Optimal planning and operation of moving target defense for detecting false data injection attacks in smart grids

by

Bo Liu

B.S., Harbin Institute of Technology, China, 2013

M.S., Harbin Institute of Technology, China, 2015

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

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

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Abstract

Moving target defense (MTD) in the power system is a promising defense strategy to detect false data injection (FDI) attacks against state estimation by using distributed flexible AC transmission system (D-FACTS) devices. Optimal planning and operation are two essential stages in the MTD application. MTD planning determines the optimal allocation of D-FACTS devices, while MTD operation decides the optimal D-FACTS setpoints under different load conditions in real-time. However, most MTD works focus on studying the MTD operation methods and neglect MTD planning. It is generally assumed that all lines are equipped with D-FACTS devices, which is the most expensive MTD planning solution. This dissertation separates MTD planning and MTD operation as two independent problems by distinguishing their roles in attack detection effectiveness, MTD application costs, and MTD hiddenness. The contributions of this work are three-fold as follows.

Firstly, this dissertation proves that MTD planning can determine the MTD detection effectiveness, regardless of D-FACTS device setpoints in MTD operation. This work designs max-rank MTD planning algorithms by using the minimum number of D-FACTS devices to ensure MTD detection effectiveness and minimize the MTD planning cost. It is proved that any MTDs under proposed planning algorithms have the maximum rank of its composite matrix, a widely used metric of the MTD detection effectiveness. In addition, this work further points out the maximum rank of the composite matrix is not strictly equivalent to maximal MTD detection effectiveness. Three types of unprotected buses in MTD are identified, and attack detecting probability (ADP) is introduced as a novel metric for measuring the detection effectiveness of MTD planning. It is proved that the rank of the composite matrix merely represents the lower bound of ADP, while the number of unprotected buses determines the upper bound of ADP. Then, a novel graph-theory-based planning algorithm is proposed to achieve maximal MTD detection effectiveness.

Secondly, this dissertation highlights that MTD operation ought to focus on reducing the MTD operation cost. This work proposes an AC optimal power flow (ACOPF) model considering D-FACTS devices as an MTD operation model, in which the reactance of D-FACTS equipped lines are introduced as decision variables to minimize system losses and generation costs. The proposed model can be used by system operators to achieve economic and cybersecure system operations. In addition, this dissertation rigorously derives the gradient and Hessian matrices of the objective function and constraints with respect to line reactance, which are further used to build an interior-point solver of the proposed ACOPF model.

Finally, this dissertation designs the optimal planning and operation of D-FACTS devices for hidden MTD (HMTD), which is a superior MTD method stealthy to sophisticated attackers. A depth-first-search-based MTD planning algorithm is proposed to guarantee the MTD hiddenness while maximizing the rank of its composite matrix and covering all necessary buses. Additionally, this work proposes DC- and AC-HMTD operation models to determine the setpoints of D-FACTS devices. The optimization-based DC-HMTD model outperforms the existing HMTD operation in terms of CPU time and detection effectiveness. The ACOPF-based HMTD operation model ensures the hiddenness and minimizes the generation cost to utilize the economic benefits of D-FACTS devices. Comparative numerical results on multiple systems show the efficacy of the proposed planning and operation approaches in achieving high detecting effectiveness and MTD hiddenness.

Optimal planning and operation of moving target defense for detecting false data injection attacks in smart grids

by

Bo Liu

B.S., Harbin Institute of Technology, China, 2013

M.S., Harbin Institute of Technology, China, 2015

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

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

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Major Professor
Hongyu Wu

Copyright

Abstract

Moving target defense (MTD) in the power system is a promising defense strategy to detect false data injection (FDI) attacks against state estimation by using distributed flexible AC transmission system (D-FACTS) devices. Optimal planning and operation are two essential stages in the MTD application. MTD planning determines the optimal allocation of D-FACTS devices, while MTD operation decides the optimal D-FACTS setpoints under different load conditions in real-time. However, most MTD works focus on studying the MTD operation methods and neglect MTD planning. It is generally assumed that all lines are equipped with D-FACTS devices, which is the most expensive MTD planning solution. This dissertation separates MTD planning and MTD operation as two independent problems by distinguishing their roles in attack detection effectiveness, MTD application costs, and MTD hiddenness. The contributions of this work are three-fold as follows.

Firstly, this dissertation proves that MTD planning can determine the MTD detection effectiveness, regardless of D-FACTS device setpoints in MTD operation. This work designs max-rank MTD planning algorithms by using the minimum number of D-FACTS devices to ensure MTD detection effectiveness and minimize the MTD planning cost. It is proved that any MTDs under proposed planning algorithms have the maximum rank of its composite matrix, a widely used metric of the MTD detection effectiveness. In addition, this work further points out the maximum rank of the composite matrix is not strictly equivalent to maximal MTD detection effectiveness. Three types of unprotected buses in MTD are identified, and attack detecting probability (ADP) is introduced as a novel metric for measuring the detection effectiveness of MTD planning. It is proved that the rank of the composite matrix merely represents the lower bound of ADP, while the number of unprotected buses determines the upper bound of ADP. Then, a novel graph-theory-based planning algorithm is proposed to achieve maximal MTD detection effectiveness.

Secondly, this dissertation highlights that MTD operation ought to focus on reducing the MTD operation cost. This work proposes an AC optimal power flow (ACOPF) model considering D-FACTS devices as an MTD operation model, in which the reactance of D-FACTS equipped lines are introduced as decision variables to minimize system losses and generation costs. The proposed model can be used by system operators to achieve economic and cybersecure system operations. In addition, this dissertation rigorously derives the gradient and Hessian matrices of the objective function and constraints with respect to line reactance, which are further used to build an interior-point solver of the proposed ACOPF model.

Finally, this dissertation designs the optimal planning and operation of D-FACTS devices for hidden MTD (HMTD), which is a superior MTD method stealthy to sophisticated attackers. A depth-first-search-based MTD planning algorithm is proposed to guarantee the MTD hiddenness while maximizing the rank of its composite matrix and covering all necessary buses. Additionally, this work proposes DC- and AC-HMTD operation models to determine the setpoints of D-FACTS devices. The optimization-based DC-HMTD model outperforms the existing HMTD operation in terms of CPU time and detection effectiveness. The ACOPF-based HMTD operation model ensures the hiddenness and minimizes the generation cost to utilize the economic benefits of D-FACTS devices. Comparative numerical results on multiple systems show the efficacy of the proposed planning and operation approaches in achieving high detecting effectiveness and MTD hiddenness.

Table of Contents

List of Figures xii

List of Tables xiv

Acknowledgements xv

Dedication xvii

1 Introduction 1

1.1 Background 1

1.2 A Brief Introduction to Moving Target Defense 2

1.3 A Brief Introduction to Cyber-physical Smart Grid 4

1.3.1 Cyber-physical Attacks in the Smart Grid 6

1.3.2 Defense Approaches in the Smart Grid 7

1.4 Research Motivations 9

1.5 Research Contributions 11

1.6 Organization of This Dissertation 15

2 Fundamentals and Related Literature 16

2.1 State Estimation and Bad Data Detection 16

2.1.1 DC-SE Formulation and BDD 16

2.1.2 AC-SE Formulation 17

2.2 FDI Attacks against SE 18

2.2.1 DC-FDI Attacks 19

2.2.2 AC-FDI Attacks 20

2.3 MTD Model in the Smart Grid . . . 21
2.4 State-of-the-art MTD Literature Review . . . 23
2.4.1 MTD Planning Approaches . . . 23
2.4.2 MTD Operation Approaches . . . 24
2.4.3 Evaluation of MTD Performance . . . 25

3 MTD Planning Algorithm . . . 27
3.1 Introduction . . . 27
3.2 Preliminaries . . . 29
3.2.1 Notation . . . 29
3.2.2 MTD Detection Effectiveness Metric . . . 30
3.2.3 Graph Theory for Power System Topology . . . 31
3.2.4 Linear Sensitivity of Transmission Loss to Line Reactance . . . 32
3.3 Max-rank Planning Algorithms . . . 33
3.3.1 Necessary Conditions for a Complete MTD . . . 33
3.3.2 Sufficient Conditions for Complete and Incomplete MTDs . . . 36
3.3.3 Max-rank Planning Algorithms . . . 39
3.4 Graph-theory-based MTD Planning Algorithm . . . 43
3.4.1 Analysis of MTD Detection Effectiveness . . . 43
3.4.2 Graph-theory-based MTD Planning Algorithm . . . 50
3.5 Experiment Results . . . 53
3.5.1 Numerical Results in Max-rank Planning Algorithms . . . 53
3.5.2 Numerical Results in Graph-theory-based MTD Planning Algorithm . 60
3.6 Summary . . . 67

4 AC Optimal Power Flow-based MTD Operation Model . . . 70
4.1 Introduction . . . 70
4.2 ACOPF-based MTD Operation Model . . . 72

4.3 An Interior-Point Solver for ACOPF-based MTD Model . . . 75
4.3.1 Preliminaries in Derivatives in ACOPF Model . . . 75
4.3.2 Gradient of Power Injection Constraints . . . 78
4.3.3 Hessian Matrix of Power Injection Constraints . . . 78
4.3.4 Gradient and Hessian Matrix of Power Flow Constraints . . . 79
4.3.5 Gradient and Hessian Matrix of System Losses . . . 80
4.4 Experiment Results . . . 82
4.4.1 Comparison of Traditional ACOPF and ACOPF-based MTD . . . 82
4.4.2 Impact of MTD Planning on System Losses . . . 83
4.4.3 Impact of τ on System Losses . . . 84
4.5 Summary . . . 85

5 Hidden MTD Planning and Operation Algorithm . . . 87
5.1 Introduction . . . 87
5.2 Preliminaries . . . 90
5.2.1 MTD Hiddenness and Detection Effectiveness . . . 90
5.2.2 Max-rank MTD Planning Algorithms . . . 92
5.3 Novel HMTD Operation Condition . . . 93
5.4 HMTD Planning Algorithm . . . 94
5.4.1 Requirements of MTD Planning for HMTD . . . 94
5.4.2 Hidden MTD Planning Algorithm . . . 97
5.5 HMTD Operation Model . . . 101
5.5.1 DC-HMTD Operation Model . . . 101
5.5.2 AC-HMTD Operation Model . . . 102
5.5.3 Cost-benefit Analysis of DC- and AC-HMTD Model . . . 105
5.6 Experiment Results . . . 106
5.6.1 HMTD Planning Solution . . . 106
5.6.2 DC-HMTD Operation Solution . . . 107

5.6.3 AC-HMTD Operation Solution . . . 109
5.6.4 Comparison between Proposed and Existing DC-HMTD Operations . 111
5.6.5 Comparison between Hidden and Existing MTD Planning Algorithms 114
5.7 Summary . . . 117

6 Conclusion and Future Work . . . 119
6.1 Conclusion . . . 119
6.2 Future Work . . . 123

Bibliography . . . 124

A Graph-theory-based Topology Analysis . . . 138
A.1 Proof of Proposition 3.3.1 . . . 138
A.2 Proof of Lemma 3.3.1 . . . 138
A.3 Proof of Lemma 3.3.2 . . . 139
A.4 Proof of Corollary 3.3.1 . . . 140
A.5 Proof of Corollary 3.3.2 . . . 140

B Reuse permissions from publishers . . . 141

List of Figures

1.1 SCADA system network. 5

1.2 Cyber-physical attacks on smart grids. 6

2.1 Illustration of FDI attack model in the smart grid. 19

2.2 Illustration of MTD and FDI attack model in the smart grid. 22

3.1 A fully measured 3-bus system. 34

3.2 Relation of protected and unprotected buses in an MTD planning. 48

3.3 The PLIS of each line in the 6-bus system. 54

3.4 Max-rank planning of the 6-bus system. 54

3.5 The PLIS of each line in the IEEE 14-bus system. 55

3.6 Max-rank planning of the IEEE 14-bus system. 56

3.7 The histogram on the rank of the composite matrix. 57

3.8 ADP under the proposed planning with measurement noises in DC-SE. . . . 58

3.9 ADP under the proposed planning with measurement noises in AC-SE. . . . 60

3.10 The proposed MTD planning solution of the IEEE 14-bus system. 61

3.11 The ADP of the four MTD planning algorithms versus q in DC-SE. 64

3.12 The impact of VAIM on the MTD detection effectiveness under each planning algorithm. 66

3.13 The impact of measurement noises on the MTD detection effectiveness under each planning algorithm. 67

3.14 The comparison of ADPs under the four planning algorithms in AC-SE. . . . 68

4.1 Time-sequence diagram of FDI attacks and MTDs. 75

4.2 System losses versus τ. 85

5.1 Illustration of HMTD and smart attack model in the smart grid. 88

5.2 An illustration of D-FACTS placement solution in HMTD. 99

5.3 Hidden MTD planning of the IEEE 14-bus system. 107

5.4 ADP and DSP of HMTD and RMTD under different MTD magnitudes. . . . 108

5.5 The impact of variant loads on the hiddenness of proposed HMTD. 110

5.6 The trade-off between MTD savings and MTD hiddenness under different λ_0. 112

5.7 ADP of RW-HMTD and proposed HMTD under FDI attacks with different VAIMs. 114

5.8 ADP and DSP of five MTD planning algorithms under 0.2 MTD magnitude. 116

5.9 ADP and DSP of five MTD planning algorithms with MTD magnitude varying from 1% to 20%. 117

List of Tables

1.1 Taxonomy of the cyber-physical attacks in the smart grid 7

1.2 Taxonomy of the cyber-physical defense methods in the smart grid 9

2.1 Moving target defense in the smart grid 23

3.1 Nomenclature 30

3.2 CPU time of proposed max-rank planning algorithms 56

3.3 CPU time of proposed algorithm on different systems 62

3.4 Planning method comparison in medium- and large-scale systems 63

4.1 Costs, losses and CPU time under different flow limit conditions 83

4.2 System losses under different MTD planning algorithms 84

5.1 Generation costs in OPF and different MTD methods 105

5.2 The performance of AC-HMTD operation 111

5.3 Performance of the proposed HMTD operation 112

5.4 Performance of RW-HMTD using direct searching in the IEEE 14-bus system 113

5.5 Performance of RW-HMTD using indirect searching in the IEEE 57-bus system 113

5.6 Existing MTD planning algorithms 115

Acknowledgments

First of all, I would like to express my deepest gratitude to my advisor, Dr. Hongyu Wu. You are the best advisor I have ever met. Thank you for your excellent guidance, encouragement, and patience over the years. I learned a lot from you in the past four years.

I would like to sincerely thank my committee members, Dr. Caterina Scoglio, Dr. Anil Pahwa, Dr. Behrooz Mirafzal, Dr. Alexandru Bardas, for their precious time and helpful comments. I also appreciate Dr. Pascal Hitzler for donating his time to be the outside chairperson.

I would like to express my gratitude to my friends in our power system group, Hang Zhang, Lawryn Edmonds, Xuebo Liu, Li Wang, Dr. Haifeng Zhang, and Dr. Nazif Faqiry. Thank you for the support and the happy time we had. I also would like to thank my friends at K-State, Dr. Haotian Wu, Dr. Xin Li, Xinya Wang, Tianyu Lin, Dr. Wenji Zhang, Junyao Kuang, Chunlin Yi, and Futing Fan, for all the fun we had together.

My special thanks to my friend Lingxuan Xu for your generous support in the past ten years. I am also grateful to friends before I came to K-state, Dr. Haibo Li, Dr. Lizhi Qu, Dr. Dongliang Xiao, Jianan Liu, Haixiang Zhang, Dr. Fangzhou Chen, Dr. Yayu Peng, Junwei Cui, and Chao Jia. I would like to thank Dr. Haosen Wang, Dr. Fa Chen, and Dr. Xinkai Zhang for your support in my most difficult time. I really appreciate your help.

My sincere appreciation to my mother Xiaoming Wang, grandfather Yuhai Wang, and grandmother Guixin Yan. You provide me with the best education. Thank you for your unending support over the years and your sacrifice to raise me up.

Finally, wholehearted thanks to my wife Dr. Qihui Yang. I feel so lucky to have your love, understanding and support. Your encouragement makes me a better man. I also want to express my love for my daughter Iris. You brought so much happiness to your mother and me every day. With you and your mother, I am the happiest person in the world.

This work was supported in part by Kansas State University new faculty start-up fund, in part by the U.S. National Science Foundation under Grant No. 1929147, and in part by the U.S. Department of Energy under Award No. DE-EE0008767.

Dedication

Dedicated to my beloved wife Qihui and daughter Iris.

Chapter 1

Introduction

1.1 Background

In the last decades, the Internet of Things (IoT) technology and information and communication technology (ICT) enabled devices have been widely used in industrial control systems (ICSs) in critical infrastructures. It not only brings convenience, efficiency, and resilience, but also brings growing vulnerabilities to the physical-cyber systems[1]. These smart devices provide substantial attack surfaces to malicious attackers in critical infrastructures, such as water systems, nuclear systems, and power systems. The U.S. Department of Energy received 362 power interruption reports related to cyber-physical attacks between 2011 and 2014, in which 31 cases were reported in 2011 and 161 cases reported in 2013[2]. The number of attacks against the energy industries has been increasing over the last decades, and some incidents have caused massive economic losses, blackouts, explosions, and even life losses.

