), the attackers can make a great amount of illicit financial profits through the wide fluctuations of the electricity prices [117]. This is the focus of this chapter.

Since opportunistic attacks are unlikely to result in any physical damages to the power system, it is a challenging task for conventional IDS to identify. Moreover, opportunistic attackers can flexibly adjust their attack strategies (e.g. probability to launch an attack when there is a chance) based on system noise level to evade detection or scrutiny [21]. Hence, to identify the abnormality of any possible compromised LA, an effective way is to observe and assess their behaviors (i.e. operations and corresponding variable states) over a long period of time. In this chapter, we seek to mitigate opportunistic attacks by presenting a novel Dirichlet-based detection scheme (hereafter referred to as DDOA). The scheme allows control center to effectively identify compromised LAs by observing their operating behaviors. We regard the contributions of this work to be three-fold:

- We first divide the smart grid infrastructure into a three-tier hierarchical framework, which is designed to effectively reduce the computational burden on the control center. This framework also makes it possible to guarantee high reliability and robustness of future smart grids.
- We pioneer to study the opportunistic attacks in smart electricity market, and build up a Dirichlet-based reputation model to monitor and assess the performance of the LAs by observing their behaviors over a long period of time.
- Lastly, we propose and evaluate an adaptive detection scheme with reputation incentive mechanism, which can effectively and accurately identify potential opportunistic attackers hidden in the smart electricity market and prevent them from manipulating electricity prices. In addition, two-level detection thresholds are also employed in our DDOA scheme, which can effectively differentiate malicious activities from common system faults in smart grids.

The remainder of this chapter is organized as follows. Section 5.2 presents the system model, the threat model, and our design goals. In Section 5.3, we introduce the preliminaries required in the understanding of this work. Our proposed Dirichlet-based reputation model and detection scheme is detailed in Section 5.4, and the performance evaluation is presented in Section 5.5. Section 5.6 concludes the chapter.

5.2 Models and Design Goals

In this section, we formalize both system and threat models, as well as describe the design goals.

5.2.1 System Model

As shown in Fig. 5.1, we consider a hierarchical flocking-based framework for future smart grids as our system model. This model comprises three tiers, namely: the lowermost tier of PMU, the intermediate tier of LA, and the uppermost tier of control center. Their roles and responsibilities are illustrated as follows:

- PMUs, deployed at each bus and generator across the whole power system, are geographically flocked, forming several flockings. They collect real-time measurement data of system status in each flocking area (e.g. power generations G , power loads L , and line power flows F), and report collected data to the PDC located in the upper tier LA area.
- The LA (formed by PDC, state estimator, and local controller) in the flocking area analyzes the real-time system status of its monitored local area with the reported data, and transforms the data to the uppermost tier of control center as required. Specifically, the PDC collects reported measurement data from PMUs; the state estimator is utilized to estimate actual system status in the flocking area; and the local controller then analyzes the estimated data, determines the locational marginal price (LMP), and issues feedback commands to local generators, distributors, transformers, etc.
- The control center stores and analyzes the measurement data for various applications (e.g. state estimation, contingencies analysis, and event diagnostics). In addition, in our model, control center is also responsible for monitoring and assessing the reputation levels of the subordinate LAs to identify abnormal LA behavior.

In this work, we assume that both control center and LAs make use of state estimation to analyze the system status of either the entire region or local regions. Particularly, control center carries out state estimation with a low frequency to reduce computational requirements (see Section 5.4.3).

5.2.2 Threat Model

Unlike traditional power systems, future smart grids will delegate real-time monitoring, data analytics, and control tasks from control center to its subordinate LAs. As aforementioned, it is natural to assume that only control center is a fully trusted party, while LAs are more likely to be compromised by malicious attackers. In our model, PMUs are assumed to be honest (i.e. data reported by PMUs to PDC are assumed to be without falsification).

By successfully compromising an LA, attackers can launch FcDI attacks by issuing fake control commands to local generators, distributors, and transformers to manipulate normal demand-supply relations in a specific flocking area. Such actions could result in changes of the LMP in the area. As this is a premeditated activity, attackers can exploit the price fluctuations/changes for financial gains. For example, attackers can collude with other players in the smart electricity markets and purchase a significant amount of electricity at a low price prior to the attacks. Once the price has been artificially jacked up, attackers will seek to sell the pre-purchased electricity to users in the grid.

Fig. 5.2 presents an example of the contouring map of the distribution of electricity prices under normal conditions on the IEEE 39-bus power system. Areas covered by various colors reflect different demand-supply relations. In case of occurrence of malicious attacks, these normal relations and consequently, electricity prices will be intentionally altered. These attacks can be broadly categorized into random attacks, reckless attacks, and opportunistic attacks.

1. Random attacks are conducted with a definite attack probability $P_a \in [0, 1]$. Since such attacks are carried out in a regular mode, it is easier to identify the attacks using traditional IDS or intrusion prevention systems (IPS).
2. Reckless attacks are launched on an ad-hoc basis. Specifically, once an opportunity appears, attackers will launch an attack without hesitation and planning. Consequently, reckless attackers are usually the easiest to be identified.

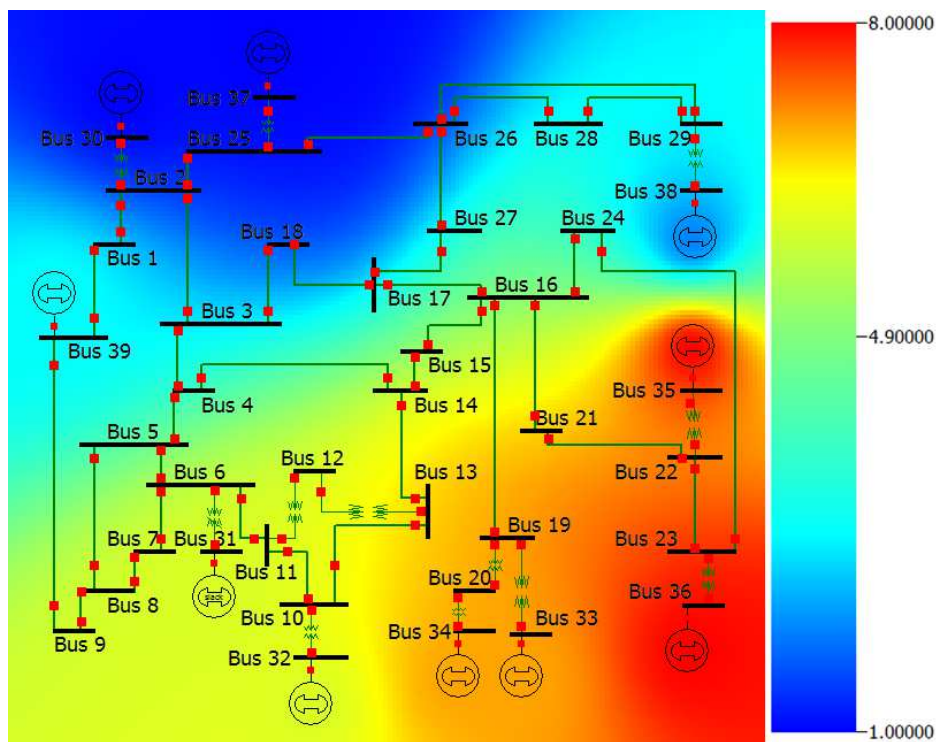


Fig. 5.2 The contouring map of electricity price distribution on IEEE 39-bus power system

3. Opportunistic attacks are carried out based on the system noise with an attack probability $P_a = C \cdot P_n^\varepsilon$, where C is a constant coefficient, and ε denotes a scalar of the system noise P_n . Particularly, $\varepsilon > 1$ indicates conservative opportunistic attackers, while $\varepsilon < 1$ indicates aggressive ones. Therefore, the larger the system noise is, the higher the attack probability will be.

It is widely believed that opportunistic attackers are the most cunning attackers, as they adapt their attack probabilities according to the system noise. Therefore, it is significantly challenging to identify such attackers using traditional detection schemes (e.g. IDS and IPS). In this chapter, we aim to propose an effective scheme to identify and detect opportunistic attackers.

5.2.3 Design Goals

The key objective of the proposed DDOA scheme is to provide an effective approach to accurately identify and detect opportunistic insider attacks in smart electricity markets. Our design goals are as follows:

1. Future smart grids are expected to be a hierarchical system, due to their capability to ensure efficiency, stability, and reliability of power system in situations with ever-increasing electricity demands, integration of renewable energy resources, and various data analytical applications. Thus, we employ a three-tier hierarchical control framework for future smart grids to support these critical requirements.
2. LAs play a prominent role in distributed flocking areas, and it is important to ensure their functionality. Since LAs cannot be fully trusted (unlike a control center), we need to be able to efficiently and accurately monitor and assess their behaviors. Thus, we present a Dirichlet-based reputation model to assess LA's operating conditions.
3. To continuously monitor all LAs' operating conditions, we propose an effective detection scheme based on our Dirichlet-based reputation model to identify LA compromised by an opportunistic attacker. In addition, we use collected real-time data in PowerWorld simulator to validate the effectiveness of our proposed DDOA scheme.

5.3 Preliminaries

In this section, we briefly introduce some preliminaries required in the understanding of the remaining of this chapter.

5.3.1 State Estimation

Although the basic concepts of state estimation have been introduced in Section 2.1, we provide additional details of what kinds of data are measured and estimated. As shown in

Table 1.1, PMUs can measure the values of a myriad of state variables. The measurement data can be classified into three main types, including power generations \mathbf{z}_G , power loads \mathbf{z}_L , line power flows \mathbf{z}_F .

According to DC power flow model, the estimated system states are given by

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{\mathbf{x}}_V \\ \hat{\mathbf{x}}_\theta \end{bmatrix} = \mathbf{H}\Lambda \begin{bmatrix} \mathbf{z}_G \\ \mathbf{z}_L \\ \mathbf{z}_F \end{bmatrix}. \quad (5.1)$$

Usually, the estimated system states only comprise voltage magnitudes $\hat{\mathbf{x}}_V$ and phase angles $\hat{\mathbf{x}}_\theta$. While, it is easy to further calculate the estimated power generations $\hat{\mathbf{x}}_G$, power loads $\hat{\mathbf{x}}_L$, line power flows $\hat{\mathbf{x}}_F$ by using $\mathbf{x} = [\hat{\mathbf{x}}_V, \hat{\mathbf{x}}_\theta]$ [115], i.e.,

$$\begin{bmatrix} \hat{\mathbf{x}}_G \\ \hat{\mathbf{x}}_L \\ \hat{\mathbf{x}}_F \end{bmatrix} = \mathbf{F} \begin{bmatrix} \hat{\mathbf{x}}_V \\ \hat{\mathbf{x}}_\theta \end{bmatrix} = \mathbf{F}\mathbf{H}\Lambda \begin{bmatrix} \mathbf{z}_G \\ \mathbf{z}_L \\ \mathbf{z}_F \end{bmatrix}. \quad (5.2)$$

where \mathbf{F} is a matrix relating the voltage magnitudes and phase angles to power generations, power loads, and line power flows.

5.3.2 Real-Time LMP

In smart electricity markets, the real-time LMP within an LA area is determined based on the estimated real-time system states. LMP is defined as the cost to serve the next unit increment of power load (say 1MWh) at each bus by comprehensively taking into account actual power generations, power loads, and line flows with respect to transmission line limits [121].

Such calculations can be formulated as an incremental linear optimization problem with state estimates as described in Eq. (5.3). The objective is to minimize the cost function subject to the power balance constraint, the generation megawatt bounds, the transaction megawatt bounds and any transmission constraints that currently exist on the system. This

optimization problem can be formulated as follows:

$$\begin{aligned}
\min \quad & \mathcal{J} = \sum C_i(\Delta G_i) - \sum C_j(\Delta L_j) \\
s.t. \quad & \sum \Delta G_i - \sum \Delta L_j = 0 \\
& \Delta G_i^{min} \leq \Delta G_i \leq \Delta G_i^{max} \\
& \Delta L_i^{min} \leq \Delta L_i \leq \Delta L_i^{max} \\
& A_{ik}\Delta G_i + D_{jk}\Delta L_j \leq 0,
\end{aligned} \tag{5.3}$$

where C_i and C_j are calculated real-time offer for generator i and real-time bid for load j , respectively [121]. A_{ik} is a matrix of shift factors for generation bus i (with respect to the reference bus) on the binding transmission constraints (k), and D_{jk} is a matrix of shift factors for load bus j (with respect to the reference bus) on the binding transmission constraints (k). The LMP values at each bus can be expressed as

$$LMP_i = \lambda - \sum A_{ik} * SP_k, \tag{5.4}$$

where λ is the marginal price of generation at the reference bus [122]. A_{ik} is a shift factor for bus i on binding constraint k , and SP_k is the shadow price of constraint k .

5.3.3 Dirichlet Distribution

Dirichlet distribution [123] is a family of continuous multivariate probability distributions, parameterized by a vector α of positive reals. Let $X = \{x_1, x_2, \dots, x_k\}$ be a discrete random variable, where $x_i > 0$ for $i = 1, 2, \dots, k$ and $\sum_{i=1}^k x_i = 1$. Suppose that $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_k]$ with $\alpha_i > 0$ for all i from 1 to k , and let $\alpha_0 = \sum_{i=1}^k \alpha_i$. Then, X is said to be a Dirichlet distribution with parameters α , which is denoted by $X \sim Dir(\alpha)$. Then, the

probability density function is expressed as

$$f(X; \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^k x_i^{\alpha_i-1} = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k x_i^{\alpha_i-1}, \quad (5.5)$$

where $B(\cdot)$ is a Beta function, and $\Gamma(\cdot)$ is a Gamma function.

The expectation and variance of $X = x_i$ are respectively given by

$$E[x_i] = \frac{\alpha_i}{\alpha_0}, \text{Var}[x_i] = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}. \quad (5.6)$$

5.4 Proposed DDOA Scheme

In this section, we elaborate our proposed DDOA scheme, which is composed of three parts: behavior rule specifications, Dirichlet-based reputation model, and detailed description of DDOA.

5.4.1 Behavior Rule Specifications

Smart grid is a large-scale interconnected cyber-physical system. The behaviors (i.e. operations and variable status) of the physical devices are an accurate reflection of their responses to the feedback commands from the control unit. Thus, assessing the behaviors of physical devices will be an efficient and reliable way to detect abnormalities in the control units. The complex interconnections within a smart grid result in multiple inter-constraints between the state variables, which can be utilized to specify a set of rule specifications for the control units' behaviors. Therefore, in this work, we define several behavior rule specifications that LAs must follow under normal operating conditions (see Table 5.1). This will allow us to identify any operating abnormality.

Let us take the first rule $R1$ as an example, G_i^t denotes the measurement value of power generation at generator i at time instant t , while \hat{G}_i^t denotes the corresponding expected value. $R1$ describes that the absolute difference between the measured value and the expected

Table 5.1 Rule specifications

Index	Rule	Description
$R1$	$ G_i^t - \hat{G}_i^t \leq \tau_G$	The absolute difference of G between measured and expected values should be below a safe threshold τ_G
$R2$	$ L_i^t - \hat{L}_i^t \leq \tau_L$	The absolute difference of L between measured and expected values should be below a safe threshold τ_L
$R3$	$ F_i^t - \hat{F}_i^t \leq \tau_F$	The absolute difference of F between measured and expected values should be below a safe threshold τ_F
$R4$	$\tau_P^{min} \leq G_i^t \leq \tau_G^{max}$	The value of G itself should be limited within a specified safe range $[\tau_G^{min}, \tau_G^{max}]$
$R5$	$\tau_L^{min} \leq L_i^t \leq \tau_L^{max}$	The value of L itself should be limited within a specified safe range $[\tau_L^{min}, \tau_L^{max}]$
$R6$	$\tau_F^{min} \leq F_i^t \leq \tau_F^{max}$	The value of F itself should be limited within a specified safe range $[\tau_F^{min}, \tau_F^{max}]$

value should be limited to a specified safe threshold τ_G . In our scheme, the expected values are defined by the values estimated by the control center (other than by LAs) using state estimation, since control center is the fully trusted party. Apart from $R1$, in real-world applications, the value of G_i^t should also be constrained within a safe range, say $[\tau_G^{min}, \tau_G^{max}]$ as described in $R4$. Similarly, parallel rules can also be specified for power loads L , power line flows F as described in other rules.

Measurement values of the state variables are revealing of the LA's behavior. Thus, it is logical to infer that deviation of these rule specifications imply abnormality. A single deviation may not sufficiently indicate that an LA is compromised, as the deviation may be due to system noise. Therefore, a conjunctive form of these rules and long-term observation of these conjunctive rules are employed in this work to effectively and accurately assess LAs' behaviors (and reduce false positive rate).

$$\mathcal{R} = R1 \cup R2 \cup R3 \cup R4 \cup R5 \cup R6 \quad (5.7)$$

The conjunctive rule \mathcal{R} is the combination of all specified rules as shown in Eq. (5.7). To simply represent whether a rule is compliant, we use “1” to denote non-compliance of a rule, while “0” to denote compliance. As such, \mathcal{R} can be represented as a binary sequence. For example, “100010” indicates that $R1$ and $R5$ are non-compliant while the remaining rules are compliant. Particularly, full compliance of the conjunctive six rules is expressed as “000000”, which is our reference sequence, seq_{ref} .

We now define the *compliance level* of each binary sequence as follows:

$$\rho = 1 - dist(seq, seq_{ref}), \quad (5.8)$$

where seq_{ref} is the binary sequence extracted from each piece of measurement data, and $dist$ function denotes the normalized distance between each binary sequence and the seq_{ref} . Many distance-based algorithms can be utilized in our scheme, like Hamming distance, Euclidean distance, etc. In this work, we use Euclidean distance to conduct our simulation experiments.

In real-world applications, multi-level systems (e.g. quaternary, octonary) can be employed instead of binary system, which will yield a more accurate compliance level of these rules. In addition, different rules may have various significance levels to the power system. Hence, distinguished weights can be assigned to each rule to enhance the accuracy of the compliance levels. However, either multi-level systems or weighted rules can impose considerable computational burden on control center and require a significant amount of storage for real-time detection applications. Therefore, if multi-level systems and/or weighted rules are to be integrated into our DDOA scheme, efficient optimization algorithms or balancing mechanisms will be required prior to deploying this enhanced scheme.

5.4.2 Dirichlet-Based Reputation Model

In our system model, control center is responsible for monitoring and assessing the behaviors of LAs, and determining whether any LA has been compromised based on a series of

historical observations. As known to us, Bayesian statistics can be used to measure the uncertainty of a decision and provide future knowledge of such decision based on a set of historical observations. In this way, a Bayesian statistics methodology is employed in our work to assist control center in making correct decisions of whether or not an LA has been compromised, and provide control center with knowledge of LAs' most possible behaviors in the future. Specifically, of the statistical techniques, Beta distribution is a viable method to determine whether a decision is correct, while a Dirichlet distribution can determine at what level a decision is correct [123]. In this chapter, to obtain a more accurate assessment of LAs' behaviors and hence, a more accurate decision, we consider a Dirichlet-based probabilistic model.

Dirichlet distribution is grounded on initial beliefs regarding an unknown event represented by a prior distribution. The initial beliefs combined with a series of historical observations can be represented by a posterior distribution. The posterior distribution is best suited for our reputation model, as the reputations are required to be updated based on historical observations. Let X be a discrete random variable denoting the compliance level ρ of the measurement data for an LA. X takes values in the set $X = \{x_1, x_2, \dots, x_k\}$, where $x_i \in [0, 1]$ and $x_{i+1} > x_i$ ($i = 1, \dots, k$). Usually, we have $x_1 = 0$, and $x_k = 1$. Let $\mathbf{p} = [p_1, p_2, \dots, p_k]$ with $\sum_{i=1}^k p_i = 1$ be the probability distribution of X , i.e. $p\{X = x_i\} = p_i$. In addition, let $\zeta = [\zeta_1, \zeta_2, \dots, \zeta_k]$ denote the cumulative historical observations and initial beliefs of X . Then, we can model \mathbf{p} with a posterior Dirichlet distribution as follows:

$$f(\mathbf{p}|\zeta) = Dir(\mathbf{p}|\zeta) = \frac{1}{B(\zeta)} \prod_{i=1}^k p_i^{\zeta_i-1} = \frac{\Gamma(\zeta_0)}{\prod_{i=1}^k \Gamma(\zeta_i)} \prod_{i=1}^k p_i^{\zeta_i-1}, \quad (5.9)$$

where $B(\cdot)$ is a Beta function, and $\Gamma(\cdot)$ is a Gamma function. $\zeta_0 = \sum_{i=1}^k \zeta_i$. Given the historical statistics ζ , the expected value of the probability of \mathbf{X} to be x_i is given by

$$E(p_i|\zeta) = \frac{\zeta_i}{\zeta_0}. \quad (5.10)$$

Let $p_i^j(t)$ denotes the probability that LA_j behaves with an compliance level x_i at time instant t , where $\sum_{i=1}^k p_i^j(t) = 1$. We model $p_i^j(t)$ using a posterior Dirichlet distribution as shown in Eq. (5.9). We define a random variable $Y^j(t)$ denoting the sum of the products of the grade and probability of each compliance level in $\mathbf{p}^j(t) = [p_1^j(t), p_2^j(t), \dots, p_k^j(t)]$ for LA_j , which is given by

$$Y^j(t) = \omega \mathbf{p}^j(t) = \sum_{i=1}^k \omega_i p_i^j(t), \quad (5.11)$$

where $\omega = [\omega_1, \omega_2, \dots, \omega_k]$ is the grade assignment for each compliance level, measuring the different impacts on LA_j 's overall operating performance. This design will significantly improve the accuracy of control center's decisions.

To assess the overall status of an LA's behaviors, we leverage the *reputation level* in our scheme. Specifically, the LA's behaviors can be described using various compliance levels. Thus, the reputation level of an LA can be defined by the graded mean value of each compliance level at time instant t as shown below:

$$R^j(t) = E[Y^j(t)] = \sum_{i=1}^k \omega_i E[p_i^j(t)] = \frac{1}{\zeta_0^j(t)} \sum_{i=1}^k \omega_i \zeta_i^j(t), \quad (5.12)$$

where $\zeta_i^j(t)$ is the cumulative historical observations of LA_j at time instant t with compliance level x_i . The variance of $Y^j(t)$ is then given by

$$\sigma^2[Y^j(t)] = \sum_{i=1}^k \sum_{l=1}^k \omega_i \omega_l \text{cov}[p_i^j(t), p_l^j(t)]. \quad (5.13)$$

Note that the covariance of $p_i^j(t)$ and $p_l^j(t)$ is given by

$$\text{cov}[p_i^j(t), p_l^j(t)] = \frac{-\zeta_i^j(t)\zeta_l^j(t)}{(\zeta_0^j(t))^2(\zeta_0^j(t) + 1)}. \quad (5.14)$$

5.4.3 Description of DDOA

In DDOA, control center first trains the initial reputation levels of the LAs based on the collected historical observations, as shown in Algorithm 5.1. Although this algorithm is designed for the training phase and can be always finished offline, it is worth noting that the time complexity of this algorithm is $\mathcal{O}(M \times N)$.

Algorithm 5.1 Reputation Level Training Algorithm

```

1: procedure DIRICHLET-BASED REPUTATION TRAINING
2:   for  $j = 1$  to  $M$ , control center do ▷  $M$  is the number of LAs
3:     1). Extracts  $N$  pieces of reported data from  $LA_j$ ;
4:     2). Computes the compliance level of each piece of data  $\rho^j(t)$ ,
5:        $t \in [1, N]$  with Eq. (5.8);
6:     for  $t = 1$  to  $N$  do
7:       for  $i = 1$  to  $k$  do
8:         if  $\rho^j(t) = x_i$  then
9:            $\zeta_i^j(t) \leftarrow \zeta_i^j(t-1) + 1$ ;
10:          break;
11:        else
12:           $\zeta_i^j(t) \leftarrow \zeta_i^j(t-1)$ ;
13:        end if
14:      end for
15:      a).  $\zeta_0^j(t) = \sum_{i=1}^k \zeta_i^j(t)$ ;
16:      b). Determines the reputation level of  $LA_j$  by
17:         $R^j(t) = \frac{1}{\zeta_0^j(t)} \sum_{i=1}^k \omega_i \zeta_i^j(t)$ .
18:    end for
19:  end for
20: end procedure

```

After the training phase, control center obtains the initial reputation level of each LA. While, these initial reputation levels only represent their historical performance. Recall that a smart grid needs to provide near real-time monitoring and control of the whole power system. As such, persistent observation and assessment of LAs' behaviors is always required to detect whether any LA may have been compromised. In the detection phase, we propose an adaptive algorithm with a *reputation incentive mechanism* to update LAs' reputation levels, whose functionality is described in Algorithm 5.2.

Algorithm 5.2 DDOA Algorithm

```

1: procedure REPUTATION UPDATING AND INTRUSION DETECTION
2:   Initialization:
3:    $T_{max}, T_{min}, T_S, T_W, H_s > H_m, N_{count} = 0,$ 
4:    $\mu_1 > \mu_2 > \dots > \mu_k, T_1 = T_2 = \dots = T_M = T_{max},$ 
5:    $\omega_1 = \bar{\omega}_1, \omega_2 = \bar{\omega}_2, \dots, \omega_k = \bar{\omega}_k$ 
6:   for  $j = 1$  to  $M$ , control center do with a frequency of  $1/T_j$ 
7:     1). Input:  $\rho^j(t), R^j(t-1), \omega_1^j, \omega_2^j, \dots, \omega_k^j$ 
8:     2). Classification:
9:        $LA_j \in \begin{cases} \mathbb{N}, & \text{if } R^j(t-1) > H_s \\ \mathbb{S}, & \text{if } H_s \geq R^j(t-1) \geq H_m \\ \mathbb{M}, & \text{if } R^j(t-1) < H_m \end{cases}$ 
10:    3). Judgement:
11:    switch  $LA_j$  do
12:      case:  $LA_j \in \mathbb{N}$ 
13:        a).  $LA_j$  is benign;
14:        b).  $\omega_k^j \leftarrow \min\{\omega_k^j e^{\mu_k}, 1\};$ 
15:        c).  $\omega_i^j \leftarrow \bar{\omega}_i, \forall i = 1, 2, \dots, k-1;$ 
16:        d).  $T_j \leftarrow T_{max};$ 
17:      case:  $LA_j \in \mathbb{S}$ 
18:         $T_j \leftarrow \max\{T_j/2, T_{min}\};$ 
19:        if  $\rho^j(t) = x_k$  then  $\triangleright x_k = 1$ 
20:           $\omega_k^j \leftarrow \min\{\omega_k^j e^{\mu_k}, 1\};$ 
21:           $T_{count} \leftarrow T_{count} + 1;$ 
22:          if  $T_{count} > T_S$  then
23:             $T_j \leftarrow \min\{T_j * 2, T_{max}\};$ 
24:             $T_{count} \leftarrow 0;$ 
25:          end if
26:        else
27:           $\omega_k^j \leftarrow \omega_k^j e^{-\mu_k};$ 
28:          if  $\rho^j(t) = x_i (i \neq k)$  then
29:             $\omega_i^j \leftarrow \omega_i^j e^{-\mu_i};$ 
30:          end if
31:           $T_{count} \leftarrow 0;$ 
32:        end if
33:      case:  $LA_j \in \mathbb{M}$ 
34:         $LA_j$  is compromised.
35:    4). Updates  $\zeta_i^j$  for  $i = 1, 2, \dots, k$  with reference to Algorithm 5.1.
36:    5). Determines  $R^j(t)$  using Eq. (5.12) with observation window
37:        $T_W.$ 
38:  end for
39: end procedure

```

Based on historical experiences, control center first specifies two thresholds H_s and H_m for the reputation level as the detection criteria, where H_s indicates suspicious threshold while H_m indicates malicious threshold. In a real-world scenario, occasional occurrence of system faults in smart grids is unavoidable and consequently, causes wide fluctuations of state variables. Such incidents impact (and reduce) both compliance and reputation levels. If a single detection threshold is utilized, we could possibly have a high false positive rate. However, two levels of threshold can successfully tolerate these system faults; thus, it can considerably reduce the false positive rate and further improve the detection rate.

By comparing the current reputation levels with the two specified thresholds, LAs can be classified into one of the three distinct groups, namely: normal, suspicious, and malicious group.

- *normal group* (\mathbb{N}): for those who reside in the normal group, we consider them as benign LAs. Thus, no further actions will be taken.
- *suspicious group* (\mathbb{S}): for those who fall into the suspicious group, reputation incentive mechanism will be triggered to adjust the monitor frequency and grades for different compliance levels.
- *malicious group* (\mathbb{M}): for those who belong to the malicious group, we consider them as malicious LAs that have been compromised by opportunistic attackers.

From a social perspective, one needs to spend a considerable amount of time performing good behaviors consistently in order to build up a good reputation, and only a few instances of bad behaviors will cause doubt on the individual's personality and result in a rapid fall in social reputation [103]. Similarly, for LAs in the suspicious group, we employ a reputation incentive mechanism to achieve adaptive assessment of their behaviors. In this mechanism, we increase the grade ω_k in response to an input of $x^j(t) = x_k$ (the full compliance level), and decrease both ω_k and ω_i responding to an input of $x^j(t) = x_i, i \neq k$. In addition, when LA_j falls in the suspicious group \mathbb{S} , control center will increase the monitor frequency of

LA_j twofold (i.e. $T_j \leftarrow T_j/2$) to pay closer attention to it. Under normal circumstances, control center monitors LA_j with a constant period T_{max} . If control center observes that LA_j behaves perfectly with all full compliance levels within a safe observation time period T_S , the monitor frequency will be reduced by half (say $T_j \leftarrow T_j * 2$). Particularly, in the case that any LA returns from group \mathbb{S} to the normal group \mathbb{N} , the monitor frequency and all the grades, with the exception of ω_k , will be recovered to the initial values. Note that the time complexity of Algorithm 5.2 is $\mathcal{O}(M)$.

In this work, we observe LA's behavior over a long period of time, rather than their entire operating history, as the latter will reduce the response speed of the reputation levels and consequently reduce the detection accuracy. Hence, we employ a relatively long observation window T_W as our reference observation period. In other words, control center only needs to assess LA's behavior within a time period of $[t - T_W, t]$.

This incentive mechanism is designed to encourage non-malicious LAs, who reside in the normal group or may fall into suspicious group due to system noise, to keep up with their good behaviors in order to increase their reputation levels, as well as rapidly decrease a suspicious LA's reputation level due to non-compliance behaviors.

5.5 Performance Evaluation

We conducted a set of experiments to evaluate the effectiveness of our proposed scheme. First, we carried out Time Step Simulation experiments using the PowerWorld simulator [105] to collect extensive real-time data from the IEEE 39-bus power testing system. Then, a series of simulations were conducted in MATLAB 2014b to analyze the collected data.

5.5.1 Data Collection in PowerWorld

The IEEE 39-bus power system, used as our testing system (see Fig. 5.3), is geographically partitioned into m areas (in our simulations, $m = 6$), which we referred to as LAs. In

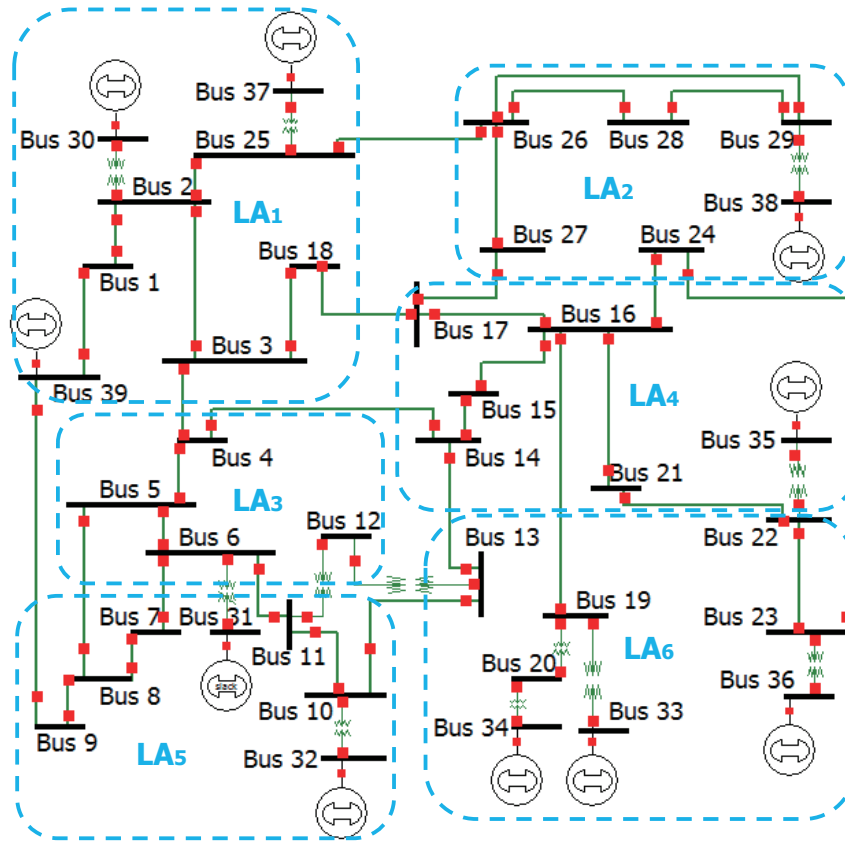


Fig. 5.3 IEEE 39-bus power system with example flocking areas

PowerWorld, we make use of Time Step Simulation to collect massive real-time data for around 20,000 minutes, including power generations of each generator G , power loads of each bus L , and line power flows of each transmission line F , etc. The first 1,500 minutes of data is used for the training phase, and the remaining data is used for the detection phase.

We randomly inserted fictitious data into the collected data to simulate the behaviors of LAs under different scalar ϵ and system noise P_n .

5.5.2 Data Analytics in MATLAB

With our proposed reputation level training algorithm, we analyze the reputation levels using the collected data. In the training phase, the effects of different ϵ and system noise P_n are first evaluated. Fig. 5.4 plots the reputation levels with respect to ϵ along the training period.

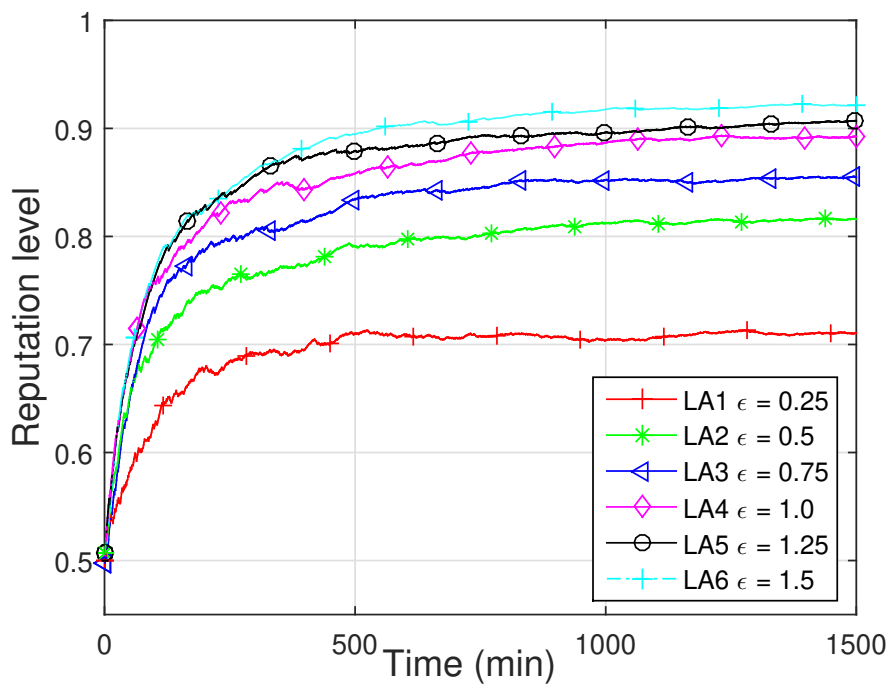


Fig. 5.4 Reputation level vs. different values of ϵ with $P_n = 0.1$

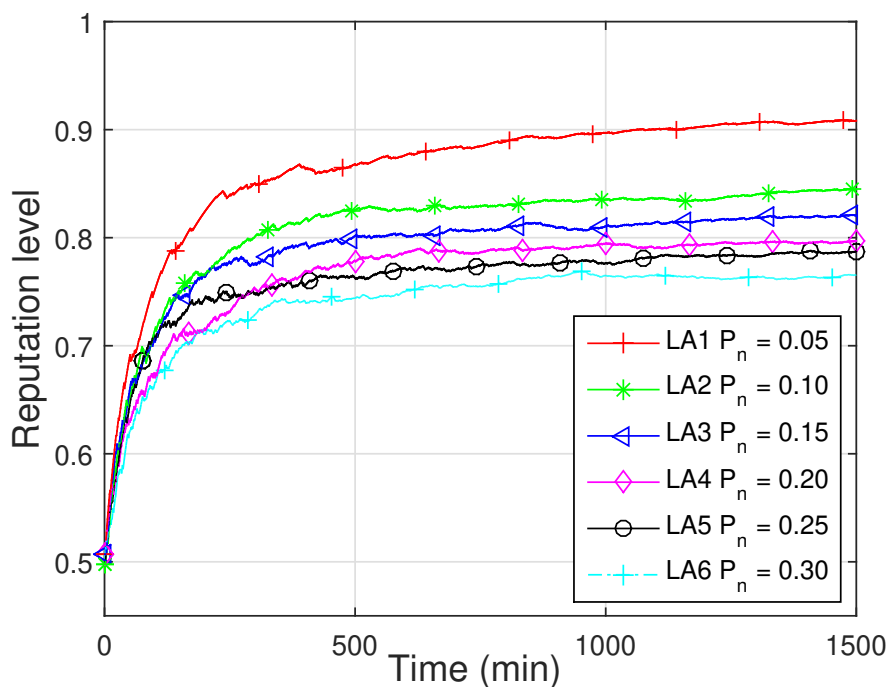


Fig. 5.5 Reputation level vs. different values of P_n with $\epsilon = 0.75$

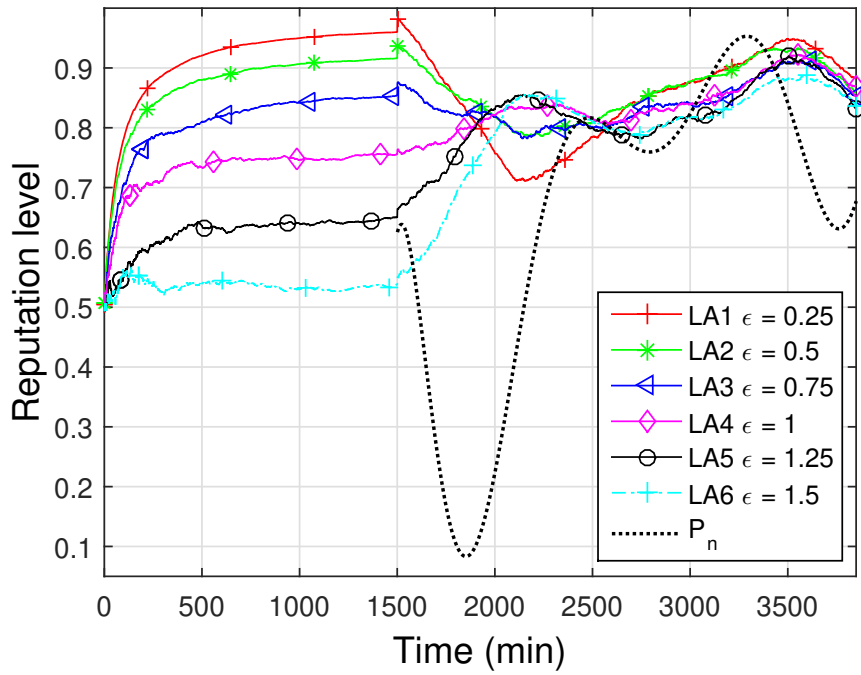


Fig. 5.6 Reputation level vs. different values of ϵ with daily dynamic P_n

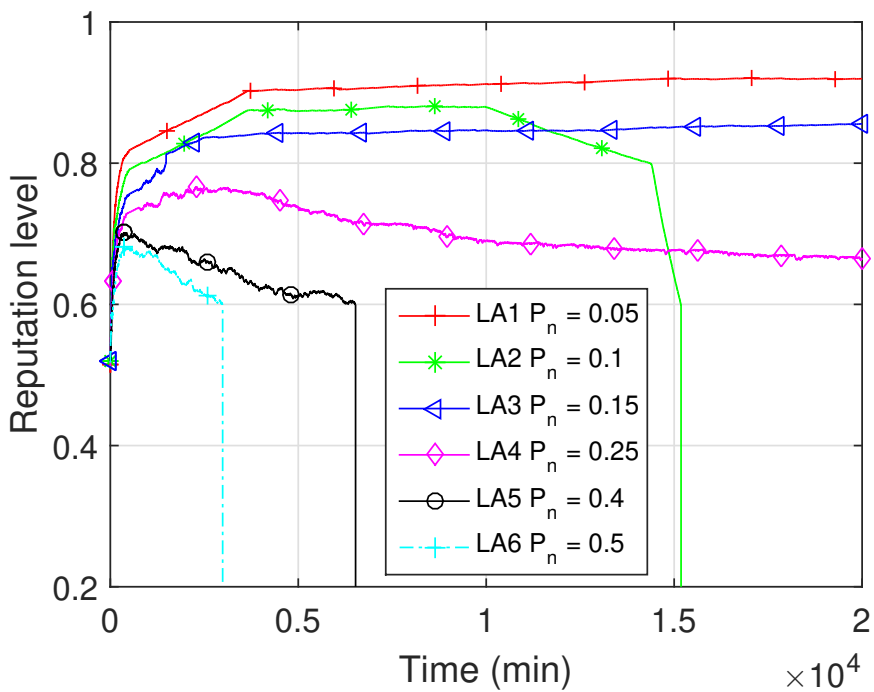


Fig. 5.7 Reputation level with an aggregative attacker with $\epsilon = 0.75$

It can be observed that the reputation level converges to a constant value as time progresses, and the higher the ϵ , the higher the reputation. This is because, as explained in Section 5.2.2, a higher ϵ indicates a lower attack probability, hence leading to a higher reputation level. The reputation levels under different system noise P_n along the training period are plotted in Fig. 5.5. Similar to the effect of ϵ , the reputation level asymptotically converges to a constant value, while the lower the system noise, the higher reputation level (recall lower system noise results in lower attack probability).