The major incidents in the energy sector are summarized as follows. In 1999, hackers attacked the supervisory control and data acquisition (SCADA) system of gasoline pipeline through code manipulation, which caused a fireball and killed three people. In 2003, the Slammer worm attacked the David-Besse nuclear plant in Ohio, the U.S., which resulted in the critical parameter display system off for five hours[3]. In 2007, the Aurora attack manipulated a circuit breaker of a generator in the U.S., and repeated on-and-off switching

operation, which resulted in the explosion of a one million dollar diesel generator[4]. In 2010, the Stuxnet attack damaged the Iranian nuclear fuel-enrichment facility through a computer worm in the control room. The Stuxnet worm compromised the control signals to push the system to unsafe conditions and meanwhile injected fake sensor measurements to cover the ongoing attack[5].

There are also a huge wave of cyber-physical attacks targeting electric power grids. In 2011, BlackEnergy malware compromised the ICSs of several national critical infrastructures in the U.S.[6]. In 2013, a denial-of-service (DoS) attack hit JEA, an electric utility in the U.S., resulting in a shutdown of payment systems for a few days. In 2015, the cyber-attack on three Ukrainian electric power distribution companies resulted in power outages for several hours, affecting 225,000 customers[7]. In 2016, Ukraine's power grid was hacked again. Hackers sent malware to employees via e-mails, stole login credentials, shut down substations, and shut off 200 MW generation for one hour[8].

These attacks and their severe consequences were a series of wake-up calls to the power and energy industry. Governmental agencies, industry companies, and academic institutions realized the emergence of threats to cyber-physical power systems and have been making tremendous efforts ever since to understand, analyze, and improve the cybersecurity of power systems. This dissertation focuses on improving the cybersecurity of power systems against data integrity attacks by using an emerging defense mechanism, i.e., moving target defense (MTD). Theoretical foundations established and simulation results obtained in this dissertation provide a novel understanding of guiding principles in the MTD planning and operation. The remainder of this chapter briefly reviews MTD methods, the cyber-physical system (CPS) of smart grids, the cyber-physical attacks in smart grids, and detection methods in smart grids.

1.2 A Brief Introduction to Moving Target Defense

MTD is a concept of proactively controlling change across multiple system dimensions in order to increase uncertainty and apparent complexity for attackers, reduce their window

of opportunity, and increase the costs of their probing and attack efforts[9]. The central premise of MTD is that it is impossible to provide complete and perfect security for a given system[10]. Given this, the goal of MTD is to defend against and thwart attacks through altering attack surfaces. It has been proven that MTD is a promising defense approach in information technology (IT) systems[11], which usually operate in a static configuration and provide attackers with ample reconnaissance time. Common reconnaissance defenses rely on network firewalls to prevent the attacker from discovering devices or services. However, these defenses don't protect the identity of devices that require a constant and public external presence, such as web and email servers[12]. Static nodes in network are vulnerable to denial-of-service attacks, man-in-the middle attacks, and replay attacks[13]. As a dynamic defense, MTD approaches can provide improved security by constantly altering the attack surface of a target.

There has been a large body of MTD related research efforts in various aspects of IT systems, such as network address shuffling (MTD in network address)[12], address space layout randomization (MTD in memory layouts)[14], moving target Internet protocol version 6 (IPv6) defense (MTD in the network layer of the protocol stack)[13], MTD platform for cloud-based IT systems (MTD in cloud-based IT system)[11], and instruction-set randomization (MTD in instruction sets)[15]. This dissertation briefly summarizes these technologies as follows:

1) Network address shuffling is an MTD method that dynamically alters an organization's network by remapping the static association between addresses and systems[12]. For the Internet, addresses are a combination of IP and transport layer information (protocol and port numbers). Network address shuffling can shuffle either or both types of information. Taking IP address shuffling as an example, the address of protected devices within a network is replaced with a pseudorandom IP address chosen from the address space available to the administrator. Thus, address shuffling shortens the intervening periods, which can be used for attack reconnaissance.

2) In address space lay-out randomization, computer memory is dynamically remapped to prevent an attacker from reliably discovering the exact layout of a targeted program in memory[14]. Address space lay-out randomization has similarities to network address shuffling

(dynamically moving a targeted item around a fixed space).

3) MTD applied in IPv6 leverages the huge address space of IPv6 to implement dynamic addressing[13]. This moving target IPv6 defense can not only limit an attacker's time to find a vulnerable attack vector, but also add privacy and anonymity to communicating hosts.

4) The MTD platform for cloud-based IT systems leverages the advantage of the cloud-automation framework to make automated changes to the IT system[11]. This is because the cloud-automation framework can capture an IT system's setput parameters and dependencies using a high-level abstraction. Thus, the MTD platform for cloud-based IT systems can replace the running components of the system with fresh new instances.

1.3 A Brief Introduction to Cyber-physical Smart Grid

A smart grid can be treated as a self-sufficient system, which can integrate the generation sources of any type and scale for providing sustainable, reliable, safe and high-quality electricity to all consumers[16]. As a smart grid is much more dynamic than a traditional grid, it requires significantly more control functions and data communication[17]. A smart grid is one of the most complex CPSs to support advanced technologies and dynamic nature. The landscape of the cyber-physical smart grid is undergoing a radical transformation, characterized by growing renewable energy resources, demand diversification, and the integration of information and communication technologies (ICTs)[16]. All these new dynamics underline the importance of sensing, data acquisition, communication and control technologies such as energy management systems (EMS), state estimation (SE), and SCADA systems in the cyber-physical smart grid.

The SCADA system in a power system collects, analyzes, and visualizes the power system data. One key function of the SCADA system is to supervise the whole system in real-time, including meter readings and statuses of sensors. The SCADA system is composed of host servers, human-machine interface, communication devices, outstations hardware, and local substation processors[18]. Local substation processors include Remote Terminal Unite (RTU), Programmable Logic Controller (PLC), and Intelligent Electronic Devices (IES).

The data collected by SCADA is used by SE to estimate the system status such as nodal voltage in the power grid. Based on the system status, various applications in EMS generate control commands, and then the SCADA system sends these commands to remote substation control devices, such as outstation devices. Thus, the SCADA system forms the basis of the EMS[18]. As the SCADA system evolves, there is a growing interest in exploring the security vulnerabilities of the SCADA system over the communication network and internet technologies[19].

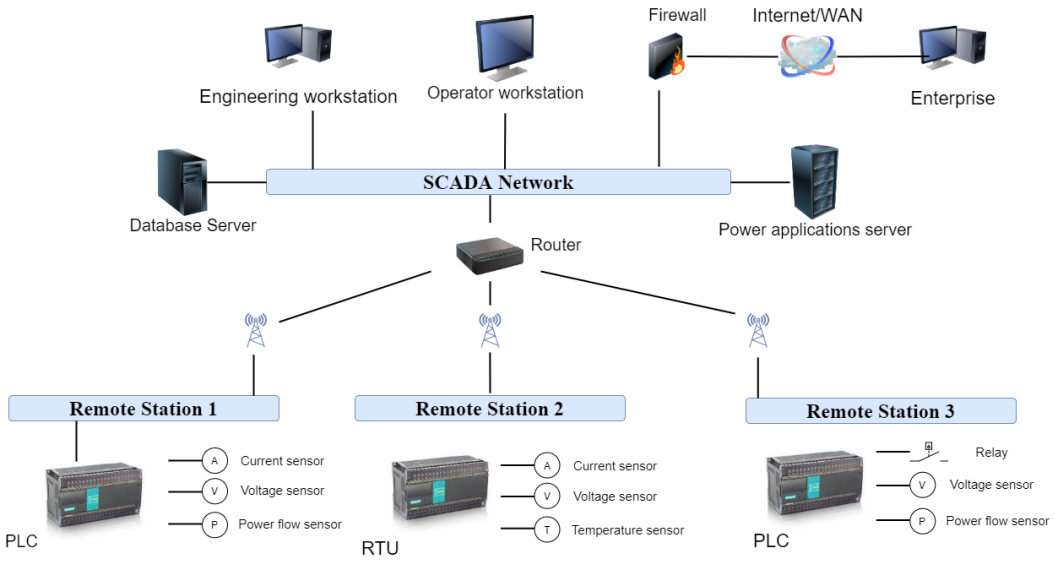


Figure 1.1: *SCADA system network.*

EMS is a system of computer-aided tools used by power grid operators to control and optimize the performance of the grid[20]. As more renewable energy resources, plug-in electric vehicles, and energy storage systems are integrated into the grid, EMS controls the energy flow among the various sources. In addition, EMS is essential to maintain supply-demand balance, satisfying all the system constraints to achieve economic, reliable, and secure operation of the power system[21]. The functions in EMS include contingency analysis, optimal power flow, network reconfiguration, security-constrained unit commitment, automatic generation control (AGC), load forecasting, etc.

SE provides an optimal estimate of the power system voltage based on the available measurements from the SCADA system on the assumed system model[22]. In addition, SE

performs observability analysis, detects bad data, and adds pseudo data if necessary. The estimated state information will be passed on to EMS applications such as AC optimal power flow (ACOPF), contingency analysis, and load forecasting. Furthermore, the same information is presented in corporate offices for further system planning and analysis[22]. Therefore, a reliable SE is critical to maintain the normal function of EMS, which determines the operational security of the power grid.

1.3.1 Cyber-physical Attacks in the Smart Grid

The coupling of the power grid infrastructure with complex computer networks greatly expands the attack surface in the smart grid[23]. Cyber-physical attacks targeting smart grids have resulted in both tremendous economic loss and security issues[24]. To address cybersecurity-related issues in the smart grid, National Institute of Standards and Technology (NIST) provides quantitative notions of risks, threats, vulnerabilities, and attack consequences for power grids[25]. National Electric Sector Cybersecurity Organization Resource (NESCOR) discusses a dozen attack scenarios through impact analyses and assessment[26].

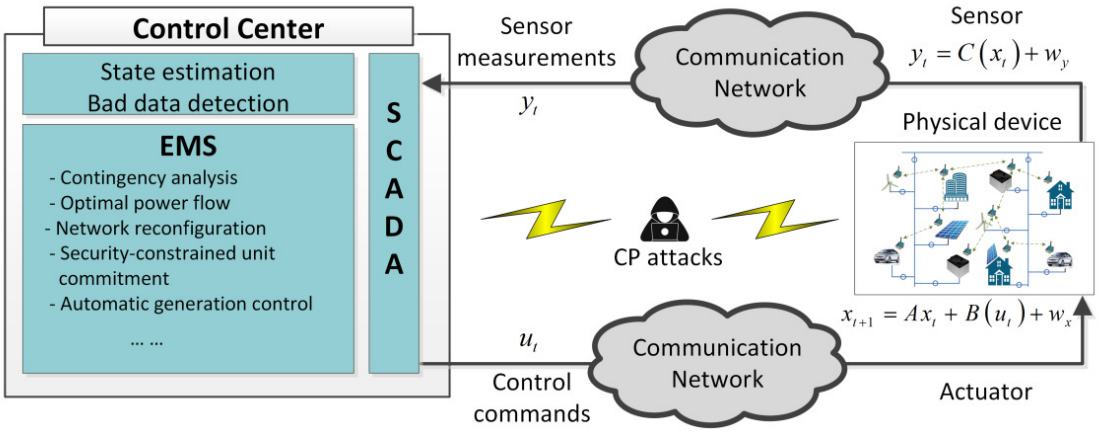


Figure 1.2: *Cyber-physical attacks on smart grids*[27].

The cyber-physical attacks on the smart grid are shown in Figure 1.2. The control center of the power grid sends commands to actuators in grids and receives measurements from field sensors through the communication system. Attackers can perform data integrity attacks by modifying commands or measurements in the communication system. Therefore, the

cyber-physical attacks against smart grids, depending on how a attack is delivered, can be categorized into four types, i.e., control signal attacks, measurement attacks, control-signal-measurement attacks, and communication network attacks. In each type, attacks can be further classified according to their attack objectives, as illustrated in Table 1.1. It is necessary to note that a coordinated attack is more likely to happen in practice. Sophisticated attackers can attack terminal display systems or congest the communication system first to cover ongoing attacks[18]. Additionally, multiple attacks combined together can further exacerbate the adverse consequence of the cyber-physical attacks[2].

Table 1.1: *Taxonomy of the cyber-physical attacks in the smart grid*

Attack Type	Attacks Name	Objective
Commu. network attacks	Byzantine attacks[28]	Reduce the overall communication network performance
	DoS attacks[29;30]	Cause congestion in network communication
Control signal attacks	Aurora attacks[31]	Cause damage to generators, motors and transformers
	Pricing attacks[32;33]	Profitability and mismatch between the generated and the consumed power
Measurement attacks	AGC attacks[33;34]	Rapid decline in the system frequency
	False data injection (FDI)[35–39]	Incorrect voltage estimation
	Blind FDI attacks[40;41]	Incorrect voltage estimation
	Load redistribution attacks[42–45]	Incorrect voltage estimation
	Topology attacks[46;47]	Incorrect topology estimation
	GPS spoofing attack[48;49]	Incorrect timestamp of PMU measurements
Control-signal-measurement attacks	Line outage masking attacks[50–52]	Measurement manipulation to mask line-outage
	Stuxnet-like attacks[5;53]	Incorrect control and measurement signal

1.3.2 Defense Approaches in the Smart Grid

With emerging threats from cyber-physical attacks, cyber-physical defense approaches have been widely studied in the literature. Following the categorization in reference[27], this dis-

sertation classifies these defense approaches into five categories based on their defense mechanisms, as shown in Table 1.2. Here, this chapter briefly introduces them below and more details can be found in reference[27].

Model and algorithmic enhancement as well as data-driven approaches are the cyber-physical defense approaches deployed in the application layer of smart grids. Model and algorithmic enhancement focuses on advancing mathematical detection models, such as adaptive cumulative detector[54], sparsity-based detector[55], and infinity norm-based measurement residual analysis[56]. Data-driven approaches use supervised learning classifiers to detect FDI attacks, such as support vector machine, k-nearest neighbor, and extended nearest neighbor[57]. In addition, deep learning technologies have been applied to detect attacks, such as deep reinforcement learning and recurrent neural network[58].

Securing critical field devices constitutes a natural defense approach. This is because most attacks summarized in Table 1.1 inject malicious data (i.e., control commands and sensor measurements) into field devices through exploiting protocol vulnerability or leakage of confidential information. Reference[59] focused on selecting the sets of protected measurements to detect FDI attacks. Reference[60] took advantage of phasor measurement unit (PMU) devices to detect FDI attacks, and reference[61] studied the optimal placement of PMU devices considering attack detection effectiveness. However, protecting a larger number of field devices or deploying a large number of PMU devices could be prohibitively costly in a real-world cyber-physical smart grid due to its enormous attack surface.

Watermarking and MTD are defense approaches deployed in the physical layer of smart grids. In watermarking defense methods, a known noise is injected into the input of the system as a prob, and then an expected effect of this prob can be found in the system output. Thus, attacks without considering the watermarking can be detected by defenders. Unlike in IT systems, MTD in the physical layer of a smart grid is a proactive defense approach that actively modifies the power grid configurations to invalidate the knowledge of attackers about the system using controllable devices, such as distributed flexible AC transmission system (D-FACTS) devices and load tap changer. In this case, cyber-physical attacks constructed based on the outdated system configuration can be detected by defenders.

The blockchains are defense approaches deployed in the cyber layer of smart grids. Blockchain technology transfers field measurement data and local transaction data in a peer-to-peer manner. The data in blockchain can not be compromised by attackers, unless the adversaries own more than 51% of the devices in the system. Therefore, blockchain technology protects data integrity, confidentiality, availability and accountability in smart grids. The application of blockchain in smart grid can be classified into three levels, i.e., field measurement and communications[62;63], power generations and transmissions[64;65], as well as power distributions and utilization[66–68].

Table 1.2: *Taxonomy of the cyber-physical defense methods in the smart grid*

Detection Type	Deployed Layer
Model and algorithmic enhancement	Application layer
Data-driven approaches	Application layer
Securing measurement sensors	Physical layer
Watermarking	Physical layer
Moving target defense	Physical layer
Blockchain	Cyber layer

1.4 Research Motivations

Cyber-physical incidents mentioned in Background 1.1 and cyber-physical attacks summarized in Table 1.1 pose significant threats to cyber-physical smart grids. These attacks can undermine the grid control system, cause damage to generators, cause blackout in a city, and even trigger cascading failures. As power systems are critical infrastructures to support a functioning society, it is of paramount significance to enhance the cybersecurity of power systems. Governments worldwide have recognized the threats of cyber-physical attacks against the power grid. The U.S. has invested more than $210 million in cybersecurity research since 2010[69]. Canada invested $40 billion in improving power system infrastructures to achieve a reliable and secure power system by 2020[69]. Therefore, it is urgent to study the defense methods to improve the cybersecurity of smart grids. This dissertation help achieve this goal by advancing MTD approaches in the physical layer of smart grids.

MTD planning and MTD operation are two essential stages in the construction of MTDs in the smart grid. First, system operators need to install D-FACTS devices on an appropriately selected subset of transmission lines given certain constraints (e.g., budget). This is referred to as an MTD planning problem. Then, a system operator ought to optimally determine the D-FACTS setpoints under varying load conditions in real-time power system operations. This is referred to as an MTD operation problem. However, the above two distinct stages have not been clearly distinguished in the literature, which can thwart the application of MTD in real-world power grids. Specifically, some MTD-related works simultaneously generate MTD planning and operation solutions according to the load condition. This is impractical since the location of D-FACTS devices would keep changing in a short time period due to the variation in the load. Some other works in the literature ignore the MTD planning stage by assuming all lines are equipped with D-FACTS devices, which is called a full MTD planning. This full MTD planning solution neglects the realistic budget constraints of system operators, resulting in a low utilization rate of the D-FACTS devices. Therefore, addressing MTD planning and MTD operation in a separate yet synergistic fashion is an urgent task.