In addition, to demonstrate how opportunistic attackers can adapt their attack probabilities according to the system noise, we profile the daily system noise level based on real-time daily load pattern in Fig. 5.6. Chertkov *et al.* have demonstrated a significant correlation between system noise and load pattern in [124]. Under such circumstances, the reputation level versus system noise level under different ϵ is also presented. From this figure, we observe that the reputation level fluctuates conversely with the system noise, due to the same reason (i.e. system noise has inverse impacts on the reputation level).

In the detection phase, we study two scenarios to demonstrate the effectiveness of our proposed scheme. In the first scenario (see Fig. 5.7), we assume that at time instant 10,000 minutes, LA_2 is compromised by a malicious attacker. Since LA_2 belongs to the normal group in the beginning, we observe that after it is compromised, the reputation level decreases slightly to the suspicious group threshold H_S . With our reputation incentive mechanism, once the reputation level drops below H_S , it is regarded as suspicious and the reputation level decreases rapidly to the malicious group threshold H_M with respect to continuous non-compliance behaviors. Thus, the compromised LA_2 has been identified. By contrast, LA_5 and LA_6 are designed to be compromised from the very beginning. A notable difference is that LA_6 suffers from a higher system noise than LA_5 , and the reputation level of LA_6 decreases faster than LA_5 .

Modelling a different opportunistic attacker, we insert a temporal system fault to LA_3 at time instant 10,000 minutes in scenario two to highlight the different performance between

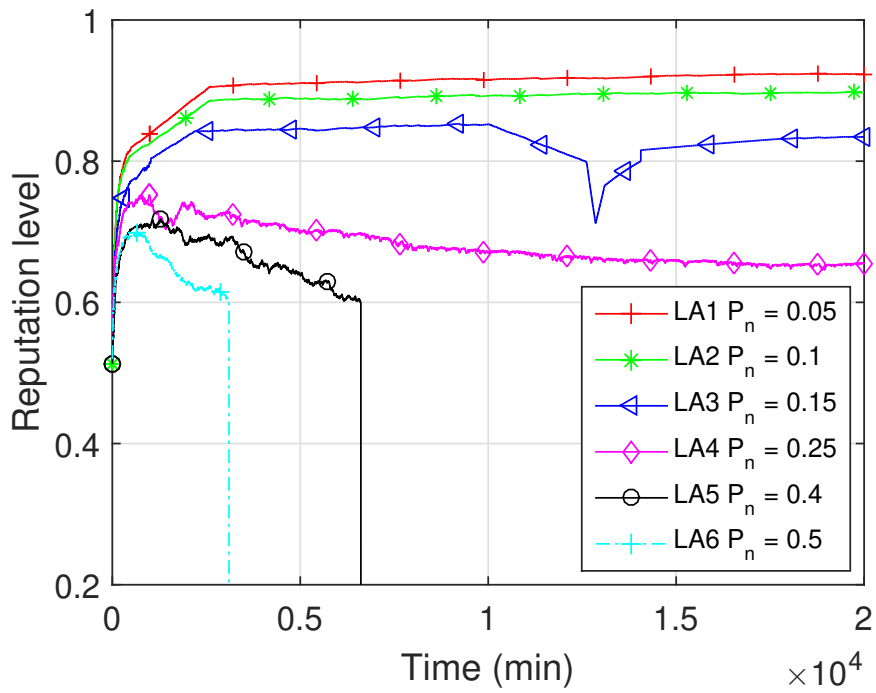


Fig. 5.8 Reputation level with an inserted temporal system fault with $\epsilon = 0.75$

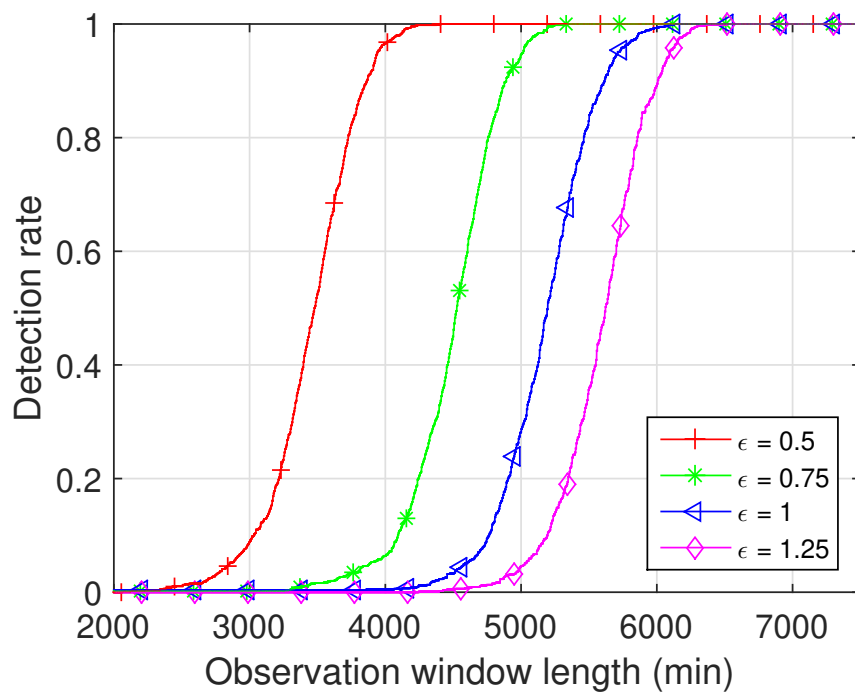


Fig. 5.9 Detection rate vs. different lengths of observation window with $P_n = 0.1$

attackers and system faults, and the corresponding reputation level variation is shown in Fig. 5.8. We observe that due to the system fault, the reputation level of LA_3 first decreases from the normal group to the suspicious group with a low decrease rate in normal group and a high decrease rate in suspicious group. This is because the proposed reputation incentive mechanism adaptively changes the decrease rate accordingly. After that, the reputation level gradually recovers and, finally, converges to a steady level. It is clear that the system fault will not change the behavior of the LA, and although the reputation decreases within a short period of time, our scheme is able to recover the reputation.

Finally, the detection rate versus the length of the observation window T_W is presented in Fig. 5.9. We observe that within a specific period (say [2000, 4000]), the detection rate increases with the growth of the observation window length, as a longer observation window can provide additional evidence to identify the hidden attackers. Compared with conservative attackers (with $\epsilon > 1$), it is quicker to identify the aggressive attackers (with $\epsilon < 1$) using our proposed scheme.

In summary, we have demonstrated that a potential class of opportunistic attackers in smart grids can adapt their attack probabilities according to the dynamic system noise level P_n , and our proposed DDOA scheme can effectively detect and identify these opportunistic attackers (e.g. state-sponsored actors). In addition, our scheme has been shown to accommodate occasional system faults due to the two specified thresholds H_s and H_m . We have also shown that our scheme achieves a high detection rate with long observation windows. Therefore, our proposal is an effective and promising solution to detect opportunistic attackers in smart grid cyber-physical systems.

5.6 Summary

In this chapter, we have presented a three-tier hierarchical framework for future smart grids, and highlighted the importance of resilience against financially-motivated opportunistic

attackers (seeking to manipulate smart electricity prices). To defend against opportunistic attacks, we have proposed a Dirichlet-based detection scheme (DDOA) to identify and detect potential attackers. Using simulations of extensive real-time data collected from the IEEE 39-bus power testing system, we demonstrated the practicality of DDOA simulations.

Chapter 6

PFDD: On Feasibility and Limitations of Detecting FmDI Attacks Using D-FACTS

Recent studies have investigated the possibilities of proactively detecting the high-profile FmDI attacks on smart grids by using the D-FACTS (distributed flexible AC transmission system) devices - the PFDD approach. However, the feasibility and limitations of such an approach have not been systematically studied in the existing literature. In this chapter, we pioneer to explore the feasibility and limitations of adopting the PFDD approach to thwart FmDI attacks on smart grids. Specifically, we thoroughly study the feasibility of using PFDD to detect FmDI attacks by considering single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. We prove that PFDD can detect all these three types of FmDI attacks targeted on buses or super-buses with degrees larger than 1, as long as the deployment of D-FACTS devices covers branches at least containing a spanning tree of the grid graph. The minimum efforts required for activating D-FACTS devices to detect each type of FmDI attacks are respectively evaluated. In addition, we also discuss the limitations of this approach, and it is strictly proved that PFDD is not able to detect FmDI attacks targeted on buses or super-buses with degrees equalling 1.

6.1 Introduction

Among the cyber threats on power grids, the high-profile FmDI attacks have drawn extensive research attentions from both power and security communities [49, 51, 28, 125, 107, 126–128]. The success of an FmDI attack is based upon attackers' knowledge of power grid connections and configurations. Unfortunately, from FmDI attackers' perspective, knowledge harvesting of power grids has been remarkably facilitated by the rapid integration of information and communications technologies as well as global proliferation of powerful hacking tools [129]. As mentioned earlier in Section 2.2, various channels can be exploited by FmDI attackers to illegally obtain valuable information of power grids.

Armed with valuable information of power grids, the knowledgeable FmDI attackers are capable of constructing attack vectors that can easily circumvent the conventional state estimation based false data detection (FDD) defenses [125, 130, 131]. This may make many of existing FDD defenses no longer feasible. We term such FDD defenses as passive approaches. A few recent studies have demonstrated the possibilities of achieving proactive FDD - termed as PFDD - in power grids by using distributed flexible AC transmission system (D-FACTS) devices [132, 133, 22]. Morrow *et al.* are presumably the pioneers to develop the idea of using D-FACTS devices to achieve topology perturbation for detecting either fault-induced or maliciously-injected bad data in the power grid [132]. Following this work, Rahman *et al.* investigated the moving target defense approaches to harden the security of the power system state estimation, one of which is to perturb the power line admittance by using D-FACTS devices [133]. More recently, Tian *et al.* proposed a hidden moving target defense approach that can maintain the power flows after changing the line susceptance, to avoid alerting the attackers who can compute the state estimation residuals [22].

Despite of these encouraging developments, some important issues of PFDD remains largely open, such as how much D-FACTS devices and where to deploy them across the power system can help with PFDD, especially when the FmDI attack strategies are evolving rapidly and appearing in a more sophisticated and coordinated mode [130]. In this paper,

we aim to systematically explore the feasibility and limitations of using PFDD to detect FmDI attacks in smart grids. We consider three types of FmDI attacks, namely single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. We adopt an arguably more realistic assumption that the adversaries can falsify the measurement data but cannot compute the state estimation residuals by themselves on real-time basis. It is based on the consensus that attackers are usually equipped with limited capabilities, unable to obtain the real-time global knowledge and measurement data of the entire power grid [28].

Despite of these encouraging developments, it remains largely an open issue how much D-FACTS can help with PFDD when the FmDI attack strategies are evolving rapidly and appearing in a more sophisticated and coordinated mode [130]. In this chapter, we aim to systematically explore the feasibility and limitations of using PFDD to detect FmDI attacks in smart grids. We consider three types of FmDI attacks, namely single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. We adopt an arguably more realistic assumption that the adversaries can falsify the measurement data but cannot compute the state estimation residuals by themselves on real-time basis. It is based on the consensus that attackers are usually equipped with limited capabilities, unable to obtain the real-time global knowledge and measurement data of the entire power grid [28].

The main contributions of this chapter are four-fold:

- First, we design a framework to detect FmDI attacks on smart grids by using the PFDD approach. The rationale behind this framework is also presented.
- Second, we explore the feasibility of using PFDD to detect the aforementioned three types of FmDI attacks on smart grids. We prove that PFDD can detect the existence of all these FmDI attacks targeted on buses or super-buses with degrees larger than 1 as long as the deployment of D-FACTS devices covers at least a spanning tree of the power grid graph.

- Third, we obtain the profiles of the minimum efforts required for D-FACTS devices to identify FmDI attacks with respect to the FmDI injected offsets on the system states, for all three types of FmDI attacks respectively. These profiles are valuable for system defenders to make appropriate efforts to counter FmDI attacks.
- Last, the limitations of using PFDD are also discussed. It is strictly proved that PFDD is unable to detect FmDI attacks targeted on buses or super-buses with degrees 1.

The remainder of this chapter is organized as follows. In Section 6.2, we present an overview of the D-FACTS devices, and detail our system model as well as the adversary model. The PFDD framework and its feasibility explorations are elaborated in Section 6.3, followed by discussions on its limitations in Section 6.4. Section 6.5 closes this chapter with the conclusion.

6.2 Overview and Models

6.2.1 Overview of D-FACTS Devices

In power grids, the flexible AC transmission system (FACTS) devices have been proven as a feasible solution to control power flows in the transmission & distribution system by altering the impedance of the power lines or changing the phase angle of the voltage applied across the lines [134]. However, high costs and reliability concerns have limited the deployments of FACTS devices. A distributed solution of FACTS, named D-FACTS, has therefore emerged and got widely accepted because of its smaller scale, lower costs, and better performance compared to FACTS devices [134]. It can be reasonably expected that D-FACTS devices will be widely deployed across smart grids in the near future due to its increasing capabilities and decreasing installation costs [22, 24].

Representatives of D-FACTS devices include distributed static series compensator (DSSC), distributed series reactor (DSR), and synchronous voltage source (SVS) [135]. These D-

FACTS devices can be used to support a myriad of applications such as contingency response, loop flow control, phase balancing, transient stability response, renewable energy transfers, etc. In this chapter, we employ D-FACTS devices with the expectation to achieve adaptable power grid configurations, which ultimately allows proactive detection of FmDI attacks.

6.2.2 System Model

In our system model, we consider the DC power flow based state estimation involving bad data detection procedure (see Fig. 6.1). Note that though we mainly focus on DC power flow model for its suitability and simplicity, we also extend our analysis to AC power flow model in Section 6.3.2 to show the universality of our studies.

The basic concepts of state estimation is introduced in Section 2.1. In this chapter, we provide more details of how the H matrix is constructed.

H Matrix Construction

Let $\mathbf{A} \in \mathbb{R}^{l \times n}$ denote the branch-bus connection matrix. l is the number of branches (power lines), and n is the number of buses that is the same as the number of system states in DC state estimation. The entries of \mathbf{A} are given by

$$a_{ki} = \begin{cases} 1, & \text{if branch } k \text{ starts at bus } i \\ -1, & \text{if branch } k \text{ ends at bus } i \\ 0, & \text{otherwise,} \end{cases} \quad (6.1)$$

where $k \in \mathcal{L} = \{1, 2, \dots, l\}$ and $i \in \mathcal{N} = \{1, 2, \dots, n\}$. In addition, let $\mathbf{D} \in \mathbb{R}^{l \times l}$ denote a diagonal matrix whose diagonal entries are the admittance (negative susceptance in DC

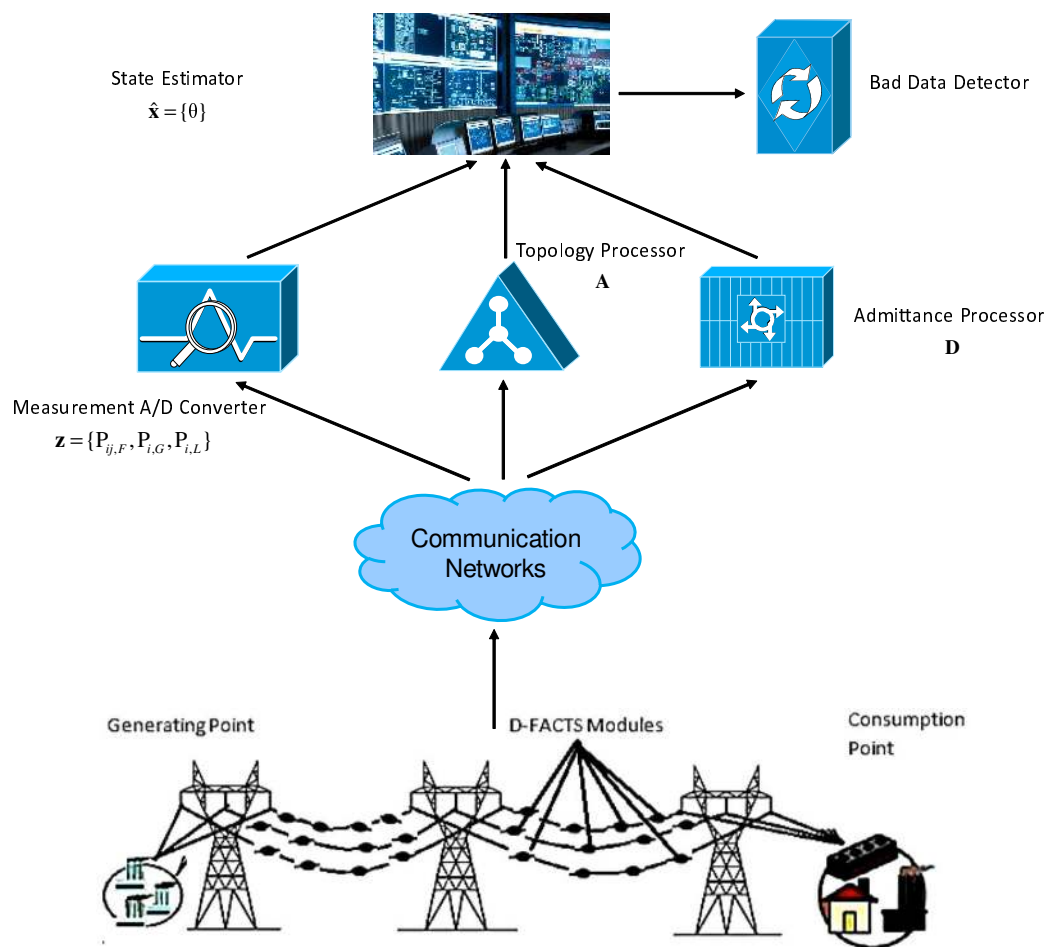


Fig. 6.1 The system model - DC state estimation in smart grids

power flow model) values of branches. Then the Jacobian matrix \mathbf{H} is constructed as [136]

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}^\top \mathbf{D} \mathbf{A} \\ \mathbf{D} \mathbf{A} \\ -\mathbf{D} \mathbf{A} \end{bmatrix}. \quad (6.2)$$

6.2.3 Adversary Model

As mentioned in Section 2.2 that, to launch a successful FmDI attack, the knowledge of \mathbf{H} matrix is demanded for the attackers. This is also described in Lemma 1.

Lemma 1. [28] *Suppose the original measurements \mathbf{z} can pass the bad measurement detection. The malicious measurements $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$ can pass the bad measurement detection if \mathbf{a} is a linear combination of the column vectors of \mathbf{H} (i.e., $\mathbf{a} = \mathbf{H}\mathbf{c}$).*

For attackers without the knowledge of \mathbf{H} matrix, it is hard for them to select an attack vector \mathbf{a} that can, fortunately, happen to lead to a success in passing the BDD test. In contrast, when it comes to knowledgeable attackers, they can choose any non-zero arbitrary vector \mathbf{c} and construct an attack vector by $\mathbf{a} = \mathbf{H}\mathbf{c}$ if they have sufficient knowledge of the \mathbf{H} matrix. As aforementioned, the recent years have seen the growing knowledge of the attackers; therefore, the existence of knowledgeable attackers are increasingly becoming a challenge that we must carefully handle. In this chapter, we consider three types of FmDI attacks in terms of the attackers' capabilities and their knowledge levels of the \mathbf{H} matrix, including

- Single-bus FmDI attacks: this type of FmDI attacks are only planned and carried out on a specific single bus, e.g., $c_i = \theta_a$ for $i \in \mathcal{N}$ and $c_j = 0$ for $\forall j \in \mathcal{N} \setminus i$, where θ_a is a constant number of voltage phase angle. The attackers are with weak attack capability and basic knowledge level of \mathbf{H} matrix (i.e., the susceptance information of one specific bus's incident branches).
- Uncoordinated multiple-bus FmDI attacks: this type of FmDI attacks can be simultaneously but independently planned and constructed on multiple buses in an uncoordinated

mode, e.g., $\mathbf{c} = (0, \underbrace{\theta_{a1}, 0, 0, \theta_{a2}, \theta_{a3}, 0, \dots}_n)^\top$, where θ_{a1} , θ_{a2} , and θ_{a3} are distinct constant numbers of voltage phase angle. The attackers are with intermediate attack capability and advanced knowledge level of \mathbf{H} matrix (i.e., the susceptance information of several buses' incident branches).

- Coordinated multiple-bus FmDI attacks (also called super-bus FmDI attacks [130]): this type of FmDI attacks can be simultaneously carried out on multiple buses in a coordinated mode, e.g., $\mathbf{c} = (\underbrace{\theta_a, \theta_a, 0, 0, \theta_a, 0, \theta_a, \dots}_n)^\top$. The attackers are with strong attack capability and expert knowledge level of \mathbf{H} matrix, i.e., the susceptance information of a super bus. Note that, a super-bus is defined as a union of multiple inter-connected buses, where all the buses united can be considered as a merged one. All the internal branches within a super-bus can be considered as omitted, and all the external branches to other buses are considered as the branches of the super-bus [130].

6.3 The Feasibility of PFDD

In this section, we study the feasibility of using PFDD approach to detect FmDI attacks on smart grids. Prior to delving into the detailed findings, we give the definition of the degree of a bus in power grids.

Definition 2. Given the graph of a power grid topology, the degree of a bus is defined as the number of connections (branches) it has to other buses. Similarly, the degree of a super-bus is defined as the number of connections (branches) the super-bus has to other buses.

Note that there is usually no isolated buses with degrees 0 in real-world power grids. In this chapter, we consider cases where the degree of any bus in a power grid is no less than 1. Our discussions contain two parts. In the first part, we show that the PFDD approach can detect FmDI attacks targeted on those buses with degrees larger than 1, of which the details will be presented later in this section. In the second part, we show that for buses with

degrees 1, the PFDD approach cannot detect FmDI attacks targeted on them; the details will be presented in Section 6.4. Note that hereafter in this section, all buses we are talking about are with degrees strictly larger than 1.

Our discussions in the rest part of this section will be presented as follows: first, we develop a framework for PFDD approach and show the rationale behind it. Then we evaluate the minimum efforts required for D-FACTS devices to identify FmDI attacks with respect to the values of injected offset on the system states. Last but most important, we formulate and prove a theorem regarding the minimum number of branches deployed with D-FACTS devices required to successfully detect FmDI attacks.

6.3.1 The Framework for PFDD Approach and Its Rationale

The PFDD approach fulfills FDD by 1) proactively activating the D-FACTS devices deployed on the transmission lines (branches), 2) updating state estimation parameters, and 3) conducting BDD process. Activating D-FACTS devices proactively changes the system configuration information, therefore affecting the state estimation; however, the attackers are incapable of following such configuration changes in a very short time. This builds up the premises for deploying the PFDD approach in smart grids to detect FmDI attacks. Note that, unlike most of the existing FDD approaches, PFDD is applied proactively regardless of whether anomaly is observed/sensed.

We design a framework for the PFDD approach in Algorithm 6.1. The rationale behind is discussed below. Assume that when D-FACTS devices are activated, the values of line admittance are altered by

$$\mathbf{D}' = \mathbf{D} + \Delta\mathbf{D}, \quad (6.3)$$

where $\Delta\mathbf{D}$ is a matrix of the line admittance variations when D-FACTS devices are activated. Accordingly, the Jacobian matrix is changed by

$$\mathbf{H}' = \begin{bmatrix} \mathbf{A}^\top \mathbf{D}' \mathbf{A} \\ \mathbf{D}' \mathbf{A} \\ -\mathbf{D}' \mathbf{A} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^\top (\mathbf{D} + \Delta\mathbf{D}) \mathbf{A} \\ (\mathbf{D} + \Delta\mathbf{D}) \mathbf{A} \\ -(\mathbf{D} + \Delta\mathbf{D}) \mathbf{A} \end{bmatrix} = \mathbf{H} + \Delta\mathbf{H}, \quad (6.4)$$

where

$$\Delta\mathbf{H} = \begin{bmatrix} \mathbf{A}^\top \Delta\mathbf{D} \mathbf{A} \\ \Delta\mathbf{D} \mathbf{A} \\ -\Delta\mathbf{D} \mathbf{A} \end{bmatrix}. \quad (6.5)$$

By conducting state estimation, the updated Frobenius norm of the normalized measurement residuals with false data injected is given by

$$\begin{aligned} \|\bar{\mathbf{r}}'_a\| &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}'_a - \mathbf{H}'\hat{\mathbf{x}}'_a)\| \\ &= \|\sqrt{\mathbf{W}^{-1}}[\mathbf{z}' + \mathbf{a} - \mathbf{H}'(\hat{\mathbf{x}}' + \Delta\mathbf{x})]\| \\ &= \|\sqrt{\mathbf{W}^{-1}}(\underbrace{\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}'}_{\text{original}} + \underbrace{\mathbf{a} - \mathbf{H}'\Delta\mathbf{x}}_{\text{injected}})\|, \end{aligned} \quad (6.6)$$

where \mathbf{z}' , $\hat{\mathbf{x}}'$, $\Delta\mathbf{x}$ are the updated measurement vector, estimated system state vector, and the injected offset on system state vector, respectively. Recall that the attackers are incapable of immediately harvesting the full knowledge of the updated Jacobian matrix \mathbf{H}' right after D-FACTS devices are activated. Hence, the attack vector is still constructed by $\mathbf{a} = \mathbf{H}\mathbf{c}$ with the original knowledge of \mathbf{H} . In this case, the reported falsified measurement data $\mathbf{z}'_a = \mathbf{z}' + \mathbf{H}\mathbf{c}$ can be easily identified as being abnormal. This is because that in most cases, vector $\sqrt{\mathbf{W}^{-1}}(\mathbf{a} - \mathbf{H}'\Delta\mathbf{x}) = \sqrt{\mathbf{W}^{-1}}(\mathbf{H}\mathbf{c} - \mathbf{H}'\Delta\mathbf{x})$, the *injected* part of Eq. (6.6), no longer equals 0. When the entry values of this vector are sufficiently large, it leads to $\|\bar{\mathbf{r}}'_a\| > \tau$ and triggers the false data alarm. Subsequent sections will provide more details on in what cases, vector $\sqrt{\mathbf{W}^{-1}}(\mathbf{a} - \mathbf{H}'\Delta\mathbf{x})$ shall be equal to 0 or not, respectively.

Algorithm 6.1 Framework for PFDD Approach

```

1: procedure
2:   1). Activate the D-FACTS devices deployed on branches of interest;
3:   2). Update  $\mathbf{D}$  matrix by Eq. (6.3);
4:   3). Update  $\mathbf{H}$  matrix by Eq. (6.4);
5:   4). Conduct state estimation by Eq. (2.4) using updated  $\mathbf{D}'$  and  $\mathbf{H}'$ ;
6:   5). Execute BDD procedure by Eq. (2.8):
7:   if  $\|\bar{\mathbf{r}}'_a\| > \tau$  then
8:     output: FmDI attack is detected.
9:   else
10:    output: No FmDI attack is detected.
11:   end if
12: end procedure

```

6.3.2 Evaluation of the Minimum Efforts Required for D-FACTS Devices to Detect Effective FmDI Attacks

We start our discussions by making the following definitions.

Definition 3. An *effective FmDI attack* is the FmDI attack that, if not detected and prevented, is capable of injecting falsified measurement data and eventually lead to impacts/changes on the power flows. In contrast, an *ineffective FmDI attack* is the FmDI attack that is capable of injecting falsified measurement data but cannot eventually lead to impacts/changes on the power flows. Note that an FmDI attack is defined as an *ineffective FmDI attack*, if the entry values of the injected offsets \mathbf{c} are within the tolerance threshold of system state errors/faults. Such small-value false data cannot lead to impacts/changes on the power grid more significant than those caused by measurement noises, and therefore can be tolerated.

Another representative example for an *ineffective FmDI attack* would be a coordinated multiple-bus FmDI attack on all the buses, i.e., $\mathbf{c} = (\theta_a, \theta_a, \dots, \theta_a)^\top$. In this case, though the attacker is capable of successfully injecting falsified measurement data $\mathbf{z}_a = \mathbf{z} + \mathbf{H}\mathbf{c}$ and leading to $\mathbf{x}_a = \mathbf{x} + \mathbf{c}$, it cannot make any impact/change on the power flows. This is because, according to Eq. (6.8), the power flow is proportional to the voltage phase difference between buses. The attack injecting a same value of voltage phase angle to all buses with no

phase difference being created between any two buses, therefore, cannot cause any impact on the power flows.

Although PFDD is theoretically feasible, it is still not clear enough in the power community whether proactively activating D-FACTS devices will contribute to potential hidden impacts or instability issues. To help provide necessary knowledge of making a consensus on this question, it is valuable to figure out the minimum efforts needed for using D-FACTS to detect *effective* FmDI attacks.

Optimization Problem Formulation under DC Model

We define the required *efforts*, resulting from activating D-FACTS devices to detect the existence of *effective* FmDI attacks, as $\|\Delta\mathbf{D}\|$, the Frobenius norm of the line admittance variations on all the branches. This optimization problem is to minimize the *efforts* subject to constraints of D-FACTS capabilities and power flow balance requirements, which is formulated by

$$\min_{\Delta\mathbf{D}} \quad \|\text{diag}(\Delta\mathbf{D})\| \quad (6.7a)$$

$$\text{s.t.} \quad \|\bar{\mathbf{r}}'_a(\Delta\mathbf{D})\| > \tau \quad (6.7b)$$

$$d_k^{\min} < \Delta d_{kk} < d_k^{\max}, \quad k \in \mathcal{L} \quad (6.7c)$$

$$P_{i,G} - P_{i,L} = \sum_{j \in \mathcal{N}_i} P'_{ij,F}, \quad i, j \in \mathcal{N}, \quad (6.7d)$$

where $\text{diag}(\cdot)$ returns a vector containing the diagonal elements of a square matrix. Since $\Delta\mathbf{D}$ is a diagonal matrix, $\|\Delta\mathbf{D}\|$ is equivalent to $\|\text{diag}(\Delta\mathbf{D})\|$. Δd_{kk} is the k -th element of vector $\text{diag}(\Delta\mathbf{D})$. $\bar{\mathbf{r}}'_a(\Delta\mathbf{D})$ denotes the updated normalized estimation residuals with false data injected, which is a function of $\Delta\mathbf{D}$. d_k^{\min} and d_k^{\max} serve as the lower and upper bounds of Δd_{kk} , respectively, implying the range of admittance variations that D-FACTS devices deployed on the k -th branch can achieve. $P_{i,G}$ and $P_{i,L}$ denote the power generations and power loads at bus i , respectively. Further, we denote the neighbour buses of bus i by \mathcal{N}_i ,

and $P'_{ij,F}$ the updated power flow between buses i and j , which in DC model is calculated by

$$P'_{ij,F} = d'_{kk}(\theta'_i - \theta'_j) = -b'_{ij}(\theta'_i - \theta'_j), \quad (6.8)$$

where θ'_i and θ'_j are the updated voltage phase angles on buses i and j . b'_{ij} is the updated susceptance of branch (i, j) , which is also the k -th branch; hence $b'_{ij} = -d'_{kk}$.

With regards to the constraints of this optimization problem, formulas (6.7c) and (6.7d) specify the capability constraints of D-FACTS devices and for the optimal power flow balance requirements, respectively. More importantly, formula (6.7b) is specified for the successful identification of FmDI attacks via the BDD procedure. The updated estimated system state vector with false data $\hat{\mathbf{x}}'_a$ being injected is equivalent to the true updated system states added by the injected offsets, which is given by

$$\hat{\mathbf{x}}'_a = \hat{\mathbf{x}}' + \Delta \mathbf{x}. \quad (6.9)$$

Also, according to Eq. (2.4), we have

$$\hat{\mathbf{x}}'_a = \Lambda' \mathbf{z}'_a = \Lambda'(\mathbf{z}' + \mathbf{a}) = \hat{\mathbf{x}}' + \Lambda' \mathbf{a}. \quad (6.10)$$

Thus, $\Delta \mathbf{x}$ can be represented

$$\Delta \mathbf{x} = \Lambda' \mathbf{a}. \quad (6.11)$$

As a result, constraint (6.7b) can be rewritten as

$$\begin{aligned} \tau &< \|\bar{\mathbf{r}}'_a(\Delta \mathbf{D})\| \\ &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + \mathbf{a} - \mathbf{H}'\Delta \mathbf{x})\| \\ &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + \mathbf{Hc} - \mathbf{H}'\Lambda'\mathbf{Hc})\| \\ &= \|\sqrt{\mathbf{W}^{-1}}[\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + (\mathbf{I} - \mathbf{H}'\Lambda')\mathbf{Hc}]\| \\ &= \|\sqrt{\mathbf{W}^{-1}}\{\mathbf{z}' - (\mathbf{H} + \Delta \mathbf{H})\hat{\mathbf{x}}' + [\mathbf{I} - (\mathbf{H} + \Delta \mathbf{H})\Lambda']\mathbf{Hc}\}\|. \end{aligned} \quad (6.12)$$

Note that for clearance purpose, we will not substitute $\Delta\mathbf{H}$ by $\Delta\mathbf{D}$, but recall that $\Delta\mathbf{D}$ fully represents the variations of $\Delta\mathbf{H}$ with reference to Eq. (6.5).

The formulated optimization problem and inequality shown in Eq. (6.12) allow us to evaluate the relationship between the minimum $\|\text{diag}(\Delta\mathbf{D})\|$ and \mathbf{c} , and obtain a general profile given a specific measurement system with original designs of \mathbf{A} , \mathbf{D} , \mathbf{W} and τ . At first glance, it seems that this relationship depends on real-time measurements \mathbf{z} and system states \mathbf{x} . However, with reference to Section 6.2.2, we notice that $\|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - (\mathbf{H} + \Delta\mathbf{H})\hat{\mathbf{x}}')\| < \tau$ holds all the time under normal circumstances. Therefore, it can be claimed that the entry values of vector $\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - (\mathbf{H} + \Delta\mathbf{H})\hat{\mathbf{x}}')$ should be sufficiently small, and thus can be reasonably neglected. In this way, by Eq. (6.12), we only need to consider

$$\tau < \|\sqrt{\mathbf{W}^{-1}}([\mathbf{I} - (\mathbf{H} + \Delta\mathbf{H})\Lambda']\mathbf{H}\mathbf{c})\|. \quad (6.13)$$

Our numerical results demonstrate the validity of this claim by showing that, the entry values of vector $\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}')$ are in the magnitude of 10^{-6} and $\sqrt{\mathbf{W}^{-1}}\{[\mathbf{I} - (\mathbf{H} + \Delta\mathbf{H})\Lambda']\mathbf{H}\mathbf{c}\}$ are usually in the magnitude 10^{-3} or higher when we randomly construct an *effective* FmDI attack. This, therefore, enables the very existence of a general profile between the minimum $\|\text{diag}(\Delta\mathbf{D})\|$ and \mathbf{c} .

Optimization Problem Formulation Under AC Model

The objective to minimize the efforts subject to constraints of D-FACTS capabilities and power flow balance requirements can also be formulated under AC power flow model, which

is given by

$$\min_{\Delta \mathbf{D}} \quad \|\Delta \mathbf{D}\| \quad (6.14a)$$

$$\text{s.t.} \quad \|\bar{\mathbf{r}}'_a(\Delta \mathbf{D})\| > \tau \quad (6.14b)$$

$$d_k^{min} < \Delta d_k < d_k^{max}, \quad k \in \mathcal{L} \quad (6.14c)$$

$$P_i = P_{i,G} - P_{i,L} = \sum_{j \in \mathcal{N}_i} P'_{ij}, \quad i, j \in \mathcal{N}, \quad (6.14d)$$

It seems the formulas are similar to Eqs. (6.7a)-(6.7d), but it should be noted that the definitions of $\Delta \mathbf{D}$ and $\bar{\mathbf{r}}'_a$ are different under AC power flow model. Specifically, $\mathbf{D} \in \mathbb{R}^{l \times 1}$ is defined as the admittance vector and, therefore, $\Delta \mathbf{D}$ is the vector of admittance variations when D-FACTS devices are activated, which is given by

$$\Delta \mathbf{D} = (\Delta d_1, \Delta d_2, \dots, \Delta d_l)^\top. \quad (6.15)$$

With regard to $\bar{\mathbf{r}}'_a$ under AC power flow model, since the measurement data $\mathbf{z}'_a = \mathbf{z}' + \mathbf{a}$ and system states \mathbf{x}'_a are related by

$$\mathbf{z}'_a = \mathbf{z}' + \mathbf{a} = \mathbf{z}' + \mathbf{h}(\mathbf{c}) = \mathbf{h}'(\mathbf{x}'_a) + \boldsymbol{\eta}, \quad (6.16)$$

when FmDI attacks are in presence and D-FACTS devices are activated, the normalized measurement residual vector $\bar{\mathbf{r}}'_a$ is then given by

$$\bar{\mathbf{r}}'_a = \sqrt{\mathbf{W}^{-1}}(\mathbf{z}'_a - \hat{\mathbf{z}}'_a) = \sqrt{\mathbf{W}^{-1}}[\mathbf{z}'_a - \mathbf{h}'(\hat{\mathbf{x}}'_a)]. \quad (6.17)$$

The current system states $\hat{\mathbf{x}}'_a$ vector is now estimated by

$$\begin{aligned}\hat{\mathbf{x}}'_a &= \min_{\mathbf{x}'_a} [\mathbf{z}'_a - \mathbf{h}'(\mathbf{x}'_a)]^\top \mathbf{W}^{-1} [\mathbf{z}'_a - \mathbf{h}'(\mathbf{x}'_a)] \\ &= \min_{\mathbf{x}'_a} [\mathbf{z}' + \mathbf{a} - \mathbf{h}'(\mathbf{x}'_a)]^\top \mathbf{W}^{-1} [\mathbf{z}' + \mathbf{a} - \mathbf{h}'(\mathbf{x}'_a)] \\ &= \sum_{i=1}^m \frac{(z'_i + h_i(\mathbf{c}) - h'_i(\mathbf{x}'_a))^2}{\sigma_i^2}.\end{aligned}\quad (6.18)$$

Note that the matrix \mathbf{h}' involves the information of vector $\Delta\mathbf{D}$, the relationship of which is highly nonlinear under AC power flow model and is hard to be clearly presented. As we can see the optimization problem to find the minimum efforts by activating D-FACTS devices to detect FmDI attacks can also be applicable to AC power flow model. Solving the optimization problem under AC power flow model however is computationally expensive because this problem is highly nonlinear.