There are three metrics in the literature to evaluate the performance of an MTD, i.e., attack detection effectiveness, application cost, and hiddenness. However, the roles that MTD planning and MTD operation each plays in improving those metrics are still unknown. Consequently, a lack of such understanding will result in improper objectives in devising MTD planning and operation methods. For example, if both MTD planning and operation methods focus on maximizing attack detection effectiveness, the potential economic benefits of D-FACTS devices for reducing system losses will be reduced. Therefore, it is worthwhile to distinguish the role of MTD planning and MTD operation, and to find a good trade-off among those metrics.

This dissertation seeks to address the following fundamental yet unsolved research questions that plague the application of MTDs in the real-world power grid. Note that the rank of the composite matrix is a widely used metric of MTD detection effectiveness. An MTD with the maximum rank of the composite matrix is referred to as a max-rank MTD.

Question 1: How can system operators separate the MTD planning and MTD operation as two independent problems?

Question 2: How does the MTD planning affect the rank of the composite matrix? How can system operators optimally place D-FACTS devices to achieve maximal MTD detection effectiveness while fully considering the economic benefits from D-FACTS devices?

Question 3: How does the MTD operation impact attack detection effectiveness and operational costs of the power grid? How can system operators optimally dispatch the setpoints of D-FACTS devices to achieve economic and cybersecure system operation?

Question 4: How can system operators design MTD planning and operation methods to achieve a proper trade-off among MTD hiddenness, detection effectiveness, and economic benefits simultaneously?

1.5 Research Contributions

This dissertation contributes to separating the MTD planning and MTD operation problems as two independent problems based on graph theory. It clarifies how MTD planning determines the MTD attack detection effectiveness, MTD application costs, and the existence of MTD hiddenness. This dissertation also demonstrates how MTD operation influences these three metrics. Furthermore, it proposes multiple novel MTD planning methods and MTD operation methods, which maximize MTD detection effectiveness, reduce MTD application costs, and achieve the MTD hiddenness. The major contributions of this dissertation are summarized in response to the questions raised in Section 1.4.

Question 1: This dissertation separates the MTD planning and MTD operation based on graph theory analysis.

- Prove that an MTD is a max-rank incomplete MTD if no D-FACTS devices work in idle states, and there exists no loop in either D-FACTS graph or non-D-FACTS graph.
- Propose MTD planning sufficient conditions for both complete and incomplete MTDs to achieve the maximum rank of the composite matrix based on graph theory. Under

the MTD planning satisfying these sufficient conditions, any MTD can guarantee its detection effectiveness, regardless of the non-idle setpoints of D-FACTS devices.

I discuss these contributions in Chapter 3 and in the following article:
B. Liu and H. Wu, "Optimal D-FACTS Placement in Moving Target Defense Against False Data Injection Attacks," *in IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4345-4357, Sept. 2020[70].

Question 2.1: Max-rank MTD planning algorithms are designed for ensuring attack detection effectiveness and minimizing MTD planning cost.

- Propose a necessary condition for a complete MTD and further derive its requirements on the D-FACTS placement. These requirements are used to quickly determine if a complete MTD can be attained and to guide a systematic placement of D-FACTS devices.
- Mathematically prove that the maximum rank of the composite matrix for an incomplete MTD equals the number of transmission lines in a power system.
- Design novel MTD planning algorithms for both complete and incomplete MTDs to achieve the maximum rank of the composite matrix with the minimum number of D-FACTS devices identified. Additionally, the proposed algorithms leverage the concept of power loss sensitivity to account for the economic benefits of D-FACTS devices on the system loss reduction.

I discuss these contributions in Chapter 3 and in the following article:
B. Liu and H. Wu, "Optimal D-FACTS Placement in Moving Target Defense Against False Data Injection Attacks," *in IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4345-4357, Sept. 2020[70].

Question 2.2: A graph-theory-based planning algorithm is designed for maximizing attack detection effectiveness and considering economic benefits of D-FACTS devices.

- Prove the rank of the composite matrix only reflects the minimum number of protected buses. This work shows, for the first time, the rank of the composite matrix, a widely used indicator of the MTD detection effectiveness, is merely the lower bound of attack detection probability (ADP).

- Identify and prove three types of unprotected buses in MTDs, i.e., end buses, non-D-FACTS buses, and buses fully covered by D-FACTS lines, whose reactances are modified using a unity factor. This work highlights the importance of eliminating unprotected buses in MTD planning for improving the ADP upper bound.

- Mathematically derive and prove the ADP range of different MTD planning methods, which is verified by extensive simulations. Furthermore, this work identifies a class of MTD planning solutions with a fixed ADP.

- Propose a novel graph-theory-based MTD planning method which simultaneously maximizes the ADP lower and upper bound though maximizing the rank of the composite matrix and eliminating unprotected buses.

I discuss these contributions in Chapter 3 and in the following article:

B. Liu and H. Wu, "Systematic Planning of Moving Target Defense to Maximize Detection Effectiveness against False Data Injection Attacks in Smart Grid", in *IET Cyber-Physical Systems: Theory and Applications*, in press.

Question 3: ACOPF-based MTD operation model is proposed for reducing MTD operation costs.

- Propose a novel ACOPF-based MTD operation model with the objective of minimizing system losses and generation costs to optimally determine the setpoints of D-FACTS devices. The proposed MTD operation model can be easily integrated into the existing ACOPF model in energy management system in power systems.

- Develop an interior-point solver to resolve the proposed ACOPF-based MTD model by modifying and extending Matlab Interior-Point Solver (MIPS) in MATPOWER

developed for the conventional ACOPF. Specifically, this work derives the gradient and Hessian matrices of the objective function and the constraints in the proposed ACOPF-based MTD model with respect to the line impedance.

I discuss these contributions in Chapter 4 and in the following articles:

B. Liu and H. Wu, "Optimal D-FACTS Placement in Moving Target Defense Against False Data Injection Attacks," in *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4345-4357, Sept. 2020[70].

B. Liu, L. Edmonds, H. Zhang and H. Wu, "An Interior-Point Solver for Optimal Power Flow Problem Considering Distributed FACTS Devices," *2020 IEEE Kansas Power and Energy Conference (KPEC)*, Manhattan, KS, USA, 2020, pp. 1-5[71].

Question 4: Hidden MTD planning algorithm and operation models are proposed for achieving MTD hiddenness and maximizing attack detection effectiveness. For the hidden MTD (HMTD) planning, this work makes the following contributions.

- Derive a sufficient condition to ensure the existence of HMTD and the maximum rank of the composite matrix based on graph theory analysis.
- Propose a depth-first-search-based MTD planning algorithm, in which an HMTD that has the maximum rank of the composite matrix and covering all necessary buses can be constructed.

For the HMTD operation, this work makes the following contributions.

- Derive a novel and explicit hiddenness condition in HMTD, which can be easily integrated into MTD operation methods.
- Demonstrate the characteristics of voltage angle changes in HMTD, which bridge the HMTD operation and HMTD planning.

- Propose an optimization-based DC-HMTD operation model that maximizes the reactance changes. This model overcomes the drawbacks of the existing HMTD operation algorithm and obtains the D-FACTS setpoints more efficiently.
- Propose an ACOPF-based HMTD operation model that minimizes the generation cost and presents a trade-off between the generation cost and the MTD hiddenness.

I discuss these contributions in Chapter 5 and in the following article:

B. Liu and H. Wu, "Optimal Planning and Operation of Hidden Moving Target Defense for Maximal Detection Effectiveness," Manuscript under the second review in *IEEE Transactions on Smart Grid*.

1.6 Organization of This Dissertation

Chapter 2 provides the background knowledge of SE, FDI attacks, and MTD model as preliminaries for the follow-up sections. Besides, Chapter 2 presents a literature review on existing MTD work. In Chapter 3, max-rank MTD planning algorithms and graph-theory-based MTD planning algorithm are proposed to maximize the MTD detection effectiveness. Chapter 4 proposes an ACOPF-based MTD operation algorithm and derives the gradient and Hessian matrix of objective and constraints with respect to line reactance. Chapter 5 proposes a novel hidden MTD operation condition, a hidden MTD planning algorithm, and DC- and AC-hidden MTD operation models. Concluding remarks and future research directions are discussed in Chapter 6.

Chapter 2

Fundamentals and Related Literature

This chapter provides the background knowledge of state estimation (SE), bad data detection (BDD), false data injection (FDI) attacks, and moving target defense (MTD) as preliminaries for the follow-up sections. This chapter also provides state-of-the-art literature review on existing MTD works in power systems.

2.1 State Estimation and Bad Data Detection

This section provides SE formulation in both DC and AC power system models for two reasons. Firstly, SE plays an important role in power systems, as it provides the estimated grid voltage to applications in energy management system (EMS). Secondly, SE is the attack target of FDI attacks. Having the knowledge of SE is critical to understand the construction and detection of FDI attacks.

2.1.1 DC-SE Formulation and BDD

DC flow analysis is faster and more robust than its AC counterpart[37;72], and has been widely used in power system planning with demonstrated accuracy. In the DC-SE, the system states, i.e., nodal voltage angles $\mathbf{x} \in \mathbb{R}^{n-1}$ are estimated by a set of measurements $\mathbf{z} \in \mathbb{R}^m$ corresponding to nodal power injections and branch power flows. The measurements and

states are related as

$$\mathbf{z} = \mathbf{H} \cdot \mathbf{x} + \mathbf{e}$$

where $\mathbf{e}$ is the measurement noises assumed to be Gaussian distributed with zero mean and a diagonal covariance matrix $\mathbf{W} = diag(\sigma_1^{-2}, \sigma_2^{-2}, ..., \sigma_m^{-2})$, and $\mathbf{H}$ is the measurement Jacobian matrix containing the system topology and system configuration information. The system states can be estimated by solving the following weighted least square (WLS) optimization[22]:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} (\mathbf{z} - \mathbf{H} \cdot \mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H} \cdot \mathbf{x})$$

DC-SE has a closed-form solution as follows:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{z}$$

After estimating the voltage, BDD is applied to examine the existence of faulty measurements, which could result in significant errors in the estimated states. This dissertation uses the 2-norm estimation residual to determine the existence of bad data. As the measurement noises follow a Gaussian distribution, the estimation residual will follow a chi-square distribution $\chi^2_{(m-n)}$, where $m - n$ is the degree of freedom. More specifically, measurements are free of bad data if the inequality $\gamma = ||\mathbf{z} - \mathbf{H} \cdot \hat{\mathbf{x}}||_2 < \gamma_{th}$ holds, where $\gamma_{th} = \chi^2_{(m-n),\alpha}$ is a preset threshold to ensure BDD to have a false alarm rate at $1 - \alpha$. If $\gamma = ||\mathbf{z} - \mathbf{H} \cdot \hat{\mathbf{x}}||_2 > \gamma_{th}$ holds, it indicates the existence of bad measurement in $\mathbf{z}$.

2.1.2 AC-SE Formulation

In the AC power flow model of an n-bus power system, the nonlinear measurement model relating the measurements vector $\mathbf{z} \in \mathbb{R}^m$ and the state vector $\mathbf{x} \in \mathbb{R}^{2n-1}$ comprising voltage magnitudes and angles can be formulated as:

$$\mathbf{z} = h(\mathbf{x}) + \mathbf{e}$$

where $h(\mathbf{.}) : \mathbb{R}^{2n-1} \rightarrow \mathbb{R}^m$ is a vector of nonlinear functions which are based on the type of measurements. To provide the optimal estimate of the system states, the WLS estimator[22] minimizes the weighted least square criterion as follows:

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

It is customary to use the Gauss-Newton iterative algorithm to solve the optimization problem as the following equation. The iterative process converges when the difference between the system states in two iterative is smaller than a pre-determined threshold.

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

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

2.2 FDI Attacks against SE

FDI attacks have become a growing threat to modern power systems. The illustration of FDI attacks is demonstrated in Figure 2.1. FDI attacks inject malicious data $\mathbf{a}[t]$ into the supervisory control and data acquisition (SCADA) measurements $\mathbf{z}[t]$ based on the DC- or AC-FDI attack model. The compromised measurements $\mathbf{z}_a[t]$ received by the SCADA system can bypass BDD without alerts and result in a bias in the estimated voltage. The estimated voltage will be used in the applications of EMS, including the load estimation and optimal power flow (OPF) model.

In the literature, many studies have focused on modeling FDI attacks in DC and AC power system models[36;37;73], constructing FDI attacks with concrete attack objectives[45;74], detecting and preventing FDI attacks[59–61;75;76], and evaluating the consequences of FDI attacks[74]. The attack surfaces in smart grids for FDI attacks are huge, and FDI attacks could result in severe consequences. Firstly, falsified state estimation results could potentially mislead the operation and the auto-control mechanism of EMS[33]. Secondly, FDI attacks can attack

energy markets[77;78], which can represent a serious financial deviation. Finally, sophisticated FDI attacks can cause consequences on power grid operation conditions, such as transmission line over-load[74] and nodal voltage violation[45], resulting in blackout and even cascading failures. Therefore, it is of great significance to detect and thwart FDI attacks in power systems.

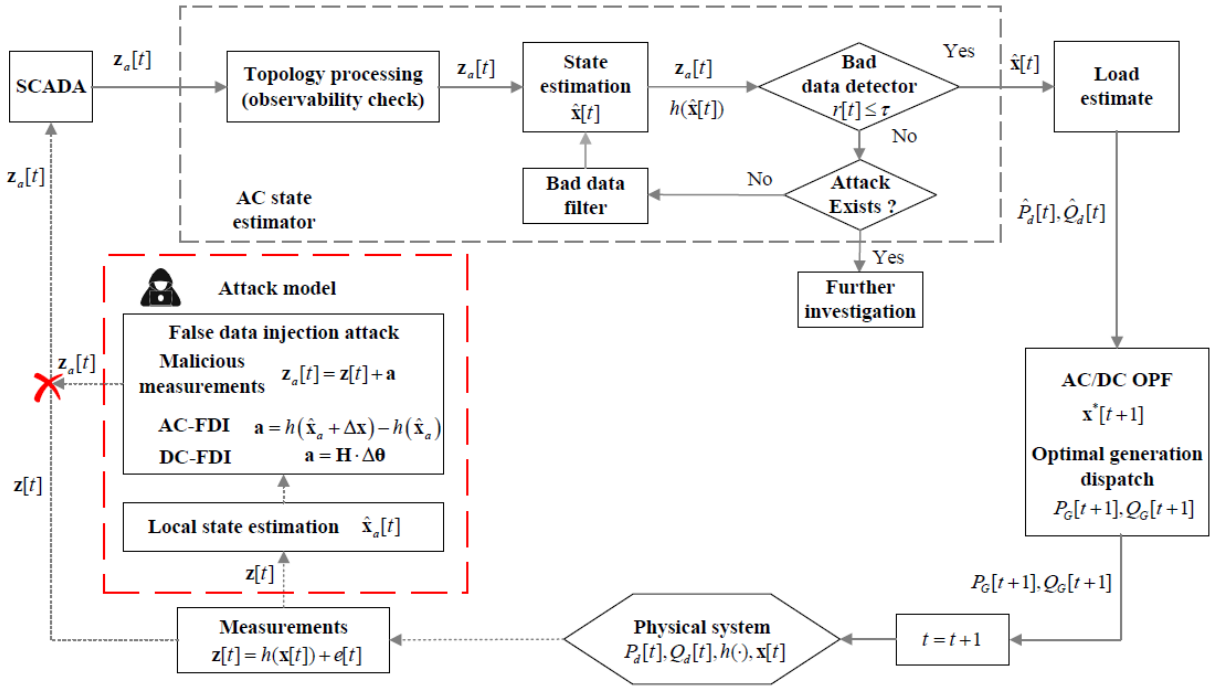


Figure 2.1: *Illustration of FDI attack model in the smart grid.*

2.2.1 DC-FDI Attacks

To construct and launch a successful FDI attack, attackers need to specify a state increment, i.e., $\Delta\theta$. Specifically, attackers must determine which buses to inject malicious voltage angle increment and specify the concrete value of this voltage angle increment. Then, attackers need to calculate the attack vector $\mathbf{a}$ and inject this attack vector into SCADA measurements, i.e., $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$. In the DC power system model, an FDI attack can compromise estimated states without being detected by BDD[36], if the attack vector $\mathbf{a}$ is calculated by $\mathbf{a} = \mathbf{H} \cdot \Delta\theta$. This is because the estimated residual remains the same before and after FDI attacks, shown

as follows:

$$\begin{aligned}\left\|\mathbf{z}_a - \mathbf{H}\cdot\hat{\theta}_a\right\|_2 &= \left\|\mathbf{z}+\mathbf{H}\Delta\theta - \mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}(\mathbf{z}+\mathbf{H}\Delta\theta)\right\|_2 \\ &= \left\|\mathbf{z} - \mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z}\right\|_2 = \gamma < \gamma_{th}\end{aligned}$$

The FDI attack can bypass BDD and falsify SE as long as the attack vector belongs to the columns space of $\mathbf{H}$, i.e., $\mathbf{a} \in col(\mathbf{H})$. This requires the attacker's knowledge about $\mathbf{H}$ or the estimated $\mathbf{H}$ from historical measurements[40]. In either case, the attack vector can be equivalently expressed as $\mathbf{a} = \mathbf{H}\cdot\Delta\theta$[79].

2.2.2 AC-FDI Attacks

The construction of AC-FDI attacks is similar to that in the DC model. In the AC model, system states include both voltage magnitudes and voltage angles. After the state increment $\Delta\mathbf{x}$ is chosen by attackers, attack vector $\mathbf{a}$ can be calculated as follows:

$$\mathbf{a} = h\left(\hat{\mathbf{x}}+\Delta\mathbf{x}\right) - h\left(\hat{\mathbf{x}}\right)$$

In this case, the estimation residual under the FDI attack is the same as the estimation residual without FDI attacks, which ensures that the FDI attack is stealthy to the system operator.

$$\begin{aligned}&\left\|\mathbf{z}_a - h\left(\hat{\mathbf{x}}+\Delta\mathbf{x}\right)\right\|_2 = \left\|(\mathbf{z}+\mathbf{a}) - h\left(\hat{\mathbf{x}}+\Delta\mathbf{x}\right)\right\|_2 \\ &= \left\|\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2+\mathbf{a}_2\end{pmatrix} - \begin{pmatrix}h_1\left(\hat{\mathbf{x}}_1\right)\\ h_2\left(\hat{\mathbf{x}}_1,\hat{\mathbf{x}}_2+\Delta\mathbf{x}\right)\end{pmatrix}\right\|_2 \\ &= \left\|\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix} - \begin{pmatrix}h_1\left(\hat{\mathbf{x}}_1\right)\\ h_2\left(\hat{\mathbf{x}}_1,\hat{\mathbf{x}}_2\right)\end{pmatrix}\right\|_2 \\ &= \left\|\mathbf{z} - h\left(\hat{\mathbf{x}}\right)\right\|_2 < \gamma_{th}\end{aligned}$$

where $\hat{\mathbf{x}}_1$ represents part of the system states which are not altered by the attacker and $\hat{\mathbf{x}}_2$ denotes the system states which are maliciously manipulated by the attacker with an state

increment vector $\Delta\mathbf{x}$; $\mathbf{z}_1$ is the vector of measurements free from manipulation, and $\mathbf{z}_2$ is the vector of measurements maliciously altered by the attacker with an FDI attack vector $\mathbf{a}_2$, i.e., $\mathbf{a}_2 = h_2(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 + \Delta\mathbf{x}) - h_2(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2)$.

Based on the DC- and AC-FDI attack models, the knowledge and capability of attackers are summarized as follows. The attackers have the capability to eavesdrop on the measurements and inject the attack vector into the SCADA measurements. The attackers need the knowledge of the power system topology and line parameters. Note that attackers need system states to construct AC-FDI attacks, while system states are not needed in the construction of DC-FDI attacks[37].

2.3 MTD Model in the Smart Grid

Recently, the concept of MTD has been introduced in the physical layer of power systems in the face of emerging FDI attacks. Different from the MTD in IT systems which highlights the changes in the network layer or the data link layer, MTD in power systems requires physical devices and extra control. MTD in power systems actively perturbs the branch equivalent impedance to invalidate attackers' knowledge about the power system configurations, which are essential for constructing stealthy FDI attacks[70]. Note that all devices which can modify the branch equivalent impedance in real-time can be used in MTD in smart grids. This dissertation takes distributed flexible AC transmission system (D-FACTS) device as an example. D-FACTS devices, such as Static Var Compensators (SVC), Thyristor Controlled Series Capacitors (TCSC), and Static Synchronous Series Compensators (SSSC), are originally utilized to control power flows, manage the power congestion, and minimize system losses by altering the impedance of power lines[80]. With the proliferation of D-FACTS devices[81], their add-on cyber-physical security benefits via MTD have attracted increasing attention in the research community.

MTD takes advantage of D-FACTS devices to create uncertainties for attackers through periodically modifying the setpoints of D-FACTS devices. The illustration of the MTD and FDI attack model is shown in Figure 2.2. The system operators determine the lines to install

D-FACTS devices at the MTD planning stage and then determine the D-FACTS setpoints in real-time at the MTD operation stage. Encrypted commands from the system operator's control room can be securely transmitted to the D-FACTS gateway for its setpoint changes through DNP3, IEC 61850, and 60870-5-104[81]. Accordingly, the incremental reactance of line i–j can be modified by the D-FACTS device within the following range:

$$|x_{ij} - x_{ij}^0| \leq |\eta x_{ij}^0|$$

where x_{ij}^0 and x_{ij} are the line reactance before and after MTD, respectively; the upper bound $\eta = 20\%$, generally referred to as the MTD magnitude, reflects the physical capacity of D-FACTS devices[72;79]. Thus, the susceptance of the transmission lines, the reciprocal of the line reactance, become $b[t+1]$ at time $t+1$.

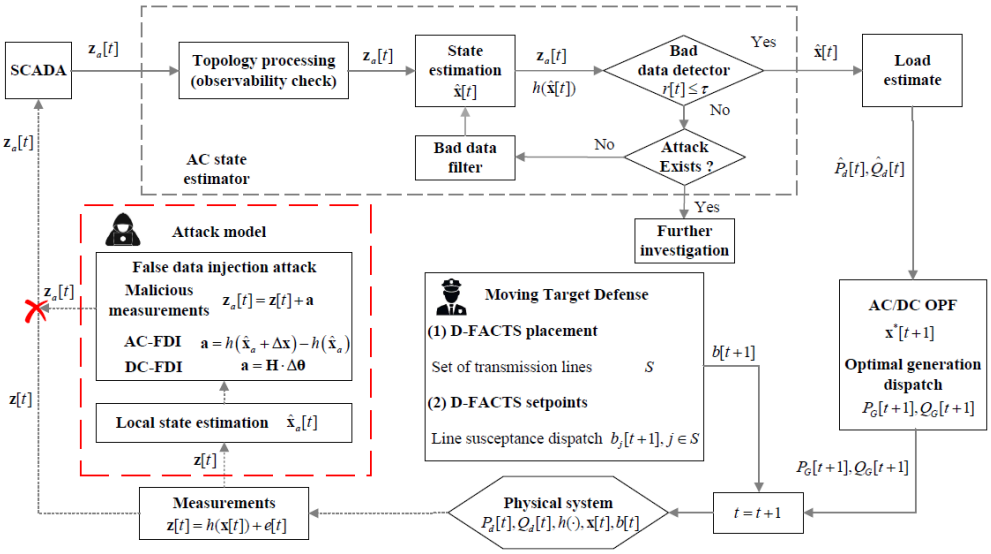


Figure 2.2: *Illustration of MTD and FDI attack model in the smart grid.*

Consequently, the measurement matrix $\mathbf{H}$ used in the DC-SE becomes a time-variant matrix. If attackers construct FDI attacks based on outdated knowledge of $\mathbf{H}$, the estimation residual in the defender's BDD can become larger, which gives the defender chance to detect FDI attacks.

2.4 State-of-the-art MTD Literature Review

The concept of MTD was first introduced into the physical layer of the power system by Morrow et al.[82] and Davis et al.[83]. Most MTD approaches in the literature are designed to detect FDI attacks against SE[72;79;82–88], as summarized in Table 2.1. In addition, the MTD approaches have been recently applied to detect coordinated FDI attacks[89] and Stuxnet-like attacks against power grids[53], in which fake sensor measurements are injected to cover the ongoing attacks on the control signals. Besides D-FACTS devices, MTD also utilizes inverter-based distributed energy resources (DERs) in distribution systems to create low magnitude perturbation signals in voltage, and then a developed detection mechanism can check the perturbation sequence in sensors[90].

Table 2.1: *Moving target defense in the smart grid*

MTD Algorithm	MTD planning	MTD operation	Characteristics
Random MTD[84DC]	Arbitrary planning	Random selection	Detection effectiveness is not considered
OPF-based MTD[85DC]	N/A	DCOPF-based operation	Minimize generation cost and guarantee detection effectiveness[85]
Spanning-tree MTD[87DC]	Spanning-tree planning	Random selection	Cover all buses, but max-rank MTD is not ensured
Max-rank MTD[79DC, 88DC]	Full planning[79], max-rank planning[88]	Optimization-based operation[79]	Minimize system losses[79]. Guarantee max-rank MTD based on numerical methods[79;88]
Hidden MTD[72DC, 86DC, 91AC]	Planning enumeration[72]; max-rank planning using protected meters[86]	Random selection subject to hidden condition[72]	Max-rank hidden MTD, but paper[86] uses extra protected meters

2.4.1 MTD Planning Approaches

In D-FACTS-based MTD approaches, MTD planning and MTD operation are two essential stages in the construction of an MTD. In the MTD planning, system operators need to install

D-FACTS devices on an appropriately identified subset of transmission lines, namely solving the problem of D-FACTS planning. Arbitrary planning and full planning are the two simplest D-FACTS planning strategies. Arbitrary planning randomly selects a subset of lines to install D-FACTS devices[84]. Full planning is the most expensive method in which D-FACTS devices are installed on every transmission line[79]. However, the detection effectiveness of MTDs under these two MTD planning methods is not considered. Spanning-tree planning[87] installs D-FACTS devices on the lines which form a spanning tree of the system. MTDs under the spanning-tree planning are effective in detecting single-bus, uncoordinated multiple-bus, and coordinated multiple-bus FDI attacks. However, this dissertation will demonstrate in Chapter 3 that the detection effectiveness in the MTD under the spanning-tree planning is not guaranteed, which is also influenced by the D-FACTS setpoints. Max-rank planning[88] can make MTDs achieve the maximum rank of the composite matrix (i.e., max-rank MTDs), a metric of the detection effectiveness.

2.4.2 MTD Operation Approaches

After the allocation of D-FACTS devices, the system operators need to continuously determine the D-FACTS setpoints under different load conditions in the MTD operation. The MTD operation mainly includes four methods. Firstly, random selection is the simplest operation method without any computational overhead, in which the D-FACTS setpoints are randomly perturbed[84]. Secondly, as D-FACTS devices are originally used to control the power flow, an OPF-based MTD operation method in the DC model integrates the D-FACTS devices into the OPF model to minimize generation costs[85]. Thirdly, the optimization-based MTD operation takes both the economic cost and the detection effectiveness into account, in which the metric of detection effectiveness is maximized or taken as constraints[79;88]. Finally, the hidden MTD operation method delicately selects D-FACTS setpoints such that all measurements remain the same after the HMTD is applied[72;86;91]. In this case, vigilant attackers cannot detect the MTD in place using BDD. To find suitable MTD planning for the hidden MTD operation, reference[72] enumerates all combination of D-FACTS device placements,

while reference[86] uses the max-rank planning[88] with the help of protected meters.

2.4.3 Evaluation of MTD Performance

In the literature, there are three important metrics to evaluate the performance of an MTD. First of all, attack detection effectiveness is the most important metric for a defense algorithm. As not all of MTDs are effective in detecting FDI attacks, the feasibility and the limitation of MTD are discussed in reference[87]. Many works focus on improving the MTD attack detection effectiveness through the MTD planning[87;88] and MTD operation[72;79;85]. Two metrics are proposed to measure the detection effectiveness of MTD, namely the Lebesgue measure[85] and the rank of the composite matrix[72;79;88]. The composite matrix rank is superior to the Lebesgue measure in the evaluation of MTD detection effectiveness since it demonstrates the inherent nature of MTD on the FDI attack detection and provides an explicit objective for constructing an effective MTD.

Secondly, the cost of the MTD application is a must-concern for system operators. The MTD cost consists of the planning cost and the operation cost. In the planning cost, the number of D-FACTS devices used in MTD determines the capital cost and labor fee. However, the number of D-FACTS devices is not considered in the existing MTD planning methods. In the operation cost, the D-FACTS setpoints impact generation costs and system losses, as these setpoints can change power flow in the system. The DCOPF-based operation in reference[85] presents a trade-off between generation costs and MTD detection effectiveness. A larger MTD detection effectiveness can be gained at the sacrifices of more generation costs. However, generation costs are not considered in the existing AC-MTD operation models.

Thirdly, the hiddenness provides a superior function for MTD as it makes MTD stealthy to attackers. Vigilant attackers use BDD to detect the existence of MTD before launching any attacks. If attackers detect any MTD in place, they may stop FDI attacks and invest more resources to launch data exfiltration attacks and obtain the latest system configuration[72]. MTDs with the hiddenness can mislead these attackers to launch detectable attacks based on incorrect line parameters. In summary, the most desirable MTD should be a hidden MTD

with maximal detection effectiveness and low operation cost.

Chapter 3

MTD Planning Algorithm

This chapter studies the relationship between moving target defense (MTD) planning and MTD detection effectiveness. This work provides us insights into how MTD detection effectiveness can be determined by MTD planning. Furthermore, three MTD planning algorithms are proposed to achieve maximal detection effectiveness step by step.

3.1 Introduction

Although the arbitrary planning, full planning, spanning-tree planning, max-rank planning have been studied in the literature, MTD planning has not been fully addressed for the three reasons. Firstly, MTD detection effectiveness is not or not systemically considered in these planning methods. In arbitrary planning and full planning, the detection effectiveness is not considered at all. Spanning-tree planning claims to be effective in detecting single-bus, uncoordinated multiple-bus, and coordinated multiple-bus false data injection (FDI) attacks. However, its detection effectiveness can be influenced by the distributed flexible AC transmission system (D-FACTS) setpoints. Secondly, the number of D-FACTS devices is not considered in these planning methods. While the price of a D-FACTS device is lower than that of a conventional FACTS device, the cost of D-FACTS devices is still not negligible, especially in a system with thousands of lines[92;93]. Furthermore, even if the D-FACTS

devices are installed on all transmission lines, some of them may only be needed in certain time periods while not at all in others, resulting in a low utilization rate of the D-FACTS devices[72;79;84]. Thirdly, the economic benefits of D-FACTS devices are not considered in these planning methods. System operators install D-FACTS devices mainly for economic considerations in their current operational practices. However, it is conceivable that the cyber-defense benefits against FDI attacks can be viewed as an important by-product of D-FACTS devices in the future smart grid. Therefore, it is necessary to optimally combine the economic and cyber-defense benefits at the planning stage of D-FACTS devices.

This chapter tries to fill these gaps by answering the following fundamental yet unresolved research questions.

1) How is MTD planning related to the MTD detection effectiveness? How can system operator achieve the maximum rank of the composite matrix through MTD planning?

2) In a power system, what is the minimum number of D-FACTS devices required to achieve the maximum rank of the composite matrix in MTDs? Can we establish an analytical approach to quickly attain this number?

3) If the minimum number of D-FACTS devices is identified, given the fact that thousands of candidate planning solutions may exist in a power system, how can system operators find the optimal planning methods to potentially enable the most economic and cyber-secure operation under the growing threats of FDI attacks?

4) Is the maximum rank of the composite matrix strictly equivalent to maximal MTD detection effectiveness? How can system operators accurately reflect the MTD detection effectiveness and improve detection effectiveness through MTD planning. How can system operators measure the detection effectiveness of a given MTD planning?

It is necessary to note that the first three questions are answered in Section 3.3, and the last question is answered in Section 3.4.

3.2 Preliminaries

3.2.1 Notation

We summarize variables frequently used throughout the dissertation in Table 3.1, where bold-faced lower- and upper-case letters stand for vectors and matrices, respectively. Subscript 0 represents variables before the implementation of an MTD. For example, $\mathbf{H}_0$ represents the original measurement matrix before an MTD, while $\mathbf{H}$ stands for the one after implementing the MTD. In addition, variables preceded by Δ represent changes in the variables. Furthermore, we use a superscript $(\cdot)'$ to represent a reduced matrix obtained by removing rows and columns of all zeros.

For presentation simplicity, "D-FACTS lines" and "non-D-FACTS lines" stand for the set of lines equipped with and without D-FACTS devices, denoted by subscript DF and $\overline{DF}$, respectively. "Non-D-FACTS buses" represent the set of buses only connected to non-D-FACTS lines, while the remaining buses in the system are "D-FACTS buses", referring to the set of buses connected to at least one D-FACTS line. An end bus in this paper refers to a bus that is only connected by one single line.

Let G be a graph of a power system composed of all transmission lines and buses. Let G_{DF}, termed as a D-FACTS graph, be a subgraph of G consisting of D-FACTS lines and all buses. Similarly, let $G_{\overline{DF}}$, termed as a non-D-FACTS graph, be a subgraph of G consisting of non-D-FACTS lines and all buses. Let a reduced D-FACTS graph denote a graph only composed of D-FACTS lines and D-FACTS buses.

A D-FACTS device works in an idle state if it is installed on a given line but doesn't modify the line reactance, i.e., $x_{ij} = x_{ij,0}$. For a D-FACTS device such as Static Var Compensators (SVC), Thyristor Controlled Series Capacitors (TCSC), and Static Synchronous Series Compensators (SSSC), its idle state corresponds to zero reactive power compensation. Otherwise, it works in a non-idle state.