Relationship Evaluation Between $\|\text{diag}(\Delta\mathbf{D})\|$ and \mathbf{c}

We evaluate the relationship by considering all the three types of FmDI attacks under DC power flow model. Note that our numerical results are obtained upon a 7-bus power grid (see Fig. 6.5), while the method we use to obtain the relationship, as aforementioned, applies to all power grids. Here, we solve the optimization problem by considering only activating D-FACTS devices deployed on one branch, which means only one element in $\Delta\mathbf{D}$ is non-zero. In addition, since the values of Δd_k are discrete, the searching space is rather limited within the range of $[d_k^{min}, d_k^{max}]$. It is, therefore, easy to enumerate all the possible values of Δd_k and obtain the minimum efforts in a short time.

In the first case, we consider a single-bus FmDI attack targeted on bus 2 and D-FACTS devices are deployed on branch (2, 5). Figure 6.2 shows the relationship between the minimum $|\Delta b_{25}|$ and c_2 under three measurement instants where $P_{5,L} = 130\text{MW}, 150\text{MW}, 170\text{MW}$, respectively. $|\Delta b_{25}|$ is the absolute susceptance of branch (2, 5) and c_2 is the second entry of vector \mathbf{c} . As we can see, the profiles are almost the same for different measurement

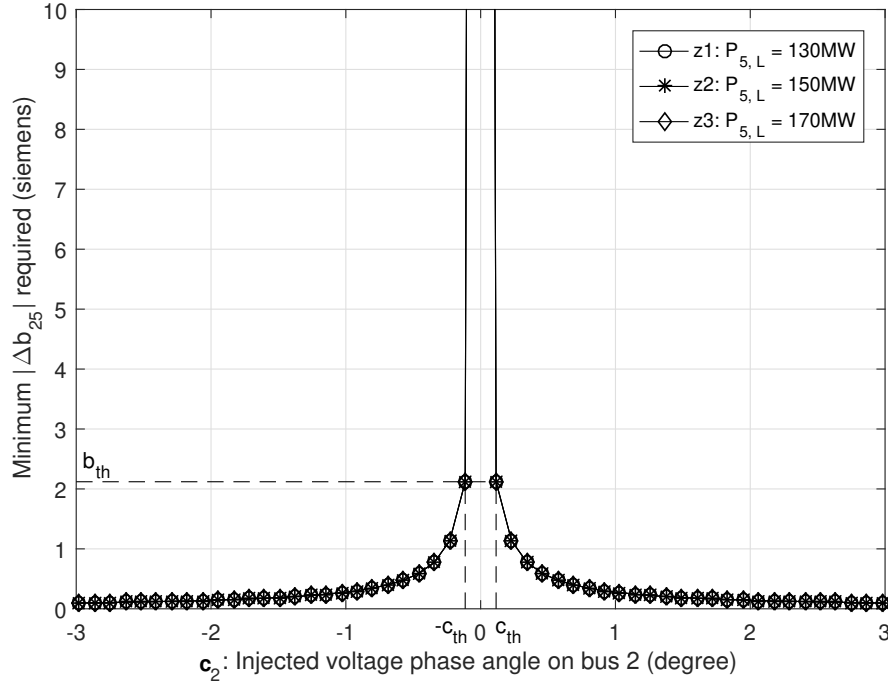


Fig. 6.2 The relationship between the minimum $|\Delta b_{25}|$ and c_2

instants. This justifies the aforementioned claim that the relationship between the minimum $\|\text{diag}(\Delta D)\|$ and c are independent of the real-time measurements z and system states x . In addition, we can also see from each profile that the larger the absolute c_2 , the lower minimum efforts are required. This indicates that it is easier for system defenders to detect FmDI attacks with reckless behaviors injecting large absolute c to expect extensive damages or profits. On the other hand, when $|c_2| < c_{th}$, either enormous efforts are required or it is impossible (beyond the adjustment capability of D-FACTS devices) to detect FmDI attacks using PFDD. Let $c_{th} > 0$ be the tolerance threshold of voltage phase angle variation, denoting the maximum value of injected voltage phase angle or measurement noises that a power grid can tolerate. The value of c_{th} can be determined by Eq. (6.12) with a given τ , and the solutions $\{c_{th}^1, c_{th}^2, \dots, c_{th}^n\}$ for different buses might be slightly different due to various configurations. For such cases, c_{th} may take the minimum solution, that is $c_{th} = \min\{c_{th}^1, c_{th}^2, \dots, c_{th}^n\}$. Correspondingly, given c_{th} , a threshold b_{th} for the minimum efforts required for D-FACTS devices to detect *effective* FmDI attacks can also be determined according to Eq. (6.12).

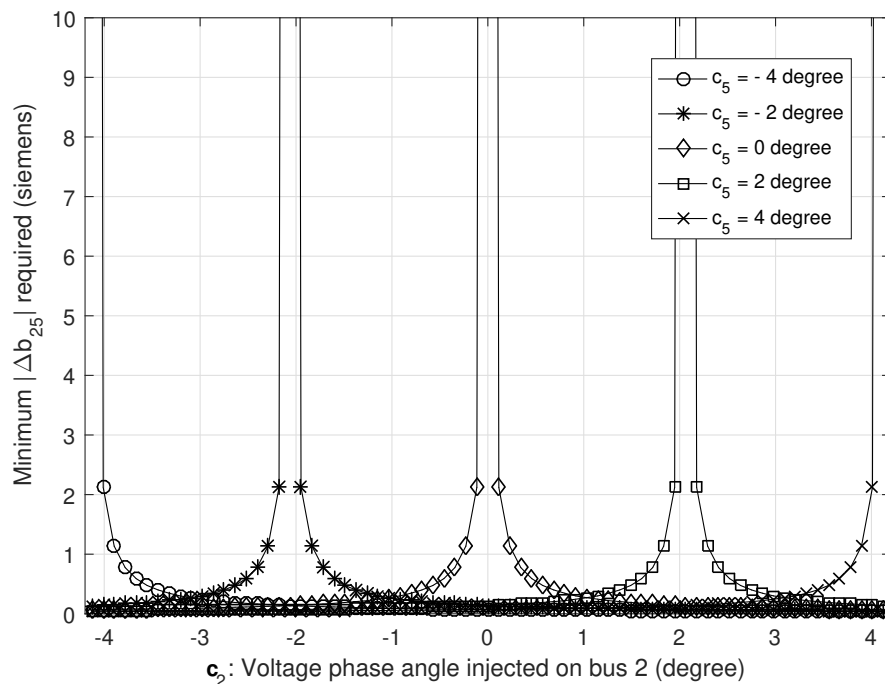


Fig. 6.3 The relationship between the minimum $|\Delta b_{25}|$ and c_2 under various values of c_5

In the second case, we consider an uncoordinated multiple-bus FmDI attack targeted on both buses 2 and 5, and branch (2, 5) is deployed with D-FACTS devices. In Fig. 6.3, we evaluate the relationship between the minimum $|\Delta b_{25}|$ and c_2 under various values of c_5 , the 5-th entry of c . As can be seen from this figure, although with different “central locations”, profiles similar to each other and to that in Fig. 6.2 are obtained under various values of c_5 . That is to say, the profile of the minimum efforts required for detecting an uncoordinated multiple-bus FmDI attack is similar to that for a single-bus FmDI attack, but the exact value is based on the injected phase difference ($c_2 - c_5$ here) between the two buses if D-FACTS devices are deployed on the branch in between.

Additionally, Fig. 6.3 also shows some numerical results for cases of detecting a coordinated multiple-bus FmDI attack by using PFDD. For those cases where the injected phase difference between c_2 and c_5 is sufficiently small (less than a tolerance threshold), infinite efforts are required; consequently, PFDD cannot identify such a coordinated multiple-bus FmDI attack by only using the D-FACTS devices deployed on branch (2, 5). Fortunately,

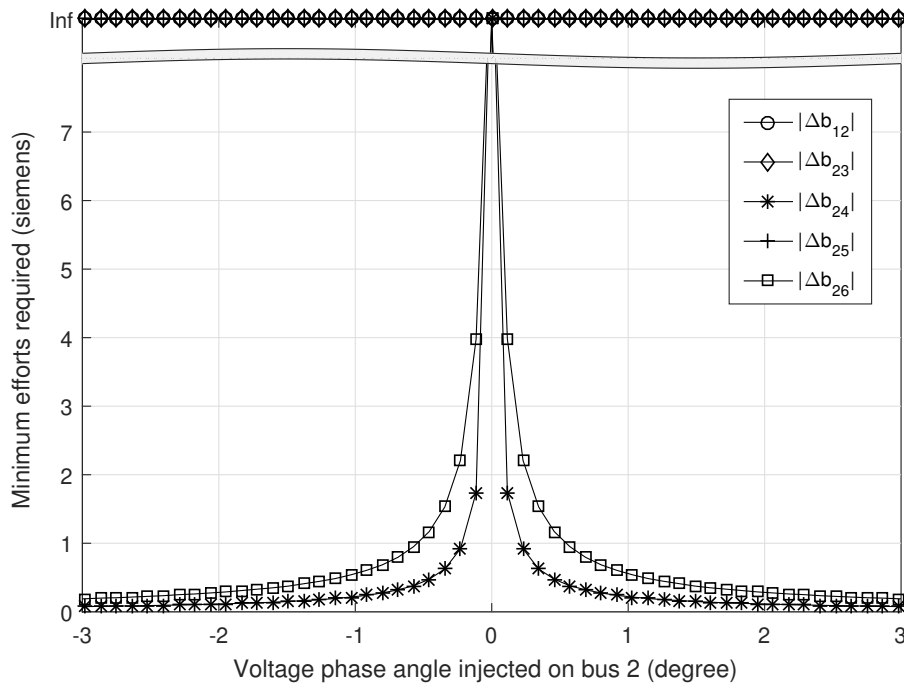


Fig. 6.4 The relationship between the minimum efforts and the injected voltage phase angle

we find that as long as any additional branch incident to the super-bus (formed by all the targeted buses in an interconnected mode, as aforementioned) is deployed with D-FACTS devices, the coordinated multiple-bus FmDI attacks can still be detected.

This finding is evidenced by the numerical results shown in Fig. 6.4. In this third case, a coordinated multiple-bus FmDI attack on buses 1, 2, 3, 5, and 7 is simulated, and suppose that D-FACTS devices are deployed on all branches incident to bus 2. As shown in Fig. 6.4, anomalies (FmDI attacks) can only be observed by activating D-FACTS devices on branches (2, 4) and (2, 6). This is because that the coordinated multiple-bus FmDI attack injects the same values of voltage phase angle (e.g. $|\theta_a| > c_{th}$) onto all the targeted buses (buses 1, 2, 3, 5, and 7 here). Hence, no injected phase difference among these coordinated buses can be observed. In contrast, sufficient difference can be observed between the un-targeted and targeted buses (e.g., between 4 and 2 or 6 and 2 here). Note that in a special case where all the buses are targeted by a coordinated multiple-bus FmDI attack, no anomaly will be

detected by using PFDD due to nonexistence of any uncoordinated bus. However, recall that such an FmDI attack is *ineffective*.

6.3.3 Minimum Deployment Requirements of D-FACTS Devices to Detect FmDI Attacks

The above discussions have shown that it is feasible to detect *effective* FmDI attacks using PFDD approach. To facilitate later discussions, we summarize this finding into Theorem 2.

Theorem 2. *In PFDD approach, D-FACTS devices deployed on a branch is able to detect the existence of effective FmDI attacks targeted on either end bus(es) (with degrees both larger than 1) of this branch, as long as the injected phase angle difference between the two end buses is larger than a tolerance threshold c_{th} .*

Proof. According to the definition of the tolerance threshold c_{th} , with a given τ for BDD test, if

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, c_{th}, 0, \dots, 0)}_n^\top, \quad (6.19)$$

then

$$\|\sqrt{\mathbf{W}^{-1}}\{\mathbf{z}' - (\mathbf{H} + \Delta\mathbf{H})\hat{\mathbf{x}}' + [\mathbf{I} - (\mathbf{H} + \Delta\mathbf{H})\Lambda']\mathbf{H}\mathbf{c}\}\| \leq \tau, \quad (6.20)$$

according to Eq. (6.12); and for any arbitrarily small positive value Δc , if

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, c_{th} + \Delta c, 0, \dots, 0)}_n^\top, \quad (6.21)$$

then

$$\|\sqrt{\mathbf{W}^{-1}}\{\mathbf{z}' - (\mathbf{H} + \Delta\mathbf{H})\hat{\mathbf{x}}' + [\mathbf{I} - (\mathbf{H} + \Delta\mathbf{H})\Lambda']\mathbf{H}\mathbf{c}\}\| > \tau. \quad (6.22)$$

In this way, let the injected phase angle difference between two ends buses $|c_0| > c_{th}$, and set

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, c_0, 0, \dots, 0)}_n^\top. \quad (6.23)$$

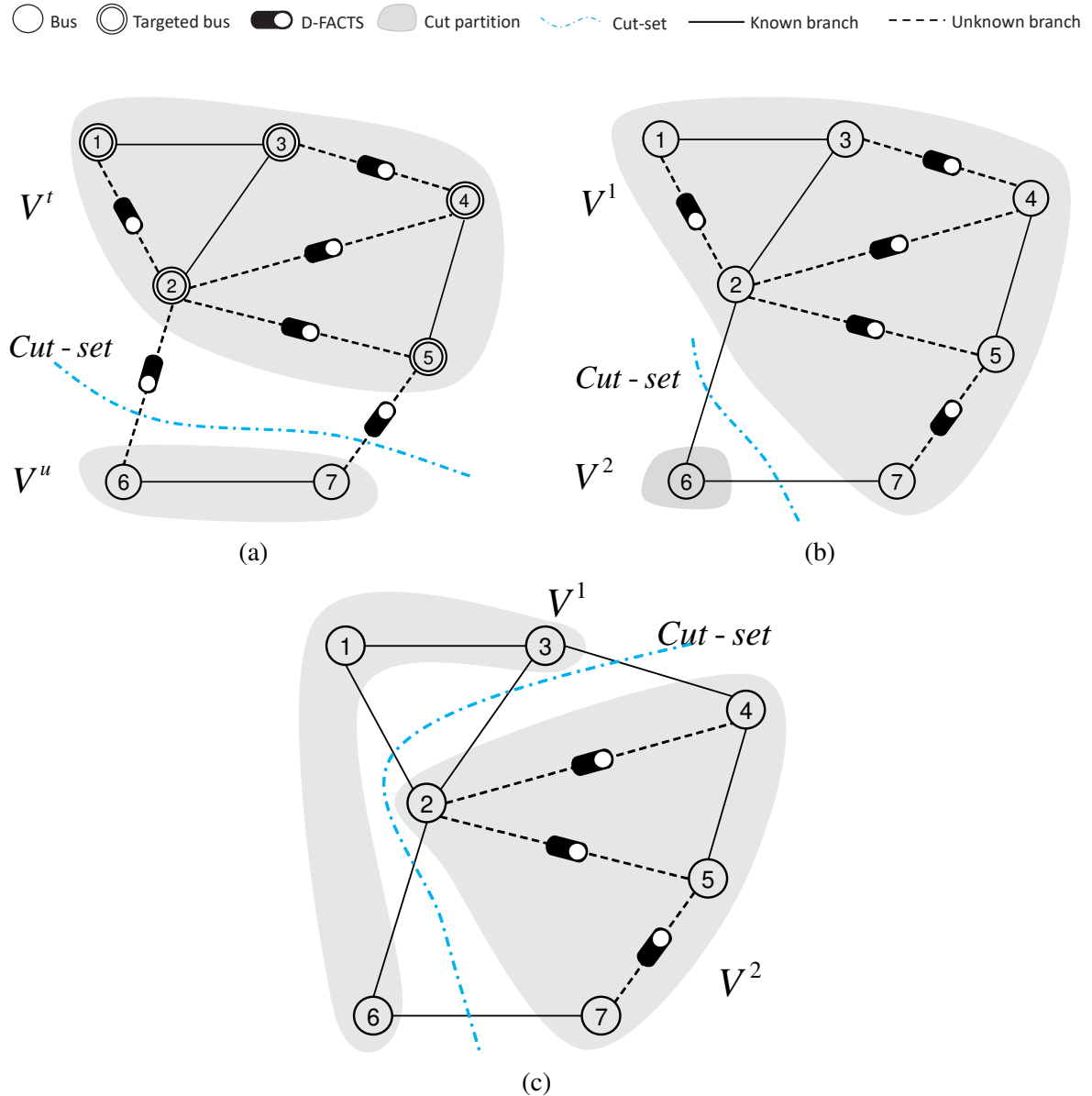


Fig. 6.5 Detection of *effective* FmDI attacks by using PFDD approach under various D-FACTS deployment strategies: (a) no *effective* FmDI attack when *unknown branches* contain a spanning tree; (b) *effective* single-bus (on V^2) or coordinated multiple-bus (on V^1) FmDI attacks when *unknown branches* fail to contain a spanning tree; and (c) *effective* single-bus (on V^1), uncoordinated multiple-bus (on V^1), or coordinated multiple-bus (on V^2) FmDI attacks when *unknown branches* fail to contain a spanning tree - less *unknown branches* compared to (b)

It is easy to obtain the same result as that in Eq. (6.22), which means that in this case the *effective* FmDI attacks can be detected.

□

In this section, we study on the minimum number of branches that need to be deployed with D-FACTS devices to guarantee the detection of all three types of *effective* FmDI attacks. A theorem shall be proposed showing that the branches installed with D-FACTS devices needs to cover at least a spanning tree of the power grid graph to ensure detection of *effective* FmDI attacks. Prior to our discussions, we make the following definitions.

Definition 4. A branch is termed as a *known branch* if its susceptance (or admittance) is unalterable and can be known to the attackers; otherwise, it is termed as an *unknown branch*.

Typically, we regard a branch deployed with D-FACTS devices as an *unknown branch* because its susceptance (or admittance) can be altered by activating D-FACTS devices; and a branch without D-FACTS devices is termed as a *known branch*.

Definition 5. A bus is termed as a *protected bus* if it is connected to at least one *unknown branch*; and an *unprotected bus* otherwise.

With these definitions, we can prove the theorem below.

Theorem 3. *The PFDD approach is feasible to detect effective FmDI attacks targeted on buses or super-buses with degrees larger than 1, if and only if the unknown branches cover at least a spanning tree of the power grid graph.*

Proof. Sufficiency: Suppose that a set of $n - 1$ branches building a spanning tree \mathcal{T} of the power grid graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ are deployed with D-FACTS devices. According to Definitions 4 and 5, these $n - 1$ branches are *unknown branches*, and all buses are *protected buses* as each of them is connected to at least one of these *unknown branches*. In this case, for any form of *effective* single-bus or uncoordinated multiple-bus FmDI attacks, there must be at least one *unknown branch* connecting to the targeted bus(es). According to Theorem 2,

it is feasible for system defenders to detect these FmDI attacks by using PFDD with given *unknown branches*. When it comes to *effective* coordinated multiple-bus FmDI attacks, at most $n - 1$ buses are targeted in such an attack, leaving at least one bus un-targeted. Thus, there must be a *cut* $\mathcal{C} = \{\mathcal{V}^t, \mathcal{V}^u\}$ that divides the buses in a grid graph into two sets - targeted buses set \mathcal{V}^t and un-targeted buses set \mathcal{V}^u , where $\mathcal{V}^t \cup \mathcal{V}^u = \mathcal{V}$. The *cut-set* of \mathcal{C} contains edges that have one endpoint in \mathcal{V}^t and the other in \mathcal{V}^u . Given the *unknown branches* contain a spanning tree (as an example shown in Fig. 6.5a), the *cut-set* must involve at least one *unknown branch* for any form of *effective* coordinated multiple-bus FmDI attacks. With these *unknown branches*, it is feasible to detect *effective* coordinated multiple-bus FmDI attacks by using PFDD (see Theorem 2).

Necessity: If *unknown branches* in a power grid do not contain a spanning tree, there must be a *cut* $\mathcal{C} = \{\mathcal{V}^1, \mathcal{V}^2\}$ that divides the buses in a grid graph into two sets \mathcal{V}^1 and \mathcal{V}^2 , and the *cut-set* of \mathcal{C} involves no *unknown branch*. Then a coordinated multiple-bus FmDI attack on all buses in \mathcal{V}^1 but none in \mathcal{V}^2 , or on all buses in \mathcal{V}^2 but none in \mathcal{V}^1 will not be detected by the PFDD approach.

□

Figures 6.5b and 6.5c show examples of FmDI attacks when *unknown branches* fail to contain a spanning tree. In the case as shown in Fig. 6.5b where *unknown branches* fail to contain a spanning tree, buses in partition \mathcal{V}^1 can be targeted by *effective* coordinated multiple-bus FmDI attacks and buses in partition \mathcal{V}^2 can be targeted by *effective* single-bus FmDI attacks. Both attacks will not be detected. In an even worse case as shown in Fig. 6.5c where fewer *unknown branches* are deployed compared to that in Fig. 6.5b, buses in partition \mathcal{V}^2 can be targeted by *effective* coordinated multiple-bus FmDI attacks and buses in partition \mathcal{V}^1 can be targeted by either *effective* single-bus or *effective* uncoordinated multiple-bus FmDI attacks. Again, none of these attacks will be detected. It can, therefore, be concluded that *unknown branches* failing to contain a spanning tree in a grid graph leave opportunities for *effective* FmDI attacks to be successful.

6.4 Discussions on PFDD Limitations

In this section, we shall discuss on the limitations of using PFDD to detect *effective* FmDI attacks targeting on buses or super-buses with degrees 1.

6.4.1 Limitations of Detecting Effective FmDI Attacks Using PFDD

In this subsection, we show our findings regarding the limitations of using PFDD to detect *effective* FmDI attacks targeting on buses or super-buses with degrees 1, which are summarized in Theorem 4 and Corollary 1.

Theorem 4. *The PFDD approach is not able to detect effective FmDI attacks targeting on buses or super-buses with degrees 1.*

Proof. Let $\epsilon_k \in \{0, 1\}^{l \times 1}$ denote a unit column vector whose k -th entry equals 1, and $\delta_i \in \{0, 1\}^{n \times 1}$ a unit column vector whose i -th entry equals 1. Define $\mu_{ij} \triangleq \delta_i - \delta_j$. In this way, matrices \mathbf{A} and \mathbf{D} in Section 6.2 can be written as

$$\mathbf{A} = \sum_{k \in \mathcal{L}, k \sim \{i, j\}} \epsilon_k \mu_{ij}^\top, \quad \mathbf{D} = \sum_{k \in \mathcal{L}, k \sim \{i, j\}} -b_{ij} \epsilon_k \epsilon_k^\top, \quad (6.24)$$

where $k \sim \{i, j\}$ denotes branch k that connects buses i and j . Then \mathbf{DA} and $\mathbf{A}^\top \mathbf{DA}$ are respectively given by

$$\mathbf{DA} = \sum_{k \in \mathcal{L}, k \sim \{i, j\}} -b_{ij} \epsilon_k \mu_{ij}^\top, \quad (6.25)$$

and

$$\mathbf{A}^\top \mathbf{DA} = \sum_{k \in \mathcal{L}, k \sim \{i, j\}} -b_{ij} \mu_{ij} \mu_{ij}^\top. \quad (6.26)$$

Let $\boldsymbol{\rho}_i \in \{0, 1\}^{(n+2l) \times 1}$ denote a unit column vector whose i -th entry equals 1. Based on Eq. (6.2), the \mathbf{H} matrix can be written as

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}^\top \mathbf{D} \mathbf{A} \\ \mathbf{D} \mathbf{A} \\ -\mathbf{D} \mathbf{A} \end{bmatrix} = \begin{bmatrix} \sum_{k \in \mathcal{L}, k \sim \{i,j\}} -b_{ij} \boldsymbol{\mu}_{ij} \boldsymbol{\mu}_{ij}^\top \\ \sum_{k \in \mathcal{L}, k \sim \{i,j\}} -b_{ij} \boldsymbol{\epsilon}_k \boldsymbol{\mu}_{ij}^\top \\ \sum_{k \in \mathcal{L}, k \sim \{i,j\}} b_{ij} \boldsymbol{\epsilon}_k \boldsymbol{\mu}_{ij}^\top \end{bmatrix} = \sum_{k \in \mathcal{L}, k \sim \{i,j\}} -b_{ij} (\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k}) \boldsymbol{\mu}_{ij}^\top. \quad (6.27)$$

For a single bus with degree 1: Suppose that an *effective* single-bus FmDI attack is targeted at bus $\zeta \in \mathcal{N}$ with degree 1, and bus $\gamma \in \mathcal{N}$ is the only neighbour of bus ζ connected by branch $\ell \in \mathcal{L}$. The attacker aims to inject θ_a to bus ζ by designing

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, \overbrace{\theta_a}^{\zeta\text{-th}}, 0, \dots, 0)^\top}_n, \quad (6.28)$$

which can be rewritten as $\mathbf{c} = \theta_a \boldsymbol{\delta}_\zeta$. In this case, the attack vector \mathbf{a} is written as

$$\begin{aligned} \mathbf{a} = \mathbf{H} \mathbf{c} &= -b_{\zeta\gamma} (\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}) \boldsymbol{\mu}_{\zeta\gamma}^\top \theta_a \boldsymbol{\delta}_\zeta \\ &= -b_{\zeta\gamma} \theta_a (\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}) (\boldsymbol{\delta}_\zeta^\top - \boldsymbol{\delta}_\gamma^\top) \boldsymbol{\delta}_\zeta \\ &= -b_{\zeta\gamma} \theta_a (\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}). \end{aligned} \quad (6.29)$$

If D-FACTS devices deployed on branch ℓ are activated, the susceptance of this branch is updated to $b'_{\zeta\gamma}$ and the \mathbf{H} matrix is updated to \mathbf{H}' . Then, we have the following major finding:

$$\mathbf{a} = \mathbf{H} \mathbf{c} = \mathbf{H}' \mathbf{c}' = -b'_{\zeta\gamma} \theta'_a (\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}), \quad (6.30)$$

where

$$\mathbf{c}' = \underbrace{(0, 0, \dots, 0, \overbrace{\theta'_a}^{\zeta\text{-th}}, 0, \dots, 0)^\top}_n, \text{ and } \theta'_a = \frac{b_{\zeta\gamma} \theta_a}{b'_{\zeta\gamma}}. \quad (6.31)$$

Based on Eqs. (6.6) and (6.11), $\|\bar{\mathbf{r}}'_a(\Delta\mathbf{D})\|$ can be rewritten as

$$\begin{aligned}
 \|\bar{\mathbf{r}}'_a(\Delta\mathbf{D})\| &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + \mathbf{a} - \mathbf{H}'\Delta\mathbf{x})\| \\
 &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + \mathbf{H}'\mathbf{c}' - \mathbf{H}'\mathbf{\Lambda}'\mathbf{H}'\mathbf{c}')\| \\
 &\stackrel{\mathbf{\Lambda}'\mathbf{H}'=\mathbf{I}}{=} \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}' + \mathbf{H}'\mathbf{c}' - \mathbf{H}'\mathbf{c}')\| \\
 &= \|\sqrt{\mathbf{W}^{-1}}(\mathbf{z}' - \mathbf{H}'\hat{\mathbf{x}}')\| < \tau.
 \end{aligned} \tag{6.32}$$

It means that no FmDI alarm will be triggered if using PFDD to detect *effective* FmDI attacks targeting on single buses with degrees 1.

For a super-bus with degree 1: Suppose that an *effective* coordinated multiple-bus FmDI attack is targeted at buses $\mathcal{S} = \{\zeta, \zeta + 1, \dots, \zeta + t\}$, where t is a positive integer. These buses form into a super-bus with degree 1, and branch ℓ is the only external branch of this super-bus connecting buses from ζ to γ . The attacker aims to inject θ_a to this super-bus by designing

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, \overbrace{\theta_a}^{\zeta\text{-th}}, \overbrace{\theta_a}^{(\zeta+1)\text{-th}}, \dots, \overbrace{\theta_a}^{(\zeta+t)\text{-th}}, 0, \dots, 0)}_n^\top, \tag{6.33}$$

which can be rewritten as $\mathbf{c} = \theta_a \sum_{i=0}^t \delta_{\zeta+i}$. In this case, the attack vector \mathbf{a} is written as

$$\begin{aligned}
 \mathbf{a} = \mathbf{H}\mathbf{c} &= \left(\sum_{k \in \mathcal{L}^{\mathcal{S}}, k \sim \{i,j\}} -b_{ij}(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k})\boldsymbol{\mu}_{ij}^\top \right) \times \left(\theta_a \sum_{i=0}^t \delta_{\zeta+i} \right) \\
 &= \left(\sum_{k \in \mathcal{L}^{\mathcal{S}}, k \sim \{i,j\}} -b_{ij}\theta_a(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k})(\boldsymbol{\delta}_i^\top - \boldsymbol{\delta}_j^\top) \right) \times \left(\sum_{i=0}^t \delta_{\zeta+i} \right),
 \end{aligned} \tag{6.34}$$

where $\mathcal{L}^{\mathcal{S}}$ denotes the branches incident to any of the buses in set \mathcal{S} . It is worth noting that $\forall i, j$ satisfying $k \sim \{i, j\}$ and $k \in \mathcal{L}^{\mathcal{S}}$, there must have $i, j \in \mathcal{S}$ except for one case where

$i = \zeta$ and $j = \gamma$. In this way, Eq. (6.34) can be rewritten as

$$\begin{aligned}
\mathbf{a} &= \left(\sum_{k \in \mathcal{L}^S, k \sim \{i,j\}} -b_{ij}\theta_a(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k}) \times (\boldsymbol{\delta}_i^\top - \boldsymbol{\delta}_j^\top) \right) \left(\sum_{i=0}^t \boldsymbol{\delta}_{\zeta+i} \right) \\
&= \left(\sum_{k \in \{\mathcal{L}^S \setminus \ell\}, k \sim \{i,j\}} -b_{ij}\theta_a(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k}) \times (\boldsymbol{\delta}_i^\top - \boldsymbol{\delta}_j^\top)(\boldsymbol{\delta}_i + \boldsymbol{\delta}_j) \right) \\
&\quad + \left(\sum_{k=\ell, \ell \sim \{\zeta,\gamma\}} -b_{\zeta\gamma}\theta_a \times (\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+\ell} - \boldsymbol{\rho}_{n+l+\ell}) \times (\boldsymbol{\delta}_\zeta^\top - \boldsymbol{\delta}_\gamma^\top)\boldsymbol{\delta}_\zeta \right) \\
&= \left(\sum_{k \in \{\mathcal{L}^S \setminus \ell\}, k \sim \{i,j\}} -b_{ij}\theta_a(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j + \boldsymbol{\rho}_{n+k} - \boldsymbol{\rho}_{n+l+k}) \times 0 \right) \\
&\quad + \left(-b_{\zeta\gamma}\theta_a(\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+\ell} - \boldsymbol{\rho}_{n+l+\ell}) \times 1 \right) \\
&= -b_{\zeta\gamma}\theta_a(\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+\ell} - \boldsymbol{\rho}_{n+l+\ell}).
\end{aligned} \tag{6.35}$$

We have the same conclusion as that shown in Eq. (6.29). Hence, we see that for an *effective* coordinated multiple-bus FmDI attack targeted on a super-bus with degree 1, we also have $\mathbf{H}\mathbf{c} = \mathbf{H}'\mathbf{c}'$, leading to failure of detecting such an FmDI attack using PFDD approach. Note that although this FmDI attack remains undetected, it is equivalent to another FmDI attack with

$$\mathbf{c}' = \underbrace{(0, 0, \dots, 0, \overbrace{\theta'_a}^{\zeta\text{-th}}, \overbrace{\theta'_a}^{(\zeta+1)\text{-th}}, \dots, \overbrace{\theta'_a}^{(\zeta+t)\text{-th}}, 0, \dots, 0)}_n^\top, \tag{6.36}$$

where $\theta'_a = b_{\zeta\gamma}\theta_a/b'_{\zeta\gamma}$ when the susceptance of branch $\ell \sim \{\zeta, \gamma\}$ is $b'_{\zeta\gamma}$. \square

Corollary 1. *Given a single bus or a super-bus ζ with degree 1 (as for a super-bus ζ represents the bus having the external branch), the external incident bus is denoted by γ , and the branch connecting these two buses is denoted by ℓ . Without knowing the susceptance of branch ℓ , as long as FmDI attackers can inject P_a to P_ζ , $-P_a$ to P_γ , and P_a to $P_{\zeta\gamma}$, this FmDI attack cannot be detected by using PFDD, where P_ζ , P_γ , $P_{\zeta\gamma}$, and P_a denote the nodal power injections of bus ζ , nodal power injections of bus γ , power flow of branch $\ell \sim \{\zeta, \gamma\}$, and a constant power value, respectively.*

Proof. Recall that \mathbf{z} is an $m \times 1 = (n + 2l) \times 1$ column vector comprising n nodal power injections $\mathbf{P}_I = \{P_1, P_2, \dots, P_n\}$ and $2l$ power flows $\mathbf{P}_F = \{P_{ij}|i, j \in \mathcal{N}, k \sim \{i, j\}, k \in \mathcal{L}\}$ and $-\mathbf{P}_F = \{-P_{ij}|i, j \in \mathcal{N}, k \sim \{i, j\}, k \in \mathcal{L}\}$. Then, \mathbf{z} can be represented by

$$\mathbf{z} = (\mathbf{P}_I, \mathbf{P}_F, -\mathbf{P}_F)^\top = \sum_{i=1}^n P_i \boldsymbol{\rho}_i + \sum_{k \in \mathcal{L}, k \sim \{i, j\}} P_{ij} \boldsymbol{\rho}_{n+k} + \sum_{k \in \mathcal{L}, k \sim \{i, j\}} -P_{ij} \boldsymbol{\rho}_{n+l+k}. \quad (6.37)$$

If FmDI attackers can inject P_a to P_ζ , $-P_a$ to P_γ , and P_a to $P_{\zeta\gamma}$, this means that the attacker can construct an attack vector

$$\mathbf{a} = P_a(\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}). \quad (6.38)$$

Based on Eq. (6.30), we can rewrite \mathbf{a} as

$$\mathbf{a} = \mathbf{H}\mathbf{c} = \mathbf{H}'\mathbf{c}' = P_a(\boldsymbol{\rho}_\zeta - \boldsymbol{\rho}_\gamma + \boldsymbol{\rho}_{n+l} - \boldsymbol{\rho}_{n+l+l}). \quad (6.39)$$

If D-FACTS devices are activated by using PFDD and $b_{\zeta\gamma}$ is updated to $b'_{\zeta\gamma}$,

$$\mathbf{c}' = \underbrace{(0, 0, \dots, 0, \overbrace{\theta'_a}^{\zeta\text{-th}}, 0, \dots, 0)}_n^\top, \text{ and } \theta'_a = \frac{P_a}{b'_{\zeta\gamma}} \quad (6.40)$$

for a single bus ζ with degree 1, and

$$\mathbf{c}' = \underbrace{(0, 0, \dots, 0, \overbrace{\theta'_a}^{\zeta\text{-th}}, \overbrace{\theta'_a}^{(\zeta+1)\text{-th}}, \dots, \overbrace{\theta'_a}^{(\zeta+t)\text{-th}}, 0, \dots, 0)}_n^\top, \quad (6.41)$$

for a super bus ζ with degree 1, comprising buses $\zeta, \zeta + 1, \dots$, and $\zeta + t$. According to Eq. (6.32), we see that such an FmDI attack targeted on either a single bus or a super-bus with degree 1 cannot be detected by using PFDD. It is worth noting that though such FmDI attacks would be successful even without the knowledge of $b_{\zeta\gamma}$, the attackers under such cases however shall have no idea of the real injected phase angle offset θ'_a , because they

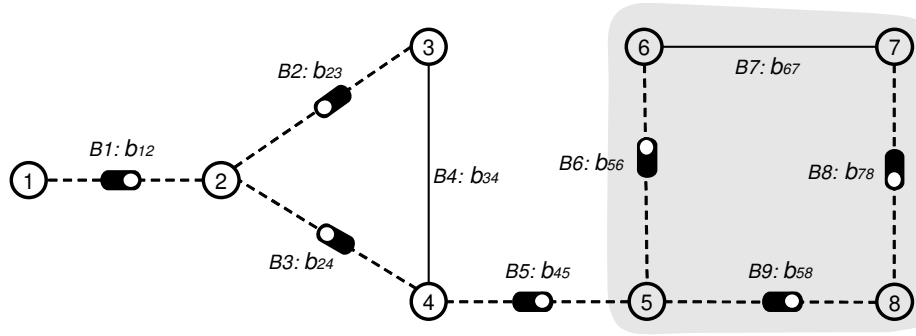


Fig. 6.6 An illustrative 8-bus power system with D-FACTS deployment covering a spanning tree.

cannot immediately obtain the value of $b'_{\zeta\gamma}$ after D-FACTS devices are activated. In other words, they do not know how much impact they can cause or how many profits they can obtain through a successful FmDI attack. \square

6.4.2 Case Study

In this subsection, we take 8-bus and 39-bus power systems (see Figs. 6.6 and 6.7) as examples to illustrate *effective* FmDI attacks targeting on a single bus and a super-bus with degree 1, respectively.

Case 1: An Effective FmDI Attack Targeted on Bus 1 in An 8-Bus System

Suppose that an FmDI attacker aims to inject θ_a to bus 1 phase angle θ_1 , he/she constructs \mathbf{c} by

$$\mathbf{c} = \theta_a \boldsymbol{\delta}_1 = (\theta_a, 0, 0, 0, 0, 0, 0, 0)^\top, \quad (6.42)$$

and an attack vector \mathbf{a} by

$$\mathbf{a} = \mathbf{H}\mathbf{c} = -b_{12}\theta_a(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2 + \boldsymbol{\rho}_{n+1} - \boldsymbol{\rho}_{n+l+1}). \quad (6.43)$$

Then, the measurement data \mathbf{z} is falsified by $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{a}$. In this case, data falsifications are equivalent to adding $-b_{12}\theta_a$ to P_1 , $b_{12}\theta_a$ to P_2 , $-b_{12}\theta_a$ to P_{12} , and $b_{12}\theta_a$ to P_{21} via

compromised meters. When system defenders activating D-FACTS devices deployed on branch $B1$, b_{12} is changed to b'_{12} . According to Eqs. (6.30) and (6.32), this FmDI attack cannot be detected by PFDD, but an offset of $\theta'_a = b_{12}\theta_a/b'_{12}$ other than θ_a is injected to θ_1 . This is equivalent to an FmDI attack with

$$\mathbf{c} = (\theta'_a, 0, 0, 0, 0, 0, 0, 0)^\top, \quad (6.44)$$

when the susceptance of branch $B1$ is b'_{12} .

Case 2: An Effective FmDI Attack Targeted on A Super-Bus Composed of Buses 5, 6, 7, and 8 in An 8-Bus System

Suppose that an FmDI attacker aims to inject θ_a to the phase angles of a super-bus composed of buses 5, 6, 7, and 8, he/she constructs \mathbf{c} by

$$\mathbf{c} = \theta_a(\boldsymbol{\delta}_5 + \boldsymbol{\delta}_6 + \boldsymbol{\delta}_7 + \boldsymbol{\delta}_8) = (0, 0, 0, 0, \theta_a, \theta_a, \theta_a, \theta_a)^\top, \quad (6.45)$$

and an attack vector \mathbf{a} by

$$\mathbf{a} = \mathbf{H}\mathbf{c} = -b_{45}\theta_a(\boldsymbol{\rho}_5 - \boldsymbol{\rho}_4 + \boldsymbol{\rho}_{n+5} - \boldsymbol{\rho}_{n+l+5}). \quad (6.46)$$

Then, the measurement data \mathbf{z} is falsified by $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{a}$. In this case, data falsifications are equivalent to adding $-b_{45}\theta_a$ to P_5 , $b_{45}\theta_a$ to P_4 , $-b_{45}\theta_a$ to P_{54} , and $b_{45}\theta_a$ to P_{45} via compromised meters. When system defenders activate the D-FACTS devices deployed on branch $B5$, b_{45} is changed to b'_{45} . According to Eqs. (6.35) and (6.32), this FmDI attack cannot be detected by PFDD and an offset of $\theta'_a = b_{45}\theta_a/b'_{45}$ is injected to $\theta_5, \theta_6, \theta_7$, and θ_8 , respectively. This is equivalent to an FmDI attack with

$$\mathbf{c} = (0, 0, 0, 0, \theta'_a, \theta'_a, \theta'_a, \theta'_a)^\top, \quad (6.47)$$

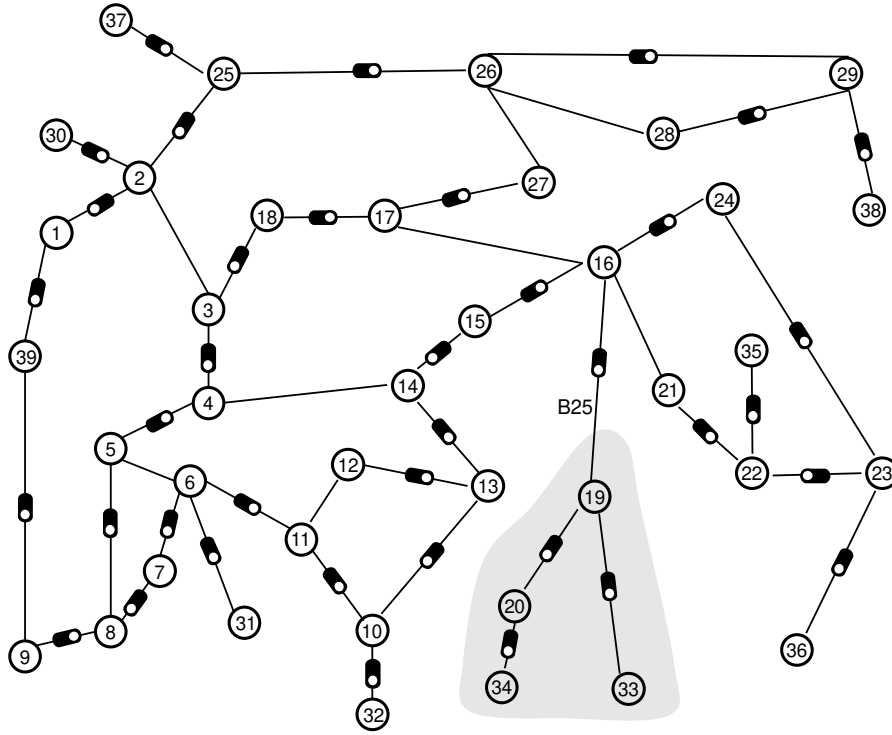


Fig. 6.7 An illustrative 39-bus power system with D-FACTS deployment covering a spanning tree.

when the susceptance of branch $B5$ is b'_{45} .