Table 3.1: *Nomenclature*

Symbol	Definition
θ	Voltage angle of buses excluding reference bus
$\mathbf{z}$	Measurement vector
$\mathbf{a}$	FDI attack vector
$\mathbf{z}_a$	Compromised measurement vector
$\mathbf{H}$	DC measurement matrix in state estimation
h_i	The i-th column in $\mathbf{H}$ (corresponding to bus i)
$\mathbf{M}$	Composite matrix of $\mathbf{H}_0$ and $\mathbf{H}$
$\mathbf{T}$	Elementary matrix in elementary column operations
$\mathbf{U}$	Elementary matrix in elementary row operations
$\mathbf{A}$	Incident matrix of power system graph
$\mathbf{X}$	Diagonal line reactance matrix
$\mathbf{D}$	Meter deployment matrix
x_{ij}	Reactance of line i–j (between bus i and bus j)
b_{ij}	Susceptance of line i–j, and $b_{ij} = 1/x_{ij}$
n	Total number of system buses
n_1	Total number of D-FACTS buses
n_2	Total number of non-D-FACTS buses
n_e	Total number of end buses
n_{1e}	Total number of end buses belonging to the D-FACTS buses
$n_{2/e}$	Total number of non-D-FACTS buses excluding end buses
m	Total number of measurements
p	Total number of lines
p_1	Total number of D-FACTS-equipped lines
p_2	Total number of lines without D-FACTS
lp	Total number of loops in a graph
t	Total number of disconnected components in a graph
$r(\cdot)$	Matrix rank operator
$col(\cdot)$	Column space operator

3.2.2 MTD Detection Effectiveness Metric

The detection effectiveness of a specific MTD under a specific FDI attack is given as a necessary and sufficient condition[79]. An FDI attack with the attack vector $\mathbf{a} = \mathbf{H}_0 \cdot \mathbf{\Delta}\theta,\ \mathbf{\Delta}\theta \neq 0$ is detectable under the MTD with measurement matrix $\mathbf{H}$ if and only if $\mathbf{H} \cdot \theta \neq \mathbf{H}_0 \cdot \mathbf{\Delta}\theta$ for any θ[79].

The detection effectiveness of a specific MTD can be measured by the rank of its composite matrix $\mathbf{M} = [\mathbf{H}_0\ \mathbf{H}]$. An MTD with a larger rank of the composite matrix has higher

detection effectiveness. Particularly, an MTD with $r(\mathbf{M}) = 2(n-1)$ can detect any FDI attack[79], which is consistent with the definition of a complete MTD in reference[72]. A complete MTD can detect any FDI attack under mild assumptions; however, the completeness is generally unattainable since it requires the total number of transmission lines to satisfies $p \geq 2(n-1)$ and the number of measurements to satisfy $m \geq 2n$[72].

In this section, we use the incremental DC measurement matrix $\Delta\mathbf{H}$, defined as $\Delta\mathbf{H} = \mathbf{H} - \mathbf{H}_0$, to reflect the influence of D-FACTS devices on the power system. Due to $r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r([\mathbf{H}_0\ \mathbf{H}])$, we make the following definition to facilitate our discussion.

Definition 3.2.1 An MTD is a complete MTD, if its rank of the composite matrix $\mathbf{M} = [\mathbf{H}_0\ \Delta\mathbf{H}]$ is equal to $2(n–1)$, i.e., $r(\mathbf{M}) = 2(n-1)$; otherwise, it is an incomplete MTD.

We find out the rank of the composite matrix can be determined by the MTD planning, which will be proved in this chapter. Arbitrary placement of the D-FACTS devices may decrease the rank and negatively influence the MTD detection effectiveness. We focus on the proper construction of $\mathbf{H}$ via the MTD planning to ensure the maximum rank of the composite matrix in MTD. For presentation simplicity, an MTD with the maximum rank of the composite matrix is referred to as a max-rank MTD hereinafter.

3.2.3 Graph Theory for Power System Topology

Graph theory can be used to bridge the D-FACTS placement topology and the MTD effectiveness (i.e., the rank of the composite matrix). We will mathematically prove in Section 3.3 that the rank of the composite matrix is related to the rank of the incidence matrix in the graph composed of D-FACTS lines. Therefore, we can utilize the rank of the incidence matrix to identify the D-FACTS placement topology based on graph theory, leading to the maximum rank of the composite matrix. According to the theorem of the Euler's formula for a disconnected graph[94], for any planar graph G with n nodes, p edges, f faces, and t components, the following equation holds:

$$n + f - p = 1 + t \tag{3.1}$$

where the number of faces equals the sum of the number of interfaces (loops in a graph) and one external face, i.e., $f = lp + 1$. Since loops in the graph correspond to linearly dependent rows of the matrix, for the planar graph G with n nodes and t components, the rank of incidence matrix $\mathbf{A}$ is $n - t$, i.e., $r(\mathbf{A}) = n - t$[95]. Using equation (3.2), the rank of $\mathbf{A}$ indicates the number of disconnected components and the number of loops in a planar graph with n nodes and p edges:

$$r(\mathbf{A}) = n - t = p - lp \tag{3.2}$$

A spanning tree of a graph is a subgraph that contains all nodes of the graph without loops. Thus, the rank of the incident matrix of a spanning tree is n–1. In an edge-weighted undirected graph, the minimum spanning tree (MST) is a spanning tree whose weight (the sum of weights on its edges) is no greater than the weight of any other spanning tree. In this chapter, the power system topology is treated as an edge-weighted graph $G(L, E)$ with buses as nodes L and lines as edges E.

3.2.4 Linear Sensitivity of Transmission Loss to Line Reactance

Traditionally, D-FACTS devices are used to manage power flows and minimize system losses in the power system operation[96]. The power loss to impedance sensitivity (PLIS), which indicates how much the system losses change due to a change in the line impedance, can help determine the most appropriate D-FACTS locations to minimize system losses. In the D-FACTS placement for loss minimization, the best k lines are chosen corresponding to the k sensitivities in PLIS, which are the furthest from zero[96]. The PLIS is calculated as:

$$\frac{dP_{loss}}{dx_{ij}} = \frac{\partial P_{loss}}{\partial P_{flow,ij}} \left[\frac{\partial P_{flow,ij}}{\partial s_{(\theta,V)}} \frac{\partial s_{(\theta,V)}}{\partial x_{ij}} + \frac{\partial P_{flow,ij}}{\partial G_{ij}} \frac{\partial G_{ij}}{\partial x_{ij}} + \frac{\partial P_{flow,ij}}{\partial B_{ij}} \frac{\partial B_{ij}}{\partial x_{ij}} \right] \tag{3.3}$$

where P_{loss} denotes the total real power loss; $s_{(\theta,\, V)}$ is a concatenated vector of all voltages of the system. Each part in equation (3.3) can be calculated using the method in reference[96]. This work assigns the absolute value of PLIS as the weight on each line in $G(L, E)$. Both the

concepts of max-rank MTD and PLIS are synergistically combined into the proposed MTD planning approach to achieve a cyber-secure and economic operation in the power system.

3.3 Max-rank Planning Algorithms

In this section, we first derive analytical necessary conditions for a complete MTD and sufficient conditions for both complete and incomplete MTDs. Then, we design MTD planning algorithms based on the proposed analytical conditions to obtain a max-rank MTD.

3.3.1 Necessary Conditions for a Complete MTD

We decompose the measurement matrix to facilitate the derivation of necessary conditions for a complete MTD. $\mathbf{H}$ and $\mathbf{H}_0$ can be expressed as follows[79]:

$$\mathbf{H} = \mathbf{D} \cdot \mathbf{X} \cdot \mathbf{A}_{-r}^T \tag{3.4}$$

$$\mathbf{H}_0 = \mathbf{D} \cdot \mathbf{X}_0 \cdot \mathbf{A}_{-r}^T \tag{3.5}$$

where $\mathbf{A}_{-r} \in \mathbb{R}^{n-1 \times p}$ is the reduced bus-branch incidence matrix of the power system by removing the row of the slack bus; $\mathbf{X}_0 \in \mathbb{R}^{p \times p}$ and $\mathbf{X} \in \mathbb{R}^{p \times p}$ are the diagonal reactance matrix before and after MTD, respectively; $\mathbf{D} \in \mathbb{R}^{m \times p}$ is the meter deployment matrix. In a fully measured power system, $\mathbf{D}$ is of full column rank since $\mathbf{D} = [\mathbf{I} \quad -\mathbf{I} \quad \mathbf{A}^T]^T$ where $\mathbf{I} \in \mathbb{R}^{p \times p}$ is an identity matrix. Note that if the system is not fully measured, the decomposition becomes $\mathbf{H} = \mathbf{C} \cdot \mathbf{D} \cdot \mathbf{X} \cdot \mathbf{A}_{-r}^T$, where $\mathbf{C}$ is a meter selection matrix defined in reference[79]. As long as $\mathbf{C} \cdot \mathbf{D}$ is of full column rank, the conclusions in a fully measured system can be extended to partially measured systems.

According to the definition of $\Delta\mathbf{H}$, $\Delta\mathbf{H} = \mathbf{D} \cdot \Delta\mathbf{X} \cdot \mathbf{A}_{-r}^T$ holds where $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}_0$. We further decompose $\Delta\mathbf{H}$ only using D-FACTS lines to gain more insights into the requirements of the D-FACTS placement. Since zero diagonal elements in $\Delta\mathbf{X}$ correspond to the non-D-FACTS lines, we separate the D-FACTS and non-D-FACTS lines by performing elementary

column and row exchanges on $\Delta\mathbf{X}$ such that $\mathbf{U}\cdot\Delta\mathbf{X}\cdot\mathbf{T}=\begin{bmatrix}\Delta\mathbf{X}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$, where $\Delta\mathbf{X}'\in\mathbb{R}^{p_1\times p_1}$ is a diagonal matrix of D-FACTS lines with full rank, i.e., $r\left(\Delta\mathbf{X}'\right)=p_1$. Accordingly, $\Delta\mathbf{H}$ can be expressed as:

$$\Delta\mathbf{H}=(\mathbf{D}\cdot\mathbf{U}^{-1})\cdot(\mathbf{U}\cdot\boldsymbol{\Delta}\mathbf{X}\cdot\mathbf{T})\cdot(\mathbf{T}^{-1}\cdot\mathbf{A}_{-r}^{T})\ =[\mathbf{D}_1\ \mathbf{D}_2]\begin{bmatrix}\Delta\mathbf{X}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{A}_1^T\\ \mathbf{A}_2^T\end{bmatrix}=\mathbf{D}_1\cdot\Delta\mathbf{X}'\cdot\mathbf{A}_1^T \quad (3.6)$$

where $\mathbf{A}_1^T\in\mathbb{R}^{p_1\times n-1}$ and $\mathbf{A}_2^T\in\mathbb{R}^{p_2\times n-1}$ are the upper and lower submatrix of $\mathbf{T}^{-1}\cdot\mathbf{A}_{-r}^T$; $\mathbf{D}_1\in\mathbb{R}^{m\times p_1}$ and $\mathbf{D}_2\in\mathbb{R}^{m\times p_2}$ are the left and right submatrix of $\mathbf{D}\cdot\mathbf{U}^{-1}$, respectively. To illustrate the aforementioned decomposition, a fully measured 3-bus system with Bus *3* being the reference bus[22] is shown in Figure 3.1, in which the dotted lines *1–2* and *2–3* are the D-FACTS lines. $\mathbf{H}_0$ can be expressed as:

$$\mathbf{H}_0=\mathbf{D}\cdot\mathbf{X}_0\cdot\mathbf{A}_{-r}^T=\begin{bmatrix}\mathbf{I}\\ -\mathbf{I}\\ \mathbf{A}\end{bmatrix}\cdot\begin{bmatrix}1/x_{12}^0 & & \\ & 1/x_{13}^0 & \\ & & 1/x_{23}^0\end{bmatrix}\cdot\begin{bmatrix}1 & -1\\ 1 & 0\\ 0 & 1\end{bmatrix} \quad (3.7)$$

Suppose the reactance of both D-FACTS lines are changed in MTD, we have $\Delta\mathbf{X}=diag(1/x_{12}-1/x_{12}^0,\ 0,\ 1/x_{23}-1/x_{23}^0)$ and $\Delta\mathbf{X}'=diag(1/x_{12}-1/x_{12}^0,\ 1/x_{23}-1/x_{23}^0)$. The column and row

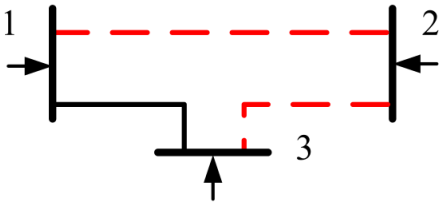


Figure 3.1: *A fully measured 3-bus system.*

operations not only separate D-FACTS and non-D-FACTS lines in $\Delta\mathbf{X}$, but also D-FACTS and non-D-FACTS lines in D and $\mathbf{A}_{-r}^T$ accordingly. Therefore, $\mathbf{D}_1$ and $\mathbf{D}_2$ are the meter deployment matrix of G_{DF} and $G_{\overline{DF}}$, respectively. $\mathbf{A}_1$ and $\mathbf{D}_2$ are the reduced bus-branch

incidence matrix of G_{DF} and $G_{\overline{DF}}$ without the row of the slack bus, respectively. From equation (3.2), the following equations hold:

$$r(\mathbf{A}_1) = n - t_{DF} = p_1 - lp_{DF} \tag{3.8}$$

$$r(\mathbf{A}_2) = n - t_{\overline{DF}} = p_2 - lp_{\overline{DF}} \tag{3.9}$$

We propose the following proposition for a complete MTD, whose proof is given in Appendix A.1.

Proposition 3.3.1. Necessary conditions for a complete MTD: For a complete MTD, the following inequality holds

$$\min\left\{r(\mathbf{D}_1), r(\Delta\mathbf{X}'), r(\mathbf{A}_1^T)\right\} \geq n-1 \tag{3.10}$$

Remarks:

- With $\Delta\mathbf{X}'$ being a diagonal matrix, $r(\Delta\mathbf{X}') = p_1 \geq n-1$ indicates that there are at least $n-1$ D-FACTS lines in a complete MTD.
- Due to $p_1 \geq n-1$ and $\mathbf{A}_1^T \in \mathbb{R}^{p_1 \times n-1}$, we have $r(\mathbf{A}_1^T) \leq n-1$. In addition, we have $r(\mathbf{A}_1^T) \geq n-1$ according to Proposition 3.3.1. Therefore, $\mathbf{A}_1^T$ is with the full column rank, i.e., $r(\mathbf{A}_1^T) = n-1$. The full column rank of $\mathbf{A}_1^T$ suggests $t_{DF} = 1$ and $lp_{DF} = p_1 - n + 1$ according to equation (3.8). Hence, G_{DF} is a connected graph with $p_1 - n + 1$ loops.
- Measurement-to-transmission-line mapping rules[97], originally designed for studying the observability in the SE, are utilized to explain the relationship between sensor deployment and D-FACTS placement in a complete MTD. In these rules, each measurement must be assigned to one line: a flow measurement of a line is assigned to itself, and an injection measurement of a bus is assigned to any of the lines connected to the bus. $r(\mathbf{D}_1) \geq n-1$ indicates that for at least $n-1$ D-FACTS lines, each of these lines must

be mapped to a different sensor. Specifically, a D-FACTS device can contribute to improving the rank of the composite matrix when a sensor is placed associated with this D-FACTS line under the measurement-to-transmission-line mapping rules. Without a mapped sensor, a line with the D-FACTS device cannot contribute to the rank of the composite matrix.

The requirements of the sensor deployment and D-FACTS placement on a complete MTD are summarized as follows: 1) the sensor deployment makes at least $n-1$ D-FACTS lines satisfy measurement-to-transmission-line mapping rules; and 2) there are at least $n-1$ D-FACTS lines that reach all buses. Specifically, when there are only $n-1$ D-FACTS lines, they should form a spanning tree. The spanning-tree topology of D-FACTS placement can be elucidated as follows. In the DC power flow model, there are $n-1$ free variables, including all voltage angles except for the slack bus. In an observable system, the number of equations between measurements and states, i.e., $\mathbf{z} = \mathbf{H} \cdot \theta$, can be reduced to $n-1$ by removing all redundant measurements. Thus, only $n-1$ independent line reactance variables are necessary to define these system states. Any $n-1$ independent lines can constitute a spanning tree for a connected grid.

3.3.2 Sufficient Conditions for Complete and Incomplete MTDs

We assume that the system is fully measured with $r(\mathbf{D}) = p$[79;84;85]. As discussed earlier, the rank of the composite matrix is indicative of the MTD effectiveness, whose upper bound for a complete MTD is $2(n–1)$. The upper bound for an incomplete MTD is discussed in the following lemmas, the proofs of which are provided in Appendix A.2 and A.3.

Lemma 3.3.1 If the number of lines in a system is less than twice the number of states, i.e., $p < 2(n–1)$, the maximum rank of the composite matrix equals the number of lines, i.e., $\max\{r([\mathbf{H}_0\ \mathbf{H}])\} = p$.

To build a max-rank MTD, we decompose $\mathbf{H}_0$ into two parts, i.e., $\mathbf{H}_0 = \mathbf{H}^0_{DF} + \mathbf{H}^0_{\overline{DF}}$,

where $\mathbf{H}^0_{DF}$ and $\mathbf{H}^0_{\overline{DF}}$ are decomposed as follows:

$$\mathbf{H}^0_{DF} = \mathbf{D} \cdot \mathbf{X}^0_{DF} \cdot \mathbf{A}^T_{-r} \tag{3.11}$$

$$\mathbf{H}^0_{\overline{DF}} = \mathbf{D} \cdot \mathbf{X}^0_{\overline{DF}} \cdot \mathbf{A}^T_{-r} \tag{3.12}$$

$$\mathbf{X}_0 = \mathbf{X}^0_{DF} + \mathbf{X}^0_{\overline{DF}} \tag{3.13}$$

where $\mathbf{X}^0_{DF} \in \mathbb{R}^{p \times p}$ is a diagonal reactance matrix whose diagonal elements are original reactance for D-FACTS lines, and zero for non-D-FACTS lines. Conversely, diagonal elements in $\mathbf{X}^0_{\overline{DF}} \in \mathbb{R}^{p \times p}$ are zero for D-FACTS lines and the original reactance for non-D-FACTS lines. For instance, we have $\mathbf{X}^0_{DF} = diag(1/x^0_{12},\ 0,\ 1/x^0_{23})$, $\mathbf{X}^0_{\overline{DF}} = diag(0,\ 1/x^0_{13},\ 0)$ in the 3-bus system as shown in Figure 3.1.