Case 3: An Effective FmDI Attack Targeted on A Super-Bus Composed of Buses 19, 20, 33, and 34 in IEEE 39-Bus System

Suppose that an FmDI attacker aims to inject θ_a to the phase angles of a super-bus composed of buses 19, 20, 33, and 34, he/she constructs \mathbf{c} by

$$\mathbf{c} = \theta_a(\boldsymbol{\delta}_{19} + \boldsymbol{\delta}_{20} + \boldsymbol{\delta}_{33} + \boldsymbol{\delta}_{34}) = \underbrace{(0, 0, \dots, 0, \overbrace{\theta_a}^{19\text{-th}}, \overbrace{\theta_a}^{20\text{-th}}, 0, \dots, 0, \overbrace{\theta_a}^{33\text{-th}}, \overbrace{\theta_a}^{34\text{-th}})}_{39}^\top, \quad (6.48)$$

and an attack vector \mathbf{a} by

$$\mathbf{a} = \mathbf{H}\mathbf{c} = -b_{16,19}\theta_a(\boldsymbol{\rho}_{19} - \boldsymbol{\rho}_{16} + \boldsymbol{\rho}_{n+25} - \boldsymbol{\rho}_{n+l+25}), \quad (6.49)$$

where 25 is the index of branch B_{25} that connects buses 16 and 19. Then, the measurement data \mathbf{z} is falsified by $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{a}$. In this case, data falsifications are equivalent to adding $-b_{16,19}\theta_a$ to P_{19} , $b_{16,19}\theta_a$ to P_{16} , $-b_{16,19}\theta_a$ to $P_{19,16}$, and $b_{16,19}\theta_a$ to $P_{19,16}$ via compromised meters. When system defenders activate the D-FACTS devices deployed on branch B_{25} , $b_{16,19}$ is changed to $b'_{16,19}$. According to Eqs. (6.35) and (6.32), this FmDI attack cannot be detected by PFDD and an offset of $\theta'_a = b_{16,19}\theta_a/b'_{16,19}$ is injected to θ_{19} , θ_{20} , θ_{33} , and θ_{34} , respectively. This is equivalent to an FmDI attack with

$$\mathbf{c} = \underbrace{(0, 0, \dots, 0, \overbrace{\theta'_a}^{19\text{-th}}, \overbrace{\theta'_a}^{20\text{-th}}, 0, \dots, 0, \overbrace{\theta'_a}^{33\text{-th}}, \overbrace{\theta'_a}^{34\text{-th}})}_{39}^\top, \quad (6.50)$$

when the susceptance of branch B_{25} is $b'_{16,19}$.

6.5 Conclusions

In this paper, we systemically investigated the feasibility and limitations of using PFDD approach to detect FmDI attacks on smart grids. Taking into account three types of FmDI attacks namely single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks respectively, we obtained the profiles of the minimum efforts required for using the D-FACTS devices in PFDD to detect FmDI attacks. We also proved that PFDD can guarantee detecting the existence of all these three types of FmDI attacks if and only if the deployment of D-FACTS devices covers branches containing at least a spanning tree of the grid graph. Last, the limitations of PFDD were investigated with findings that the PFDD approach is not able to detect effective FmDI attacks targeting on buses or super-buses with degrees 1. Further studies are needed to investigate whether PFDD by activating D-FACTS devices can cause potential instability or other hidden problems in power grids, and how often to activate D-FACTS devices would be an optimal solution to detect FmDI attacks in smart grids.

Chapter 7

Conclusions and Future Work

In this chapter, we summarize our contributions in this thesis, and propose some potential research directions for future work.

7.1 Summary of Contributions

The research focuses in this thesis lie in main topics relating to FDI attacks in smart grid CPSs including modelling and impacts evaluation of FDI attacks, novel detection approaches for FDI attacks - both FmDI and FcDI attacks. Specifically, our main research contributions are summarized as follows:

- In Chapter 3, we have developed a stochastic Petri net based analytical model to evaluate and analyze the system reliability of smart grid CPSs, specifically against topology attacks under system countermeasures (i.e., intrusion detection systems and malfunction recovery techniques). Topology attacks are evolved from FDI attacks, where attackers initialize FDI attacks by tempering with both measurement data and grid topology information. This analytical model is featured by bolstering both transient and steady-state analyses of system reliability.

- In Chapter 4, we have proposed a distributed host-based collaborative detection scheme to detect FmDI attacks in smart grid CPSs. It is considered in this work that the PMUs, deployed to measure the operating states of power grids, can be compromised by FmDI attackers, and the trusted HMs assigned to each PMU are employed to monitor and assess PMUs' behaviors. Neighboring HMs make use of the majority voting algorithm based on a set of predefined normal behavior rules to identify the existence of abnormal measurement data collected by PMUs. In addition, an innovative reputation system with an adaptive reputation updating algorithm is also designed to evaluate the overall operating status of PMUs, by which FmDI attacks as well as the attackers can be distinctly observed.
- In Chapter 5, we have proposed a Dirichlet-based detection scheme for FcDI attacks in hierarchical smart grid CPSs. In the future hierarchical paradigm of a smart grid CPS, it is considered that the decentralized LAs responsible for local management and control can be compromised by FcDI attackers. By issuing fake or biased commands, the attackers anticipate to manipulate the regional electricity prices with the purpose of illicit financial gains. The proposed scheme builds a Dirichlet-based probabilistic model to assess the reputation levels of LAs. This probabilistic model, used in conjunction with a designed adaptive reputation incentive mechanism, enables quick and efficient detection of FcDI attacks as well as the attackers.
- In Chapter 6, we have systematically explored the feasibility and limitations of detecting FmDI attacks in smart grid CPSs using D-FACTS devices. Recent studies have investigated the possibilities of proactively detecting FmDI attacks on smart grid CPSs by using D-FACTS devices - the PFDD approach. In this work, the feasibility of using PFDD to detect FmDI attacks are investigated by considering single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FmDI attacks, respectively. It is proved that PFDD can detect all these three types of FmDI attacks targeted on buses or super-buses with degrees larger than 1, as long as the deployment of D-FACTS

devices covers branches at least containing a spanning tree of the grid graph. The minimum efforts required for activating D-FACTS devices to detect each type of FmDI attacks are respectively evaluated. In addition, the limitations of this approach are also discussed, and it is strictly proved that PFDD is not able to detect FmDI attacks targeted on buses or super-buses with degrees equalling 1.

7.2 Future Work

Our studies have made significant progress in enhancing smart grid security and reliability, especially in detection and mitigation of FDI attacks. There are still plenty of work that can be done along the same direction. The following research topics will be investigated in the future as a continuation of my Ph.D. thesis.

- *Secure state estimation and proactive mitigation of FDI attacks*: Liu *et al.* mentioned that if the attackers can intrude into the control systems and obtain the knowledge of \mathbf{H} matrix, the state estimation can be bypassed leading to a success of FDI attacks. As we see, there is still a need to ensure secure state estimation even through the attackers are strong enough to get access to the control systems. As we introduced in Section 2.1, the control center needs to calculate the \mathbf{H} matrix and use it to conduct state estimation. With a closer look at Eqs. (2.4) and (2.7), we notice that to achieve state estimation and bad data detection, the control center can choose to use matrices $\mathbf{\Lambda}$ and $\mathbf{\Omega} \triangleq \mathbf{I} - \mathbf{H}\mathbf{\Lambda}$, respectively, instead of \mathbf{H} matrix itself. This makes it possible to hide \mathbf{H} matrix behind $\mathbf{\Lambda}$ and $\mathbf{\Omega}$. Since \mathbf{H} matrix is the critical information for attackers to construct FmDI attacks, it may therefore be a useful way to mitigate FmDI attacks by hiding \mathbf{H} matrix when conducting state estimation and bad data detection. We are motivated to conduct future research studies to achieve secure estimation and bad data detection without explicit content of \mathbf{H} matrix, e.g., using applied cryptography. Unlike traditional approaches for passively detecting FmDI attacks, this approach can

be considered as a proactive way to mitigate FDI attacks in smart grid CPS, which would be a breakthrough for solving such a problem and shed lights for upcoming studies.

- *Investigation of side effects by using D-FACTS devices:* In Chapter 6, we have discussed the feasibility and limitations of using PFDD to detect FmDI attacks in smart grid. However, there may be concerns about the potential negative impacts on power systems caused by using D-FACTS devices. For example, activating D-FACTS devices may, to some extent, compromise the optimal power flows and, therefore, lead to a certain power losses. If it is true, is there any suboptimal status for power flows that we may transfer to when activating D-FACTS devices. It may also be a concern that whether activating D-FACTS devices can cause transient instabilities or any other hidden problem to the electric grid. If yes, how to contain such problems remains open, which deserves extensive investigations in the future.
- *Prototyping and real-world implementation of proposed solutions:* In order to promote security and reliability of real smart grid infrastructures, we are planning to deploy the solutions presented in this thesis over a test bed, and evaluate and refine the prototypes in a real-world implementation. We expect to cooperate with those teams who have a test bed for smart grid CPS, and test all our proposed solutions.

Appendix A

Author's Publications

A.1 Book Chapters

- B1. **Beibei Li**, Rongxing Lu, and Haiyong Bao. “Behavior Rule Specification-based False Data Injection Detection Technique for Smart Grid”. In *Cyber Security for Industrial Control Systems: From the Viewpoint of Close-Loop*, pp. 119-150, Apr. 2016, CRC Press.

A.2 Journal Papers

- J1. **Beibei Li**, Gaoxi Xiao, Rongxing Lu, Ruilong Deng, Haiyong Bao. “On Feasibility and Limitations of Detecting False Data Injection Attacks on Smart Grids Using D-FACTS Devices”. Submitted to *IEEE Transactions on Industrial Informatics*, 2018.
- J2. Yandong Zheng, Rongxing Lu, **Beibei Li**, Jun Shao, Haomiao Yang, Kim-Kwang Raymond Choo. “Efficient Privacy-Preserving Data Merging and Skyline Computation over Multi-source Encrypted Data”, Submitted to *Information Sciences*, 2018.

- J3. **Beibei Li**, Rongxing Lu, Gaoxi Xiao, Haiyong Bao, Ali A. Ghorbani, “Towards Insider Threats Detection in Smart Grid Communication Systems”. Submitted to *IET Communications*, 2018.
- J4. **Beibei Li**, Rongxing Lu, Kim-Kwang Raymond Choo, Wei Wang, Sheng Luo. “On Reliability Analysis of Smart Grids under Topology Attacks: A Stochastic Petri Net Approach”. *ACM Transactions on Cyber-Physical Systems*, vol. 3, no. 1: 10, Jan. 2019.
- J5. **Beibei Li**, Rongxing Lu, Wei Wang, Kim-Kwang Raymond Choo. “Distributed Host-based Collaborative Detection for False Data Injection Attacks in Smart Grid Cyber-Physical System”. *Journal of Parallel and Distributed Computing*, vol. 103, pp. 32-41, May 2017.
- J6. **Beibei Li**, Rongxing Lu, Wei Wang, Kim-Kwang Raymond Choo. “DDOA: A Dirichlet-based Detection Scheme for Opportunistic Attacks in Smart Grid Cyber-Physical System”. *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2415 - 2425, Nov. 2016.
- J7. Haiyong Bao, Rongxing Lu, **Beibei Li**, Ruilong Deng. “BLITHE: Behavior Rule Based Insider Threat Detection for Smart Grid”. *IEEE Internet of Things Journal*, vol. 3, no. 2, pp. 190-205, Apr. 2016.

A.3 Conference Papers

- C1. **Beibei Li**, Rongxing Lu, Gaoxi Xiao, Zhou Su, Ali Ghorbani. “PAMA: A Proactive Approach to Mitigate False Data Injection Attacks in Smart Grids”. *Proc. IEEE GLOBECOM 2018*, Abu Dhabi, UAE, Dec. 9-13, 2018.

-
- C2. Mi Wen, Donghuan Yao, **Beibei Li**, Rongxing Lu. “State Estimation Based Energy Theft Detection Scheme with Privacy Preservation in Smart Grid”. *Proc. IEEE ICC 2018*, Kansas City, MO, USA, May 20-24, 2018.
- C3. **Beibei Li**, Rongxing Lu, Gaoxi Xiao. “HMM-Based Fast Detection of False Data Injections in Advanced Metering Infrastructure”. *Proc. IEEE GLOBECOM 2017*, Singapore, Dec. 4-8, 2017.

References

- [1] K. Moslehi and R. Kumar, "A reliability perspective of the smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 57–64, June 2010.
- [2] H. Farhangi, "The path of the smart grid," *IEEE Power Energy Mag.*, vol. 8, no. 1, pp. 18–28, Jan.-Feb. 2010.
- [3] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power Energy Mag.*, vol. 3, no. 5, pp. 34–41, Sep.-Oct. 2005.
- [4] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu, "Securing smart grid: cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, Aug. 2012.
- [5] J. P. Farwell and R. Rohozinski, "Stuxnet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, Feb. 2011.
- [6] N. Falliere, L. O. Murchu, and E. Chien, "W32. Stuxnet dossier," *White paper, Symantec Corp., Security Response*, vol. 5, no. 6, p. 29, Feb. 2011.
- [7] J. Cloherty and P. Thomas, "'Trojan Horse' bug lurking in vital US computers since 2011," *ABC News*, Nov. 2014, <https://abcnews.go.com/US/trojan-horse-bug-lurking-vital-us-computers-2011/story?id=26737476> (Accessed: 2018-06-28).
- [8] N. Perlroth, "In cyberattack on Saudi firm, U.S. sees Iran firing back," *The New York Times*, Oct. 2012, <https://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html> (Accessed: 2018-06-28).
- [9] B. Johnson, D. Caban, M. Krotofil, D. Scali, N. Brubaker, and C. Glycer, "Attackers deploy new ICS attack framework 'TRITON' and cause operational disruption to critical infrastructure," *FireEye*, Dec. 2017, <https://www.fireeye.com/blog/threat-research/2017/12/attackers-deploy-new-ics-attack-framework-triton.html> (Accessed: 2018-06-28).
- [10] J. Goldman, "In cyberattack on Saudi firm, U.S. sees Iran firing back," *eSecurity Planet*, Feb. 2013, <https://www.esecurityplanet.com/network-security/florida-utility-company-hit-by-cyber-attack.html> (Accessed: 2018-06-28).
- [11] K. Zetter, "Inside the cunning, unprecedented hack of Ukraine's power grid," *Wired*, Mar. 2016, <https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/> (Accessed: 2018-06-28).

- [12] U.S., “National electric grid security and resilience action plan,” *Executive Office of the President, U.S.*, Dec. 2016, https://www.whitehouse.gov/sites/whitehouse.gov/files/images/National_Electric_Grid_Action_Plan_06Dec2016.pdf (Accessed: 2018-06-28).
- [13] Canada, “National electric grid security and resilience action plan,” *Government of Canada, Canada*, Dec. 2016, https://www.whitehouse.gov/sites/whitehouse.gov/files/images/National_Electric_Grid_Action_Plan_06Dec2016.pdf (Accessed: 2018-06-28).
- [14] Reuters, “UPDATE 1-china targets \$300 bln power grid spend over 2015-20 - report,” *Reuters*, Sep. 2015, <https://www.reuters.com/article/china-power-transmission-idUSL4N1171UP20150901> (Accessed: 2018-06-28).
- [15] Australia, “A better energy future for Australia,” *Departement of the Environment and Energy, Australia*, 2018, <https://www.energy.gov.au/government-priorities/better-energy-future-australia> (Accessed: 2018-06-28).
- [16] UK, “Energy security strategy,” *Departement of Energy and Climate Change, U.K.*, Nov. 2012, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/65643/7101-energy-security-strategy.pdf (Accessed: 2018-06-28).
- [17] B. A. Hamilton, J. Miller, and B. Renz, “Understanding the benefits of the smart grid-smart grid implementation strategy,” *United States: United States Department of Energy’s National Energy Technology Laboratory*, June 2010, https://www.netl.doe.gov/File%20Library/research/energy%20efficiency/smart%20grid/whitepapers/06-18-2010_Understanding-Smart-Grid-Benefits.pdf (Assessed: 2018-06-28).
- [18] W. Li, *Risk assessment of power systems: models, methods, and applications*. John Wiley & Sons, Mar. 2014.
- [19] R. Lu, *Privacy-enhancing aggregation techniques for smart grid communications*. Springer, May 2016.
- [20] F. C. Schweppe and D. B. Rom, “Power system static-state estimation, part II: Approximate model,” *IEEE Trans. Power App. Syst.*, no. 1, pp. 125–130, Jan. 1970.
- [21] H. Bao, R. Lu, B. Li, and R. Deng, “BLITHE: Behavior rule based insider threat detection for smart grid,” *IEEE Internet Things J.*, vol. 3, no. 2, pp. 190–205, Apr. 2016.
- [22] J. Tian, R. Tan, X. Guan, and T. Liu, “Hidden moving target defense in smart grids,” in *Proc. 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids, Pittsburgh, PA, USA, Apr. 18-21, 2017*, pp. 21–26.
- [23] C. Liu, J. Wu, C. Long, and D. Kundur, “Reactance perturbation for detecting and identifying FDI attacks in power system state estimation,” *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 763–776, Aug. 2018.

- [24] J. Tian, R. Tan, X. Guan, and T. Liu, "Enhanced hidden moving target defense in smart grids," *IEEE Trans. Smart Grid*, 2018.
- [25] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Proc. IEEE 2005 Power Engineering Society General Meeting*. San Francisco, CA, USA: IEEE, 2005, pp. 454–459.
- [26] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *Proc. 8th IEE International Conference on AC and DC Power Transmission (ACDC 2006)*. London, UK: IET, 2006, pp. 58–62.
- [27] A. Monticelli, *State estimation in electric power systems: A generalized approach*. Springer Science & Business Media, Dec. 2012.
- [28] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, p. 13, May 2011.
- [29] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in Internet of Things," *J. Network and Computer Applications*, vol. 84, pp. 25–37, Apr. 2017.
- [30] N. K. Thanigaivelan, E. Nigussie, R. K. Kanth, S. Virtanen, and J. Isoaho, "Distributed internal anomaly detection system for Internet-of-Things," in *Proc. 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, Jan. 9-12, 2016*, pp. 319–320.
- [31] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov. 2015.
- [32] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study," *IEEE Syst. J.*, vol. 9, no. 1, pp. 31–44, Mar. 2015.
- [33] T.-H. Lee, C.-H. Wen, L.-H. Chang, H.-S. Chiang, and M.-C. Hsieh, "A lightweight intrusion detection scheme based on energy consumption analysis in 6lowpan," in *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*. Springer, Nov. 2014, pp. 1205–1213.
- [34] D. Oh, D. Kim, and W. W. Ro, "A malicious pattern detection engine for embedded security systems in the Internet of Things," *Sensors*, vol. 14, no. 12, pp. 24 188–24 211, Dec. 2014.
- [35] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, "On false data-injection attacks against power system state estimation: Modeling and countermeasures," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 717–729, Mar. 2014.
- [36] Y. Huang, M. Esmalifalak, H. Nguyen, R. Zheng, Z. Han, H. Li, and L. Song, "Bad data injection in smart grid: attack and defense mechanisms," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 27–33, Jan. 2013.