Since $\mathbf{H}^0_{DF}$ and $\Delta\mathbf{H}$ are composed of the same set of D-FACTS lines, $\mathbf{H}^0_{DF}$ can be directly expressed by $\Delta\mathbf{H}$, if graph G_{DF} has no loops. The connection between $\mathbf{H}^0_{DF}$ and $\Delta\mathbf{H}$ is critical to derive the following lemma.

Lemma 3.3.2 If there exists no loop in G_{DF}, the following equation holds, $r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r([\mathbf{H}^0_{\overline{DF}}\ \Delta\mathbf{H}]) = r(\mathbf{A}_1) + r(\mathbf{A}_2)$.

Lemma 3.3.2. correlates the rank of the composite matrix with the rank of incidence matrices in G_{DF} and $G_{\overline{DF}}$ based on the assumption of a loopless graph G_{DF}. The correlation can be further extended as follows:

$$\begin{aligned} r(\mathbf{M}) &= r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r(\mathbf{A}_1) + r(\mathbf{A}_2) \\ &= (p_1 - lp_{DF}) + (p_2 - lp_{\overline{DF}}) = p - lp_{DF} - lp_{\overline{DF}} \\ &= n - t_{DF} + n - t_{\overline{DF}} = 2n - (t_{DF} + t_{\overline{DF}}) \end{aligned} \tag{3.14}$$

Based on equation (3.14), we propose the following corollaries that constitute sufficient conditions for the max-rank incomplete and complete MTD.

Corollary 3.3.1. Sufficient conditions for a max-rank incomplete MTD: In a fully measured

system with $p < 2(n-1)$, an MTD is a max-rank incomplete MTD, if the D-FACTS placement ensures that there exists no loop in either G_{DF} or $G_{\overline{DF}}$.

Corollary 3.3.2. Sufficient conditions for a max-rank complete MTD: In a fully measured system with $p \geq 2(n–1)$, an MTD is a max-rank complete MTD, if the D-FACTS placement ensures that: 1) G_{DF} is a connected graph consisting of n nodes and $n–1$ edges without any loops, i.e., a spanning tree of the system; and 2) $G_{\overline{DF}}$ is a connected graph consisting of n nodes and $p-n+1$ edges.

Again, the proofs of Corollaries 3.3.1 and 3.3.2 are given in Appendix A.4 and A.5. It is worth noting that the above sufficient conditions are built upon a reasonable assumption that the reactance of all D-FACTS lines must be perturbed in an MTD, indicating that no D-FACTS devices ought to work in an idle state. Thus, the MTD planning based on Corollaries 3.3.1 and 3.3.2 ensures a 100% utilization rate of the installed D-FACTS devices, which in turn overcomes the drawback of the low utilization rate in existing MTD studies discussed previously.

Analytically determining the minimum number of D-FACTS devices to realize a max-rank MTD is of particular use for power system planning. Combining Proposition 3.3.1 and Corollary 3.3.2, the minimum number of D-FACTS devices for a complete MTD is $n–1$. In an incomplete MTD, the minimum number of D-FACTS devices corresponds to the maximum number of non-D-FACTS lines due to $p_1 + p_2 = p$. In fact, Corollary 3.3.1 suggests that a loopless non-D-FACTS graph with the maximum number of lines is a spanning tree with $n–1$ lines. Consequently, the minimum number of D-FACTS devices for a max-rank incomplete MTD is determined as:

$$\min\{p_1\} = p - n + 1 \tag{3.15}$$

For instance, a max-rank incomplete MTD can be accomplished when only 37.7% of transmission lines are equipped with D-FACTS devices on the ACTIVSg2000 system, consisting of 2,000 buses and 3,206 transmission lines[98]. To the best of our knowledge, the proposed sufficient conditions, for the first time, provide important guidance on how the

system operator can place a minimum number of D-FACTS devices to achieve a max-rank incomplete or complete MTD, translating into cyber-secure power system operations.

The use of the linearized measurement matrix could be ideal to analyze the detection of stealthy FDI attacks against AC-SE under the MTD. However, it is extremely challenging to provide analytical solutions in AC-SE. To the best of our knowledge, almost all related work[72;79;84;85] utilizes a DC model to analyze the MTD detection effectiveness since there is no explicit metric available in AC-SE to quantitatively measure the MTD detection effectiveness. In addition, the linearized measurement matrix obtained from the Jacobian matrix may be difficult to be decomposed, or at least may not provide any analytical outcomes as useful as equation (3.14) in DC-SE. Therefore, we take the customary approach in this work for analyzing the MTD detection effectiveness, and focus on the max-rank MTD planning in a DC model while being able to quantitatively measure the MTD detection effectiveness.

3.3.3 Max-rank Planning Algorithms

D-FACTS devices can be used in the power system operation for minimizing system losses and simultaneously detecting FDI attacks in MTD. However, D-FACTS placement solutions for serving the above two objectives may be quite different. From the perspective of system economic operation, D-FACTS devices ought to be installed on the lines with the highest PLIS values[96], whereas the D-FACTS placement topology needs to meet Corollaries 3.3.1 or 3.3.2 for achieving the maximum MTD detection effectiveness. Meanwhile, D-FACTS placement for the loss minimization is merely determined by PLIS weights, while that for FDI attack detection is only determined by the D-FACTS placement topology. Therefore, the concept of MST in graph theory can be fully utilized to combine the two distinct perspectives.

We adopt the absolute value of the median PLIS as the graph weight for each line. We propose novel MTD planning algorithms that pick the lines with the large median PLIS values to construct a loopless G_{DF}. Here, we design Algorithm 1 for the incomplete MTD and Algorithm 2 for the complete MTD, based on the MST identified by the Prime's algorithm[99] and the proposed sufficient conditions for the max-rank MTD. The MST in the proposed

algorithms refers to a spanning tree in the topology of the power system, which reaches all buses and has a total PLIS weight no greater than that of any other spanning trees.

With the minimum number of D-FACTS devices identified in equation (3.15), Algorithm 1 first establishes an MST as $G_{\overline{DF}}$, suggesting that the remaining lines in G_{DF} are characterized by large PLIS values. The MST identified only ensures that $G_{\overline{DF}}$ is loopless, but there might be loop(s) in G_{DF}, leading to a D-FACTS placement that fails to meet Corollary 3.3.1. If G_{DF} is loopless, Algorithm 1 stops with the D-FACTS placement solution procured; otherwise, the algorithm iteratively updates G_{DF} and $G_{\overline{DF}}$ by adjusting weights of edges, as shown in rows 5-21 of Algorithm 1, to warrant that both G_{DF} and $G_{\overline{DF}}$ are loopless. A loop in G_{DF} is formed since all edges in the loop have relatively large weights such that none of them are included in the MST. If the weight of one edge in the loop is reduced, an updated MST will include this edge and the loop in G_{DF} will be eliminated accordingly. Particularly, when there are multiple loops in G_{DF}, all edges in the first loop are identified and sorted in an ascending order of their weights. Algorithm 1 decreases the weight of the edge on the top of the order and then updates the MST (see row 10 of Algorithm 1). If this edge belongs to the updated MST, indicating the loop is broken, the edge is moved from G_{DF} to $G_{\overline{DF}}$. If this loop remains, Algorithm 1 restores its weight to the original value and repeat the above steps on the next edge in that ascending order until the loop is broken. Algorithm 1 handles the next loop and repeats the above process until the updated G_{DF} is loopless.

For a complete MTD, Algorithm 2 first negates the weights of all edges (see row 2) and then assigns an MST to G_{DF} such that G_{DF} has the largest absolute values of the median PLIS. The MST only ensures G_{DF} is connected, but there might be isolated nodes in $G_{\overline{DF}}$, resulting in a D-FACTS placement solution that fails to meet Corollary 3.3.2. Specifically, each isolated node in $G_{\overline{DF}}$ decreases the rank of the composite matrix by one according to (3.14). If the $G_{\overline{DF}}$ constructed in row 5 is a connected graph, Algorithm 2 terminates with the D-FACTS placement solution obtained; otherwise, Algorithm 2 updates G_{DF} and $G_{\overline{DF}}$, shown in rows 6–21, to make sure that both of them have no isolated nodes. For an isolated node in $G_{\overline{DF}}$, all edges connected to this node are contained in the MST (G_{DF}) due to their relatively low weights. If the weight of one edge connected to the isolated node is increased,

Algorithm 1: Max-rank Planning Algorithm for the Incomplete MTD

Input: The edge-weighted graph $G(L, E)$ of a power grid topology
Output: DF: set of D-FACTS lines; NDF: set of non-D-FACTS lines

Initialization: $E_{lp} = \emptyset$ // set of edges in a loop
NDF = find the MST in G // NDF candidates
$DF = E - NDF$
Generate a graph G_{DF} composed of DF lines and all nodes
while G_{DF} has loops
  Add all edges in the first loop to set E_{lp}
  Arrange edges in E_{lp} in ascending order of their weights
  for each edge ε in E_{lp} // start from the lowest-weight edge
    $\varepsilon.\omega = \varepsilon.\omega \times \lambda$ // decrease the positive weight ($\lambda < 1$)
    NDF = find the MST in weight-updated G
    $DF = E - NDF$
    Update G_{DF} using new DF lines
    if G_{DF} has no loops
      return DF, NDF
    else if the same loop E_{lp} still exists in G_{DF}
      $\varepsilon.\omega = \varepsilon.\omega \div \lambda$ //restore ε, try the next edge in loop
    else
      break //loop E_{lp} doesn't exist, move to the next loop
    end if
  end for
end while
return DF, NDF

the updated MST will exclude this edge, and this isolated node in $G_{\overline{DF}}$ will be removed consequently. The basic idea of Algorithm 2 is similar to that of Algorithm 1; however, the major difference is that Algorithm 1 updates the weights to eliminate the loops in G_{DF}, while Algorithm 2 updates the weights to find a connected $G_{\overline{DF}}$.

Algorithm 2: Max-rank Planning Algorithm for the Complete MTD

Input: The edge-weighted graph $G(L, E)$ of a power grid topology
Output: DF: set of D-FACTS lines; NDF: set of non-D-FACTS lines

Initialization: $L_{iso} = \emptyset$ // set of isolated nodes in non-D-FACTS graph
Negating the weights of all edges in G
DF = find the MST in G // DF candidates
$NDF = E - DF$
Generate a graph $G_{\overline{DF}}$ composed of NDF lines and all nodes
Add all isolated nodes in $G_{\overline{DF}}$ to L_{iso}
while $\|L_{iso}\|_0 \neq 0$
  for each node $l \in L_{iso}$
    Arrange edges connected to l in descending order of the weights
    for each edge ε connected to l // start from largest-weight edge
      $\varepsilon.\omega = \varepsilon.\omega \times \lambda$ // increase the negative weight (λ<1)
      DF = find the MST in weight-updated G
      $NDF = E - DF$
      Generate a graph $G_{\overline{DF}}$ composed of NDF lines and all nodes
      $L_{iso} = \emptyset$, and add all isolated nodes in $G_{\overline{DF}}$ to L_{iso}
      if $\|L_{iso}\|_0 = 0$
        return DF, NDF
      end if
      if $l \in L_{iso}$ // node l still exists in L_{iso}
        $\varepsilon.\omega = \varepsilon.\omega \div \lambda$ // restore ε, try the next edge
      end if
    end for
  end for
end while
return DF, NDF

The time complexity of Prim's algorithm is $O(ElogL)$ in a connected edge-weighted graph with L nodes and E edges. The time complexity of Algorithm 1 and Algorithm 2 is $O(KElogL)$, where K is the times of updating the edge weights. The value of K depends on the system topology and the absolute value of the median PLIS of each line. We will show the CPU time of proposed algorithms on medium- to large-scale power systems in the next

section.

A max-rank MTD is not attainable if the number of D-FACTS devices is less than the minimum number identified in equation (3.15). Since $r(\mathbf{M})$ is determined by the number of loops in $G_{\overline{DF}}$, i.e., $r(\mathbf{M}) = p - lp_{\overline{DF}}$ in equation (3.14), each D-FACTS device placed by Algorithm 1 breaks one loop in $G_{\overline{DF}}$ and thus increases $r(\mathbf{M})$ by one. Note that equation (3.14) holds regardless of the number of D-FACTS devices as long as G_{DF} is loopless. For example, if there are k D-FACTS devices less than $\min\{p_1\}$, $r(\mathbf{M})$ will decrease by k. Therefore, when the number of D-FACTS devices is less than the minimum number required by equation (3.15), the placement is equivalent to removing k D-FACTS devices from a placement solution obtained by Algorithm 1. In order to maintain the largest median PLIS in G_{DF}, the D-FACTS devices on k lines with the lowest absolute values of PLIS are chosen to be removed.

3.4 Graph-theory-based MTD Planning Algorithm

In this section, we study the metrics of MTD detecting effectiveness from the perspective of MTD planning. We design a novel detection effectiveness metric for MTD planning to reflect the detection effectiveness more accurately, i.e., attack detecting probability (ADP). Then, we highlight the drawbacks of the max-rank planning and further propose a novel graph-theory-based planning algorithm.

3.4.1 Analysis of MTD Detection Effectiveness

We define below protected and unprotected buses in MTDs in the noiseless condition to facilitate the presentation.

Definition 3.4.1. In an MTD, Bus i is a protected bus if the corresponding column of this bus in $\mathbf{H}_0$ is linear independent to $\mathbf{H}$, i.e., $h_i^0 \notin col(\mathbf{H})$; otherwise, it is an unprotected bus in the MTD, i.e., $h_i^0 \in col(\mathbf{H})$.

The protected and unprotected buses in this dissertation are defined from the perspective

of detection effectiveness. This is superior to the definition in reference[87], where protected buses are merely associated with D-FACTS devices. The drawback of the definition in reference[87] is that while D-FACTS devices are installed on lines connected to a bus, this bus can still be an unprotected bus under certain circumstances. We propose Lemmas 3.4.1 and 3.4.2 to illustrate the characteristics of protected and unprotected buses in MTDs against single-bus attacks. Note that a single-bus attack is the simplest FDI attack requiring the minimum effort from attackers who manipulate measurements only with respect to one bus. This dissertation focuses on analyzing the MTD detection effectiveness against such single-bus attacks. However, the theoretical results of this work can be extended to analyze multiple-bus attacks since a multiple-bus attack can be treated as multiple single-bus attacks launched at the same time.

Lemma 3.4.1. Any FDI attack on a protected bus is detectable, while any FDI attack on an unprotected bus is undetectable by an MTD.

The proof of Lemma 3.4.1 is apparent and thus omitted here. It should be noted that the number of protected buses determines the MTD detecting effectiveness. We further propose Lemma 3.4.2 to demonstrate the exact relation between the rank of the composite matrix and the number of protected buses.

Lemma 3.4.2. If the rank of the composite matrix in an MTD is $r([\mathbf{H}_0\ \mathbf{H}])$, there are at least $r([\mathbf{H}_0\ \mathbf{H}]) - (n-1)$ protected buses.

Proof: Since both $\mathbf{H}$ and $\mathbf{H}_0$ are of full column rank, i.e., $r(\mathbf{H}) = r(\mathbf{H}_0) = n-1$, there are at least $r([\mathbf{H}_0\ \mathbf{H}]) - n + 1$ columns in $\mathbf{H}_0$ that are linear independent to $\mathbf{H}$ in the MTD. Thus, there are at least $r([\mathbf{H}_0\ \mathbf{H}]) - n + 1$ protected buses in the MTD with a compose matrix rank $r([\mathbf{H}_0\ \mathbf{H}])$. ■

According to Lemma 3.4.2, the rank of the composite matrix merely reflects the minimum number of protected buses in an MTD. Therefore, all existing MTD operation and planning methods in the literature based on the rank maximization of the composite matrix[70;79;88] are equivalent to attaining a maximized lower bound on the number of protected buses, i.e.,

a lower bound on the MTD detection effectiveness. This rank as the only indicator of the MTD detection effectiveness is insufficient to analytically compare the detection effectiveness of different MTDs with the same lower bound. Hence, it is necessary to introduce an upper bound of the MTD detection effectiveness. This upper bound can be determined by the number of unprotected buses according to Definition 3.4.1. One can maximize this upper bound of MTD detection effectiveness by minimizing the number of unprotested buses via MTD planning. We propose the following lemma to show what constitutes an unprotected bus.

Lemma 3.4.3. 1) An end bus, 2) a non-D-FACTS bus, and 3) a bus fully covered by D-FACTS lines whose reactances are modified using a unity factor are unprotected buses.

Proof: Since there are three types of unprotected buses, this lemma is proved as follows: 1) End Bus i has only one transmission line i–j and its reactances before and after an MTD are x_{ij}^0 and x_{ij}, respectively. It is evident that $h_i^0 x_{ij}^0 = h_i x_{ij}$ holds. In this case, for any value of x_{ij}in different MTDs, h_i^0 is always linear dependent to $\mathbf{H}$ due to $h_i^0 = (x_{ij}/x_{ij}^0) \times h_i \in col(\mathbf{H})$. Thus, an end bus is an unprotected bus in any MTDs. 2) If Bus i is a non-D-FACTS bus, $h_i = h_i^0$ holds. It is apparent that $h_i^0 \in col(\mathbf{H})$ holds. Thus, a non-D-FACTS bus is an unprotected bus. 3) Suppose all lines connected to Bus i are D-FACTS lines. An MTD modifies the reactances of these lines with the same factor k, i.e., $x_j = kx_j^0, j \in S_i$ where S_i is the set of lines connected to Bus i. It is evident that $h_i^0 = kh_i$ holds. Thus, this bus becomes an unprotected bus under the MTD. ■

It is worthwhile to mention that end buses and non-D-FACTS buses can be identified based on the power system topology or MTD planning, while a bus fully covered by D-FACTS lines whose reactances are modified using a unity factor involves both MTD planning and MTD operation. Additionally, Lemma 3.4.3 points out a limitation of the MTD that a single-bus FDI attack on any end bus is undetectable regardless of the D-FACTS setpoints and planning solutions. Therefore, leaving an end bus as a non-D-FACTS bus can reduce the number of D-FACTS devices without affecting the detection effectiveness. Installing protected sensors to secure measurements related to the end buses can thwart FDI attacks

against end buses. However, securing measurement sensors to detect and prevent FDI attacks belongs to another defense algorithm well studied in the literature[59–61;75;76], which is beyond the scope of this dissertation.

The difference between Lemmas 3.4.1–3.4.3 and Remark 3 in reference[79] are summarized here. Remark 3 claims for an MTD with $\mathbf{H}$, $r([\mathbf{H}_d\ \mathbf{H}]) = n - 1 + |S_d|$ holds, where S_d is a set of columns in $\mathbf{H}_0$ independent to $\mathbf{H}$[79]. However, the following important issues are not discussed in reference[79]. Firstly, reference[79] only provides a highly abstracted set of states without concrete instances. In this work, we instantiate the columns in S_d as protected buses from the standpoint of power systems. Secondly, $r([\mathbf{H}_0\ \mathbf{H}])$ is the MTD detection metric, but the relationship between $r([\mathbf{H}_0\ \mathbf{H}])$ and $r([\mathbf{H}_d\ \mathbf{H}])$is not directly presented in reference[79]. In Lemma 3.4.1, we clarify the relation between FDI attacks and unprotected/protected buses. In Lemma 3.4.2, we clarify the relation between $r([\mathbf{H}_0\ \mathbf{H}])$ and the number of protected buses. Thirdly, reference[79] doesn't discuss which buses are included in S_d or not in an MTD or under an MTD planning solution. In Lemma 3.4.3, we point out three types of unprotected buses.

Here, we utilize the ADP of an MTD to measure its detection effectiveness against FDI attacks accurately, which is widely used as attack detection evaluation metric in the literature[72;79;85].

Definition 3.4.2. The ADP of an MTD is defined as the ratio of the number of the FDI attacks detected by this MTD to the total number of FDI attacks.

In a specific MTD, a bus is either a protected bus or an unprotected bus in accordance with Definition 3.4.1. Without loss of generality, we assume targeted buses in single-bus FDI attacks are uniformly distributed. The reference bus is not considered in calculating the ADP since this bus cannot be a target bus of FDI attacks. Thus, the ADP of an MTD against single-bus FDI attacks is equal to the ratio of the number of protected buses (n_p) to the total number of buses excluding the reference bus, i.e., $ADP_{MTD} = n_p/(n-1)$, according to Lemmas 3.4.1 and 3.4.2. Since n_p in different MTDs varies as their D-FACTS setpoints change, we propose Theorem 3.4.1 to demonstrate the ADP range of an MTD. Note that

Lemmas 3.4.1 and 3.4.2 are used to derive the lower bound of an MTD in Theorem 3.4.1, and Lemma 3.4.3 to derive the upper bound.

Theorem 3.4.1. For an MTD with $r([\mathbf{H}_0\ \mathbf{H}])$, n_2 non-D-FACTS buses and n_{1e} end buses belonging to the D-FACTS buses, its ADP against single-bus FDI attacks satisfies:

$$\frac{r([\mathbf{H}_0\ \mathbf{H}]) - n + 1}{n-1} \leq ADP_{MTD} \leq 1 - \frac{n_2 + n_{1e}}{n-1} \tag{3.16}$$

Proof: In an MTD with $r([\mathbf{H}_0\ \mathbf{H}])$, $n_p \geq r([\mathbf{H}_0\ \mathbf{H}]) - (n-1)$ holds according to Lemma 3.4.2. In addition to $n_2 + n_{1e}$ unprotected buses identified by Lemma 3.4.3, D-FACTS setpoints can convert some protected buses to unprotected buses. Thus, $n_p \leq (n-1) - (n_2 + n_{1e})$ holds. Therefore, the ADP of the MTD satisfies equation (3.16). ■

It is worth mentioning that the ADP of a complete MTD[70;72;79] is a particular case of Theorem 3.4.1, where the ADP upper and lower bounds are both equal to 1. For the upper bound, there is neither an end bus nor a non-D-FACTS bus in the complete MTDs, i.e., $n_{1e} = 0$ and $n_2 = 0$, since the D-FACTS lines form a spanning tree[70]. For the lower bound, $r([\mathbf{H}_0\ \mathbf{H}]) = 2(n-1)$ in the complete MTD.

In addition, we define a novel metric for measuring the detection effectiveness of an MTD planning solution.

Definition 3.4.3. The ADP of an MTD planning solution is defined as the average ADP of MTDs under this planning solution.

The ADP of an MTD planning solution is not a fixed value since the number of protected buses varies depending on the D-FACTS setpoints in each MTD under this planning solution. Instead, we can calculate an ADP range of an MTD planning solution, as shown in the following theorem, to represent its detective effectiveness.

Theorem 3.4.2. For an MTD planning solution with a fixed rank of the composite matrix equal to $r([\mathbf{H}_0\ \mathbf{H}])$, n_2 non-D-FACTS buses, and n_{1e} end buses belonging to the D-FACTS

buses, its ADP against single-bus FDI attacks satisfies:

$$\frac{r([\mathbf{H}_0\ \mathbf{H}]) - n + 1}{n-1} \leq ADP_{Planning} \leq 1 - \frac{n_2 + n_{1e}}{n-1} \tag{3.17}$$

It is trivial to prove Theorem 3.4.2 using Theorem 3.4.1. Note that this dissertation focuses on the MTD planning with a fixed rank of the composite matrix rather than a varying rank relevant to the D-FACTS setpoints. The selection of D-FACTS setpoints falls into the MTD operational issue and is therefore out of the scope of this dissertation. The MTD planning with a fixed rank of the composite matrix is preferable in the power system operation. This is because such an MTD planning can provide the system operator with more freedom for dispatching D-FACTS setpoints to meet system economic and reliability criteria while assuring the MTD detection effectiveness.

To elucidate Theorem 3.4.2, we demonstrate the relation between protected and unprotected buses under an MTD planning solution in Figure 3.2. The rank of the composite matrix solely determines the minimum number of protected buses, i.e., the ADP lower bound. The system topology and MTD planning jointly decide the number of unprotected buses, i.e., the ADP upper bound. Under this MTD planning solution, the boundary between the protected and the unprotected buses (dashed line in Figure 3.2) moves between the two shadow areas subject to the specific D-FACTS setpoints in an MTD.

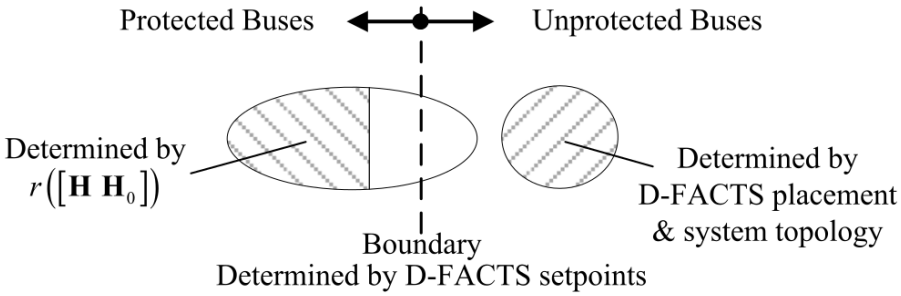


Figure 3.2: *Relation of protected and unprotected buses in an MTD planning.*

One can use Theorem 3.4.2 to analyse the MTD detection effectiveness of any MTD planning solutions with a loopless D-FACTS graph since the rank of the composite matrix in these planning is fixed regardless of D-FACTS setpoints[70]. For example, the max-rank MTD

planning[70] achieves the maximum rank of the composite matrix, i.e., $r([\mathbf{H}_0\ \mathbf{H}])$. This indicates any MTD under the max-rank planning has the maximum ADP lower bound. Besides, $G_{\overline{DF}}$ forms a spanning tree in the max-rank planning[70], leading to $n_{1e} = 0$. We illustrate the ADP range of incomplete MTDs constructed under the max-rank MTD planning in the following corollary.

Corollary 3.4.1. The ADP of the max-rank MTD planning[70] against single-bus FDI attacks satisfies:

$$\frac{p-(n-1)}{n-1} \leq ADP_{Planning} \leq 1 - \frac{n_2}{n-1} \tag{3.18}$$

Corollary 3.4.1 shows the merits and drawbacks of the max-rank MTD planning. Compared with (3.17), the max-rank MTD planning increases both the ADP lower and upper bounds. Nevertheless, the existence of non-D-FACTS buses reduces the ADP upper bound significantly.

We further propose Corollary 3.4.2 to identify a special class of MTD planning solutions with a fixed ADP (i.e., the ADP lower bound equates to their upper bound) regardless of the setpoints of D-FACTS devices. Here, we define a reduced D-FACTS graph as a graph composed of D-FACTS lines and D-FACTS buses.

Corollary 3.4.2. If the reduced D-FACTS graph is a loopless and connected graph and $G_{\overline{DF}}$ is a spanning tree, the ADP of this MTD planning solution is a fixed value regardless of D-FACTS setpoints, i.e., $ADP_{Planning} = (p-n+1)/(n-1)$.

Proof: According to the theorem of Euler's formula for a disconnected graph[100], in any planar graph with n vertices, p edges, f faces, and t components, the following equality holds: $n + f - p = 1 + t$. Since the number of faces equals to the sum of the number of interfaces (loops in the graph) and one external face, *i.e.*, $f = lp+1$, we have $n-t = p-lp$. Thus, $n_1 - t_1 = p_1 - lp_1$ holds in the reduced D-FACTS graph, where t_1 and lp_1 are the number of disconnected components and the number of loops in the reduce D-FACTS graph, respectively. If the reduced D-FACTS graph is connected and loopless, i.e., $t_1 = 1$ and $lp_1 = 0$, $n_1 - 1 = p_1$ holds. Since the sum of the number of D-FACTS and non-D-FACTS

buses equals the number of buses in the system, i.e., $n_1+n_2=n$, we have $n_1-1=(n-1)-n_2$. Therefore, $p_1=(n-1)-n_2$ holds. As $G_{\overline{DF}}$ is a spanning tree with $(n-1)$lines, the number of D-FACTS lines is $p_1=p-(n-1)$. Then, the ADP lower bound in (3.18) equals the ADP upper bound, i.e., $p-(n-1)=(n-1)-n_2$. Therefore, this class of MTD planning solutions has a fixed ADP, i.e., $ADP=(p-n+1)/(n-1)$. ■

Although the MTD under the MTD planning identified in Corollary 3.4.2 has a fixed ADP, one ought to avoid these MTD planning solutions since it has the lowest ADP among all max-rank planning solutions.

3.4.2 Graph-theory-based MTD Planning Algorithm

We propose a graph-theory-based planning method to simultaneously ensure the ADP lower bound and increase the ADP upper bound. The proposed method is composed of Algorithm 1 in the previous section and Algorithm 3. In these algorithms, we calculate and assign PLIS to each line as its weight in G, as PLIS is an indicator to determine the most appropriate D-FACTS locations to minimize system losses[96]. A line with a larger absolute PLIS value indicates installing a D-FACTS device on this line can reduce more system losses.

In Algorithm 1, we aim to maximize the rank of the composite matrix by ensuring both G_{DF} and $G_{\overline{DF}}$ loopless. However, there may exist non-D-FACTS buses in the solution obtained from Algorithm 1.

In Algorithm 3, we remain the maximal rank of the composite matrix and use extra D-FACTS lines to cover all unprotected buses, excluding end-buses. Algorithm 3 takes the result in Algorithm 1 as its input. In Algorithm 3, the following four crucial objectives are holistically accounted for: 1) reducing the investment cost of extra D-FACTS devices; 2) retaining the maximum rank of the composite matrix; 3) achieving better operational economics in the system; and 4) elevating the ADP upper bound. We establish the following three rules, each corresponding to one of the first three objectives.

Planning Rule 1. Placing a D-FACTS device only on the line whose two nodes are neither D-FACTS bus nor end bus. Thus, we can use one D-FACTS device to simultaneously

eliminate two non-D-FACTS buses for reducing the number of extra D-FACTS devices.

Planning Rule 2. Installing an extra D-FACTS device on a line if the updated D-FACTS graph containing this line remains loopless. We design Rule 2 to remain the maximum rank of the composite matrix in MTDs by keeping both $G_{\overline{DF}}$ and G_{DF} loopless.

Planning Rule 3. Sorting non-D-FACTS lines in descending order of their PLIS values and sequentially deciding whether to install a D-FACTS device on the non-D-FACTS lines. We design Rule 3 to select lines with large PLIS values.

It identifies all end buses, D-FACTS buses, and non-D-FACTS lines in the system during the initialization. In the first part (as shown in Rows 1–14) of Algorithm 3, we sort non-D-FACTS lines in descending order of their weights according to Rule 3 and iteratively check each non-D-FACTS line. We place a D-FACTS device on this line if Rules 1 and 2 can be simultaneously satisfied; otherwise, we skip this line and consider the next. After the first step, non-D-FACTS buses excluding end buses might still exist, i.e., $n_{2/e} \neq 0$, if their neighbor buses are either end buses or D-FACTS buses. Then, we identify all these nodes in set V_{NDF}. In the second part of Algorithm 3 (Rows 16–26), we convert all buses in V_{NDF} to D-FACTS buses following Rules 2 and 3. Algorithm 3 stops after eliminating all non-D-FACTS buses, excluding end buses.

The ADP range of the proposed planning solution is shown in the following corollary.

Corollary 3.4.3. The ADP of the graph-theory-based planning method against single-bus FDI attacks satisfies:

$$\frac{p-(n-1)}{n-1} \leq ADP_{Planning} \leq 1-\frac{n_e}{n-1} \tag{3.19}$$

The ADP lower bound of the proposed MTD planning is no less than that in the max-rank planning since it remains the maximum rank of the composite matrix. Since all end buses are non-D-FACTS buses in the max-rank planning, i.e., $n_2 = n_{2/e} + n_e$ holds. Compared with the ADP upper bound in (3.18), the proposed planning increases the ADP upper bound by $n_{2/e}/(n-1)$ though eliminating $n_{2/e}$ non-D-FACTS buses, excluding the end buses in the max-rank MTD planning.

Algorithm 3: Covering all unprotected buses, excluding end-buses

Input: The edge-weighted graph $G(V, E)$ of a power grid topology E^0_{DF}: set of D-FACTS lines in Algorithm 1

Output: E_{DF}: set of D-FACTS lines in the proposed MTD planning

Initialization: $E_{DF} = E^0_{DF}$
V_{end} = find all end nodes in $G(V, E)$
V_{DF} = find all nodes of edges in E_{DF} // set of D-FACTS buses
$E_{NDF} = E - E_{DF}$ // set of non-D-FACTS lines
Arrange edges in E_{NDF} in descending order of their weights // Rule 3
for each edge ε in E_{NDF} // start from the largest-weight edge
  if two nodes of ε are neither in V_{end} nor in V_{DF} // Rule 1
    Generate a graph G_1 composed of E_{DF} and ε : $G_1(V, E_{DF} + \varepsilon)$
    if G_1 has no loops // Rule 2
      Remove ε from E_{NDF}, and add ε to E_{DF}
      Add two nodes of ε to V_{DF}
    end if
  end if
end for
$V_{NDF} = V - V_{DF} - V_{end}$ // non-D-FACTS nodes failed to meet Rules1 & 2
for each node v in V_{NDF}
  Generate an empty set E_1, and add all edges connected to v to E_1
  Arrange edges in E_1 in descending order of their weights // Rule 3
  for edge ε in E_1 // start from the largest-weight edge
    Generate a graph G_2 composed of E_{DF} and ε : $G_2(V, E_{DF} + \varepsilon)$
    if G_2 has no loops // Rule 2
      Remove ε from E_{NDF}, and add ε to E_{DF}
      break // deal with the next node in V_{NDF}
    end if
  end for
end for
return E_{DF}

Based on Theorem 3.4.2, we derive the ADP range of different MTD planning methods in Corollaries 3.4.1, 3.4.2, and 3.4.3. Specifically, we present the ADP range in max-rank planning methods in Corollary 3.4.1, and derive the ADP range in the proposed planning in Corollary 3.4.3. In Corollary 3.4.2, we identify a class of planning methods that has a fixed ADP.

3.5 Experiment Results

The MTD planning algorithms are implemented using the Java programming language. The PLIS calculation, state estimation (SE), bad data detection (BDD), and FDI attacks are all programmed in MATLAB. The algorithms are performed on a laptop with Intel Core i5 processor CPU 2.70 GHz dual-core with 8 GB RAM. Data of all power systems used in this section are obtained from the MATPOWER package[101].

3.5.1 Numerical Results in Max-rank Planning Algorithms

To evaluate the effectiveness of the proposed MTD planning approaches, we perform numerical tests on a 6-bus system for the validation of a complete MTD, as well as on the IEEE 14-bus and IEEE 118-bus systems for incomplete MTDs. We examine the MTD detection effectiveness on FDI attacks in these systems.