- [37] M. Esmalifalak, Z. Han, and L. Song, "Effect of stealthy bad data injection on network congestion in market based power system," in *Proc. 2012 IEEE Wireless Communications and Networking Conference (WCNC), Paris, France, Apr. 1-4, 2012*, pp. 2468–2472.
- [38] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. First IEEE International Conference on Smart Grid Communications (SmartGridComm), Gaithersburg, MD, USA, Oct. 4-6, 2010*, pp. 226–231.
- [39] Y. Huang, H. Li, K. A. Campbell, and Z. Han, "Defending false data injection attack on smart grid network using adaptive CUSUM test," in *Proc. 45th Annual Conference on Information Sciences and Systems (CISS), The John Hopkins University, Baltimore, MD, USA, Mar. 23-25, 2011*, pp. 1–6.
- [40] M. Esmalifalak, H. Nguyen, R. Zheng, and Z. Han, "Stealth false data injection using independent component analysis in smart grid," in *Proc. IEEE Second International Conference on Smart Grid Communications (SmartGridComm), Brussels, Belgium, Oct. 17-20, 2011*, pp. 244–248.
- [41] S. Li, Y. Yilmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.
- [42] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting false data injection attacks in AC state estimation," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2476–2483, Sep. 2015.
- [43] Y. Chen, S. Nyemba, and B. Malin, "Detecting anomalous insiders in collaborative information systems," *IEEE Trans. Dependable Secure Comput.*, vol. 9, no. 3, pp. 332–344, May-June 2012.
- [44] A. Ambre and N. Shekhar, "Insider threat detection using log analysis and event correlation," *Procedia Comput. Sci.*, vol. 45, pp. 436–445, Mar. 2015.
- [45] P. A. Legg, O. Buckley, M. Goldsmith, and S. Creese, "Automated insider threat detection system using user and role-based profile assessment," *IEEE Syst. J.*, vol. 11, no. 2, pp. 503–512, June 2017.
- [46] O. Brdiczka, J. Liu, B. Price, J. Shen, A. Patil, R. Chow, E. Bart, and N. Ducheneaut, "Proactive insider threat detection through graph learning and psychological context," in *Proc. IEEE Symposium on Security and Privacy Workshops, San Francisco, CA, USA, May 24-25, 2012*, pp. 142–149.
- [47] M. J. Mayhew, M. Atighetchi, A. Adler, and R. Greenstadt, "Use of machine learning in big data analytics for insider threat detection," in *Proc. 34th IEEE Military Communications Conference (MILCOM), Tampa, FL, USA, Oct. 26-28, 2015*, pp. 915–922.
- [48] M. Ring, S. Wunderlich, D. Grödl, D. Landes, and A. Hotho, "A toolset for intrusion and insider threat detection," in *Data Analytics and Decision Support for Cybersecurity*. Springer, Aug. 2017, pp. 3–31.

- [49] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "DDOA: A dirichlet-based detection scheme for opportunistic attacks in smart grid cyber-physical system," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 11, pp. 2415–2425, Nov. 2016.
- [50] K.-K. R. Choo, "A conceptual interdisciplinary plug-and-play cyber security framework," in *ICTs and the Millennium Development Goals*. Springer Science & Business Media New York, Apr. 2014, pp. 81–99.
- [51] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system," *J. Parallel Distrib. Comput.*, vol. 103, pp. 32–41, May 2017.
- [52] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.
- [53] J. Wu, M. Dong, K. Ota, Z. Zhou, and B. Duan, "Towards fault-tolerant fine-grained data access control for smart grid," *Wireless personal communications*, vol. 75, no. 3, pp. 1787–1808, June 2014.
- [54] J. Pagliery. (2015, Oct.) ISIS is attacking the U.S. energy grid. <http://money.cnn.com/2015/10/15/technology/isis-energy-grid/> (Accessed: 2018-06-28).
- [55] W. Kröger, "Critical infrastructures at risk: A need for a new conceptual approach and extended analytical tools," *Reliability Engineering & System Safety*, vol. 93, no. 12, pp. 1781–1787, Dec. 2008.
- [56] E. A. Lee, "Fundamental limits of cyber-physical systems modeling," *ACM Transactions on Cyber-Physical Systems*, vol. 1, no. 1, pp. 3–26, Feb. 2017.
- [57] J. Kim and L. Tong, "On topology attack of a smart grid: undetectable attacks and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1294–1305, July 2013.
- [58] G. Dalton, R. F. Mills, J. M. Colombi, and R. A. Raines, "Analyzing attack trees using generalized stochastic Petri nets," in *Proc. IEEE Information Assurance Workshop (IAW'06), West Point, NY, USA, June 21-23, 2006*, pp. 116–123.
- [59] R. Mitchell and I.-R. Chen, "Modeling and analysis of attacks and counter defense mechanisms for cyber physical systems," *IEEE Trans. Rel.*, vol. 65, no. 1, pp. 350–358, Mar. 2016.
- [60] R. McDonald, F. Pereira, K. Ribarov, and J. Hajič, "Non-projective dependency parsing using spanning tree algorithms," in *Proc. Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP'05), Vancouver, British Columbia, Canada, Oct. 06-08, 2005*, pp. 523–530.
- [61] K. Jensen and G. Rozenberg, *High-level Petri nets: theory and application*. Springer Science & Business Media, Dec. 2012.
- [62] J. P. McDermott, "Attack net penetration testing," in *Proc. New Security Paradigms Workshop (NSPW'00), Ballycotton, County Cork, Ireland, Sep. 18-21, 2001*, pp. 15–21.

- [63] F. Bause and P. Kritzinger, *Stochastic Petri Nets*, Jan. 1996, vol. 26.
- [64] F. Tüysüz and C. Kahraman, “Modeling a flexible manufacturing cell using stochastic Petri nets with fuzzy parameters,” *Expert Systems with Applications*, vol. 37, no. 5, pp. 3910–3920, May 2010.
- [65] K. Jensen, *Coloured Petri nets: basic concepts, analysis methods and practical use*. Springer Science & Business Media, Apr. 2013, vol. 1.
- [66] J.-C. Laprie, K. Kanoun, and M. Kaâniche, “Modelling interdependencies between the electricity and information infrastructures,” in *Proc. International Conference on Computer Safety, Reliability, and Security (SAFECOMP), Nuremberg, Germany, Sep. 18-21, 2007*, pp. 54–67.
- [67] R. Zeng, Y. Jiang, C. Lin, and X. Shen, “Dependability analysis of control center networks in smart grid using stochastic Petri nets,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1721–1730, Sep. 2012.
- [68] T. M. Chen, J. C. Sanchez-Aarnoutse, and J. Buford, “Petri net modeling of cyber-physical attacks on smart grid,” *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 741–749, Dec. 2011.
- [69] A. Householder, K. Houle, and C. Dougherty, “Computer attack trends challenge Internet security,” *Computer*, vol. 35, no. 4, pp. sulp5–sulp7, Apr. 2002.
- [70] J. Weimer, S. Kar, and K. H. Johansson, “Distributed detection and isolation of topology attacks in power networks,” in *Proc. 1st International Conference on High Confidence Networked Systems (HiCoNS), Beijing, China, Apr. 17-18, Beijing, China, 2012*, pp. 65–72.
- [71] S. Liu, X. P. Liu, and A. El Saddik, “Denial-of-Service (DoS) attacks on load frequency control in smart grids,” in *Proc. Innovative Smart Grid Technologies (ISGT), Washington, DC, USA, Feb. 24-27, 2013*, pp. 1–6.
- [72] S. Amin, A. A. Cárdenas, and S. S. Sastry, “Safe and secure networked control systems under Denial-of-Service attacks,” in *Proc. International Workshop on Hybrid Systems: Computation and Control (HSCC), San Francisco, CA, USA, Apr. 13-15, 2009*, pp. 31–45.
- [73] G. Hug and J. A. Giampapa, “Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks,” *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, Sep. 2012.
- [74] Z. Djekic, “Online circuit breaker monitoring system,” Ph.D. dissertation, Texas A&M University, College Station, TX, Dec. 2007.
- [75] D. Boneh, B. Lynn, and H. Shacham, “Short signatures from the Weil pairing,” *J. Cryptology*, vol. 17, no. 4, pp. 297–319, Sep. 2004.
- [76] L. Huang, Y. Sun, J. Xu, W. Gao, J. Zhang, and Z. Wu, “Optimal PMU placement considering controlled islanding of power system,” *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 742–755, Mar. 2014.

- [77] H. W. Vildhøj and D. K. Wind, “Supplementary notes for graph theory 1,” pp. 26–26, 2016, <http://www.student.dtu.dk/~dawi/01227/01227-GraphTheory.pdf> (Accessed: 2018-06-28).
- [78] C. A. Hoare, “Quicksort,” *The Computer Journal*, vol. 5, no. 1, pp. 10–16, 1962.
- [79] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty *et al.*, “The IEEE reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee,” *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.
- [80] R. L. Graham and P. Hell, “On the history of the minimum spanning tree problem,” *Annals of the History of Computing*, vol. 7, no. 1, pp. 43–57, Jan.-Mar. 1985.
- [81] J. B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proceedings of the American Mathematical society*, vol. 7, no. 1, pp. 48–50, Feb. 1956.
- [82] N. J. Dingle, W. J. Knottenbelt, and T. Suto, “PIPE2: A tool for the performance evaluation of generalised stochastic Petri nets,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 36, no. 4, pp. 34–39, Mar. 2009.
- [83] S. Shin and G. Gu, “Conficker and beyond: a large-scale empirical study,” in *Proc. 26th Annual Computer Security Applications Conf. (ACSAC), Austin, Texas, USA, Dec. 06-10, 2010*, pp. 151–160.
- [84] S. Gorman, Y. J. Dreazen, and A. Cole, “Insurgents hack US drones,” *Wall Street Journal*, vol. 17, Dec. 2009.
- [85] H. Fang, L. Xu, and K.-K. R. Choo, “Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks,” *Appl. Math. Comput.*, vol. 296, pp. 153–167, Mar. 2017.
- [86] J. Chen, L. Shi, P. Cheng, and H. Zhang, “Optimal denial-of-service attack scheduling with energy constraint,” *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 3023–3028, Nov. 2015.
- [87] E. Handschin, F. Schweppe, J. Kohlas, and A. Fiechter, “Bad data analysis for power system state estimation,” *IEEE Trans. Power App. Syst.*, vol. 94, no. 2, pp. 329–337, Mar. 1975.
- [88] H. M. Merrill and F. C. Schweppe, “Bad data suppression in power system static state estimation,” *IEEE Trans. Power App. Syst.*, no. 6, pp. 2718–2725, Nov. 1971.
- [89] J. Chen and A. Abur, “Placement of PMUs to enable bad data detection in state estimation,” *IEEE Trans. Power App. Syst.*, vol. 21, no. 4, pp. 1608–1615, Nov. 2006.
- [90] W. W. Kotiuga and M. Vidyasagar, “Bad data rejection properties of weighted least absolute value techniques applied to static state estimation,” *IEEE Trans. Power App. Syst.*, no. 4, pp. 844–853, Apr. 1982.

- [91] T. V. Cutsem, M. Ribbens-Pavell, and L. Mili, "Hypothesis testing identification: a new method for bad data analysis in power system state estimation," *IEEE Trans. Power App. Syst.*, no. 11, pp. 3239–3252, Nov. 1984.
- [92] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power App. Syst.*, vol. 9, no. 3, pp. 1601–1609, Aug. 1994.
- [93] M. M. Nordman and M. Lehtonen, "Distributed agent-based state estimation for electrical distribution networks," *IEEE Trans. Power App. Syst.*, vol. 20, no. 2, pp. 652–658, May 2005.
- [94] M. Qiu, W. Gao, M. Chen, J.-W. Niu, and L. Zhang, "Energy efficient security algorithm for power grid wide area monitoring system," *IEEE Trans. on Smart Grid*, vol. 2, no. 4, pp. 715–723, Dec. 2011.
- [95] M. Qiu, H. Su, M. Chen, Z. Ming, and L. T. Yang, "Balance of security strength and energy for a PMU monitoring system in smart grid," *IEEE Commun. Mag.*, vol. 50, no. 5, pp. 142–149, May 2012.
- [96] A. Castiglione, R. Pizzolante, C. Esposito, A. De Santis, F. Palmieri, and A. Castiglione, "A collaborative clinical analysis service based on theory of evidence, fuzzy linguistic sets and prospect theory and its application to craniofacial disorders in infants," *Future Gener. Comput. Syst.*, vol. 67, pp. 230–241, Aug. 2017.
- [97] A. Patel, H. Alhussian, J. M. Pedersen, B. Bounabat, J. C. Júnior, and S. Katsikas, "A nifty collaborative intrusion detection and prevention architecture for smart grid ecosystems," *Computers & Security*, vol. 64, pp. 92–109, July 2017.
- [98] C. Esposito, A. Castiglione, F. Palmieri, and M. Ficco, "Trust management for distributed heterogeneous systems by using linguistic term sets and hierarchies, aggregation operators and mechanism design," *Future Gener. Comput. Syst.*, vol. 74, pp. 325–336, Dec. 2017.
- [99] U. S. Premarathne, I. Khalil, and M. Atiquzzaman, "Trust based reliable transmissions strategies for smart home energy consumption management in cognitive radio based smart grid," *Ad Hoc Netw.*, vol. 41, pp. 15–29, May 2016.
- [100] K. Bowman and L. Shenton, "Parameter estimation for the Beta distribution," *J. Stat. Comput. Simul.*, vol. 43, no. 3-4, pp. 217–228, Nov. 1992.
- [101] M. Srivatsa, L. Xiong, and L. Liu, "Trustguard: countering vulnerabilities in reputation management for decentralized overlay networks," in *Proc. 14th International Conference on World Wide Web (WWW), Chiba, Japan, May 10-14, 2005*, pp. 422–431.
- [102] F. G. Mármol and G. M. Pérez, "TRIP, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks," *J. Network and Computer Applications*, vol. 35, no. 3, pp. 934–941, May 2012.
- [103] Y. L. Sun, Z. Han, W. Yu, and K. R. Liu, "A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks." in *Proc. 25th IEEE INFOCOM, Barcelona, Catalunya, Spain, Apr. 23-29, 2006*, pp. 1–13.

- [104] D. Zhang, S. Li, P. Zeng, and C. Zang, "Optimal microgrid control and power-flow study with different bidding policies by using powerworld simulator," *IEEE Trans. Sustainable Energy*, vol. 5, no. 1, pp. 282–292, Jan. 2014.
- [105] PowerWorld, 2018, <https://www.powerworld.com> (Accessed: 2018-06-28).
- [106] Y. L. Yuan, Z. Y. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 382–390, June 2011.
- [107] R. Deng, G. Xiao, and R. Lu, "Defending against false data injection attacks on power system state estimation," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 198–207, Feb. 2017.
- [108] O. Vukovic and G. Dan, "Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1500–1508, July 2014.
- [109] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.
- [110] S. Iwamoto, M. Kusano, and V. H. Quintana, "Hierarchical state estimation using a fast rectangular-coordinate method," *IEEE Trans. Power Syst.*, vol. 4, no. 3, pp. 870–880, Aug. 1989.
- [111] ICS-CERT. (Nov./Dec.) NCCIC/ICS-CERT monitor november-december 2015.
- [112] K. Zetter. (2015, Aug.) A cyberattack has caused confirmed physical damage for the second time ever. <https://www.wired.com/2015/01/german-steel-mill-hack-destruction/> (Accessed: 2018-06-28).
- [113] S. Z. Bi and Y. J. Zhang, "False-data injection attack to control real-time price in electricity market," *Proc. IEEE GLOBECOM, Atlanta, GA, USA, Dec. 9-13*, pp. 772–777, 2013.
- [114] R. Tan, V. B. Krishna, D. K. Yau, and Z. Kalbarczyk, "Integrity attacks on real-time pricing in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 18, no. 2, p. 5, Dec. 2015.
- [115] L. Jia, R. J. Thomas, and L. Tong, "Impacts of malicious data on real-time price of electricity market operations," in *Proc. 45th Hawaii International Conference on System Sciences, Maui, HI, USA, Jan. 04-07, 2012*, pp. 1907–1914.
- [116] O. Kosut, L. Y. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.
- [117] M. Esmalifalak, G. Shi, Z. Han, and L. Song, "Bad data injection attack and defense in electricity market using game theory study," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 160–169, Mar. 2013.
- [118] R. X. Lu, X. H. Liang, X. Li, X. D. Lin, and X. M. Shen, "EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1621–1631, Sep. 2012.

- [119] J. Ma, Y. Liu, L. Song, and Z. Han, "Multiact dynamic game strategy for jamming attack in electricity market," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2273–2282, Sep. 2015.
- [120] R. Mitchell and I.-R. Chen, "Behavior rule specification-based intrusion detection for safety critical medical cyber physical systems," *IEEE Trans. Dependable Secure Comput.*, vol. 12, no. 1, pp. 16–30, Jan. 2015.
- [121] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.
- [122] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static-state estimation, parts I, II, and III," *IEEE Trans. Power App. Syst.*, vol. 89, no. 1, pp. 120–135, Jan. 1970.
- [123] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Multivariate Distributions, volume 1, Models and Applications*. New York: John Wiley & Sons, Apr. 2002, vol. 59.
- [124] M. Chertkov, F. Pan, and M. G. Stepanov, "Predicting failures in power grids: The case of static overloads," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 162–172, Mar. 2011.
- [125] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *Proc. IEEE GLOBECOM, Anaheim, CA, USA, Dec. 3-7*, Dec. 2012, pp. 3153–3158.
- [126] J. Valenzuela, J. Wang, and N. Bissinger, "Real-time intrusion detection in power system operations," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1052–1062, May 2013.
- [127] L. Che, X. Liu, and Z. Li, "Fast screening of high-risk lines under false data injection attacks," *IEEE Trans. Smart Grid*, 2018.
- [128] ———, "Mitigating false data attacks induced overloads using a corrective dispatch scheme," *IEEE Trans. Smart Grid*, 2018.
- [129] E. Kovacs, "Wikileaks releases details on CIA hacking tools," *SecurityWeek*, Mar. 2017, <http://www.securityweek.com/wikileaks-releases-details-cia-hacking-tools> (Accessed: 2018-06-28).
- [130] R. Deng and H. Liang, "False data injection attacks with incomplete information and countermeasures in smart grid," *IEEE Trans. Ind. Informat.*, to appear.
- [131] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, June 2011.
- [132] K. L. Morrow, E. Heine, K. M. Rogers, R. B. Bobba, and T. J. Overbye, "Topology perturbation for detecting malicious data injection," in *Proc. 45th Hawaii International Conference on System Science (HICSS), Maui, HI, USA, Jan. 4-7*, 2012, pp. 2104–2113.

-
- [133] M. A. Rahman, E. Al-Shaer, and R. B. Bobba, "Moving target defense for hardening the security of the power system state estimation," in *Proc. First ACM Workshop on Moving Target Defense, Scottsdale, Arizona, USA, Nov. 07-07*, Nov. 2014, pp. 59–68.
- [134] D. Divan and H. Johal, "Distributed FACTS - A new concept for realizing grid power flow control," *IEEE Trans. Power Electron.*, vol. 22, no. 6, pp. 2253–2260, Nov. 2007.
- [135] K. Rogers and T. Overbye, "Some applications of distributed flexible AC transmission system (D-FACTS) devices in power systems," in *Proc. 40th North American Power Symposium (NAPS), Calgary, AB, Canada, Sep. 28-30*, 2008, pp. 1–8.
- [136] K. C. Sou, H. Sandberg, and K. H. Johansson, "Electric power network security analysis via minimum cut relaxation," in *Proc. 50th Conference on Decision and Control and European Control Conference (CDC-ECC), Orlando, FL, USA, Dec. 12-15*, 2011, pp. 4054–4059.