MTD Planning Solutions

Since the PLIS calculated in (3.3) is with respect to the time-variant system operating point, we take the absolute value of the median PLIS of each line over a 24-hour scheduling horizon as its weight. The hourly system load over a 24-hour period can be found at http://motor.ece.iit.edu/data. The nodal load distribution factor is assumed to be fixed over time.

The PLIS of each line and MTD planning solution in the 6-bus system are shown in Figures 3.3 and 3.4, respectively. For each transmission line, the band inside the box, the

lower and upper ends of the whisker, displayed in Figure 3.3, are the median, minimum, and maximum PLIS over 24 hours, respectively. Lines 2, 3, 5, 8 and 9, which have the largest absolute values of the median PLIS, are the five (identified by Corollary 3.3.2) most suitable lines to install D-FACTS devices. The G_{DF} (red dotted lines in Figure 3.4) is constructed by five D-FACTS lines forming a spanning tree of the 6-bus system. The $G_{\overline{DF}}$ is a connected graph depicted by the solid lines in Figure 3.4. The MTD planning ensures that both the G_{DF} and the $G_{\overline{DF}}$ satisfy Corollary 3.3.2. Therefore, any MTD under this MTD planning is a complete MTD able to detect any FDI attack.

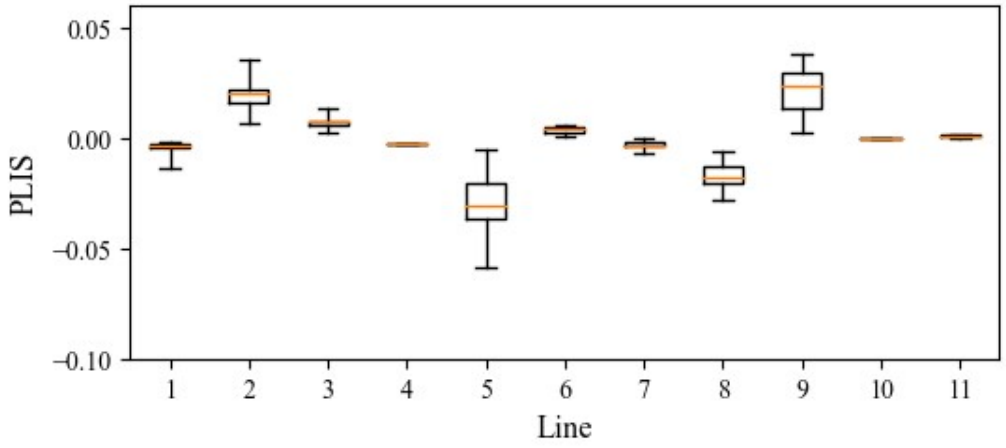


Figure 3.3: *The PLIS of each line in the 6-bus system.*

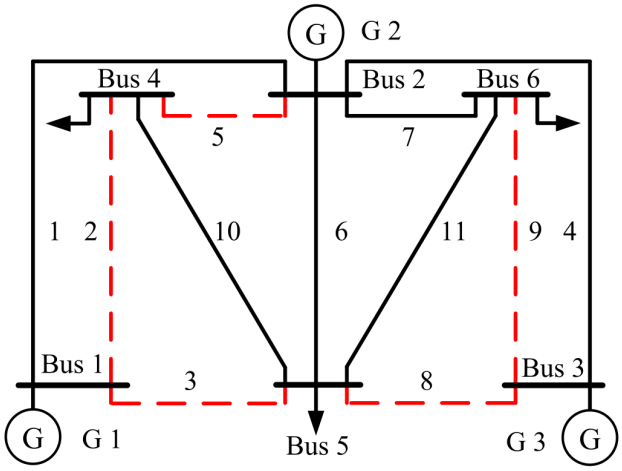


Figure 3.4: *Max-rank planning of the 6-bus system.*

The PLIS of each line and MTD planning solution on the IEEE 14-bus system are shown

in Figures 3.5 and 3.6, respectively. The D-FACTS are installed on most of the lines with large absolute values of PLIS, because $G_{\overline{DF}}$ is the MST of the IEEE 14-bus system in Algorithm 1. Figure 3.5 shows that the median PLIS of Lines 1–7 are comparatively larger than that of others in the IEEE 14-bus system. As shown in Figure 3.6, Lines 1, 3 and 4 are selected by Algorithm 1 to install D-FACTS devices, while Lines 2, 5, 6, and 7 are not chosen to avoid any loop. Instead, Algorithm 1 chooses Lines 8, 10, 12, and 13 to maintain both G_{DF} and $G_{\overline{DF}}$ loopless. Furthermore, the MTD planning solution has the largest absolute values of the median PLIS out of all feasible solutions according to Corollary 3.3.1.

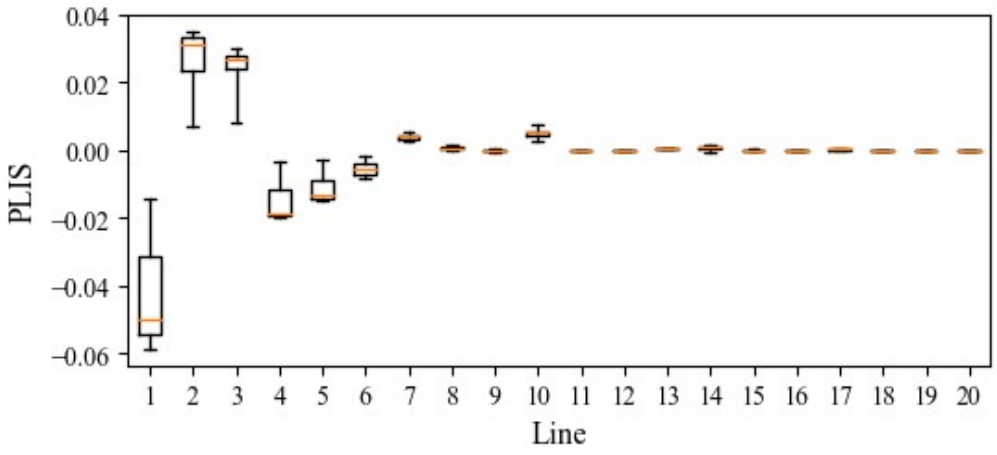


Figure 3.5: *The PLIS of each line in the IEEE 14-bus system.*

For the IEEE 118-bus system[101], D-FACTS devices are installed on 62 lines indexed by L_{DF}={2, 3, 4, 5, 12, 21, 23, 31, 36, 38, 39, 41, 42, 43, 47, 50, 51, 52, 56, 61, 63, 67, 69, 70, 73, 76, 83, 84, 86, 87, 91, 93, 94, 95, 99, 100, 101, 102, 103, 104, 106, 110, 112, 114, 119, 120, 126, 131, 133, 135, 139, 140, 141, 144, 148, 152, 154, 157, 160, 161, 163, 167}. Both G_{DF} and $G_{\overline{DF}}$ are loopless to guarantee a max-rank incomplete MTD. As a result, 34.6% of transmission lines on the IEEE 118-bus system are equipped with D-FACTS devices.

To demonstrate the computational efficiency of the proposed MTD planning algorithms, we test them in other medium- to large-scale power systems[101]. Table 3.2 illustrates the CPU time of the proposed algorithm in each system. The proposed MTD planning algorithms are used at the planning stage of D-FACTS devices, which in general does not have stringent time requirements. Due to the high efficiency of the proposed algorithms as shown

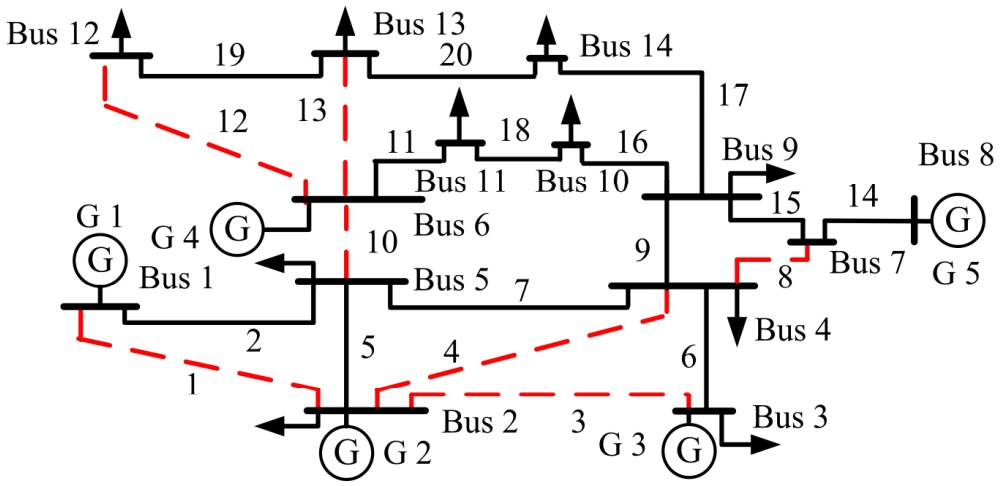


Figure 3.6: *Max-rank planning of the IEEE 14-bus system.*

in Table 3.2, they can be applied in the real-time RMTD to maximize the MTD effectiveness without running the MTD-based ACOPF model proposed in the next Chapter. This will be investigated in our future work.

Table 3.2: *CPU time of proposed max-rank planning algorithms*

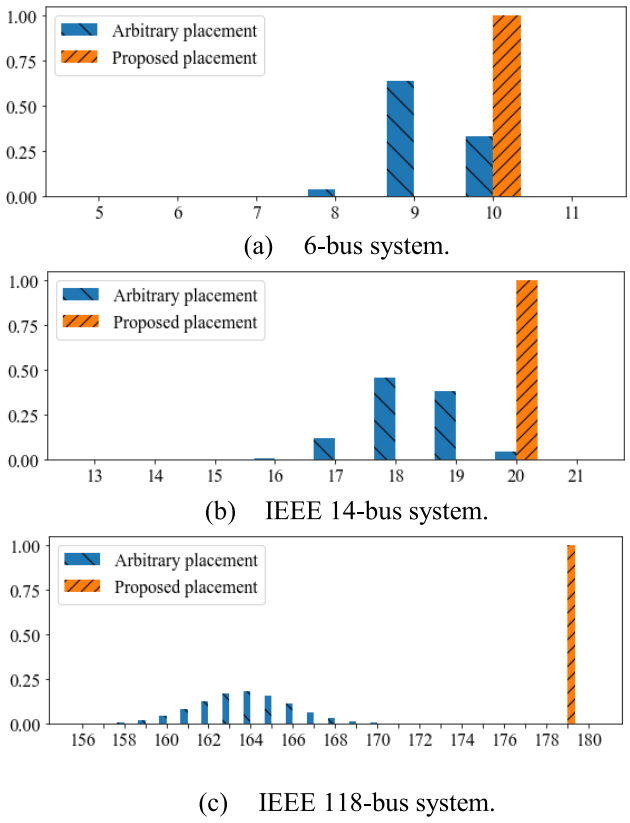
Power system	CPU time (ms)
IEEE 118-bus system	18
ACTIVSg 200-bus system[101]	20
ACTIVSg 500-bus system[101]	26
2746 bus system[101]	276

Comparison of the Composite Matrix Rank

Despite the setpoints of D-FACTS devices, any non-zero perturbation under the proposed MTD planning is a max-rank MTD. This allows for the use of the simplest MTD (i.e., RMTD) to evaluate the efficacy of the proposed MTD planning. In RMTD, the incremental reactance of the D-FACTS line is randomly generated satisfying constraint (15j). In this section, we set $\eta = 0.2$ and $\tau = 0.05$, consistent with the settings in reference[79].

In each of the above three systems, we compare the rank of the composite matrix under the proposed MTD planning with that under an arbitrary planning. The arbitrary planning

randomly selects a subset of lines to install the D-FACTS devices by using the minimum number of D-FACTS devices identified in (3.15). We implement the RMTD for 10,000 times under either of the proposed and arbitrary planning.



(a) 6-bus system.

(b) IEEE 14-bus system.

(c) IEEE 118-bus system.

Figure 3.7: *The histogram on the rank of the composite matrix.*

A histogram comparing the rank of the composite matrix over the 10,000 RMTDs is illustrated in Figure 3.7. According to Lemma 3.3.1, the maximum rank of the composite matrix in the three systems is 10, 20 and 179, respectively. In Figure 3.7, the proposed MTD planning, as expected, always leads to the max-rank MTD, while it is not the case under the arbitrary planning. In addition, the rank of the composite matrix deviates further from the corresponding maximum rank as the system size increases. The results demonstrate that the proposed MTD planning approach can guarantee the max-rank MTD and is thus much superior to the arbitrary planning especially in large systems.

MTD Detection Effectiveness for FDI Attacks

Simulations are carried out to test the RMTD effectiveness against stealthy FDI attacks under the proposed MTD planning in a noisy environment in both DC-SE and AC-SE. Without loss of generality, we consider a single time period in MTD, in which the system loads are assumed to be constant. Regarding DC-FDI attacks $\mathbf{a} = \mathbf{H}_0\Delta\theta$, we randomly generate attack vector $\Delta\theta$ with a fixed number of attack buses, i.e., $||\Delta\theta||_0 = q, q = 1, 2, ...5$. Then, we inject $\mathbf{a} = \mathbf{H}_0\Delta\theta$ into the real measurement vector. For each q, we simulate 1,000 distinct FDI attacks. Five thousand FDI attacks are constructed as an attack pool for each system to calculate the ADP of an MTD, which is defined as the true positive rate (sensitivity) of detecting FDI attacks.

Furthermore, one hundred different RMTDs are constructed as a defense pool under the proposed MTD planning. For each system, all FDI attacks in this attack pool are sequentially implemented on each of the RMTDs in the defense pool. We highlight the average ADP to reflect the effectiveness of the RMTD in detecting FDI attacks. In a noisy condition, the measurement noise is assumed to be Gaussian distributed with zero mean and the standard deviation as 3% of the actual measurement. BDD is used to detect the FDI attacks, and its threshold is set at a false positive rate of 0.05.

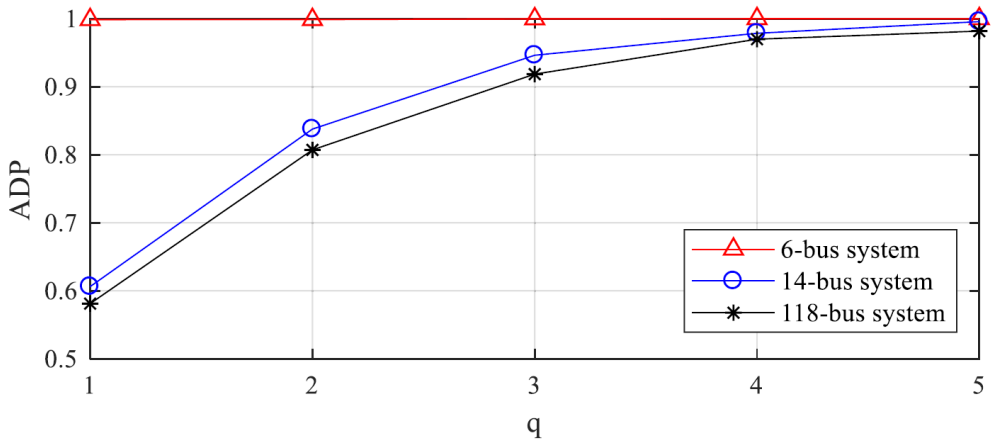


Figure 3.8: *ADP under the proposed planning with measurement noises in DC-SE.*

The average ADP of RMTDs under the proposed MTD planning in DC noisy conditions

is shown in Figure 3.8. Since the MTD planning solution in the 6-bus system (see Figure 3.4) ensures a complete MTD, the RMTD is capable of detecting any FDI attack regardless of the number and location of attack buses, leading to an ADP of 100%. The detection effectiveness is highly desirable for the power system operator as a cyber-attack defender. Nevertheless, the requirements on a complete MTD discussed in Section 3.3.1 can no longer be met on the IEEE 14-bus and the IEEE 118-bus systems. In those two systems, nearly 60% of FDI attacks with $q = 1$ can be detected by the RMTD as illustrated in Figure 3.8. The ADP increases with a growing number of attack buses and becomes close to 1 when $q \geq 4$. The results show that FDI attacks on more buses are more likely to be detected by the RMTD under the proposed MTD planning.

Further, we test the RMTD effectiveness under the proposed MTD planning in the AC power flow model. To facilitate the comparison between AC and DC power flow models, we generate AC FDI attacks by using the same attack vector $\Delta\theta$ in the DC attack pool. Different from DC-FDI attacks, attackers need to know system states (i.e., voltage magnitude and angle) to construct AC-FDI attacks. It is reasonable to assume that attackers know the actual system states after the MTD through data infiltration and use the original line parameters before the MTD to calculate the manipulated measurements, i.e., $\mathbf{z}_a = h(\mathbf{V}, \theta + \Delta\theta, \mathbf{x}_0)$. The ADP is calculated by sequentially launching FDI attacks against AC-SE, when the system operator uses the same D-FACTS setpoints in the defense pool generated previously. Regarding the sensor deployment, we adopt a 2.5 redundant factor (i.e., a ratio between the numbers of measurements and system states) to guarantee the observability in SE. Sensors for voltage magnitude, active and reactive power flow, active and reactive nodal power injection are randomly deployed in each system. The average ADP of RMTDs under the proposed MTD planning in AC noisy conditions is demonstrated in Figure 3.9. When comparing Figures 3.9 and 3.8, it is seen the MTD detection effectiveness under the proposed MTD planning in AC-SE is generally consistent with that in DC-SE. The results in Figure 3.9 demonstrates the effectiveness of the proposed MTD planning in AC-SE.

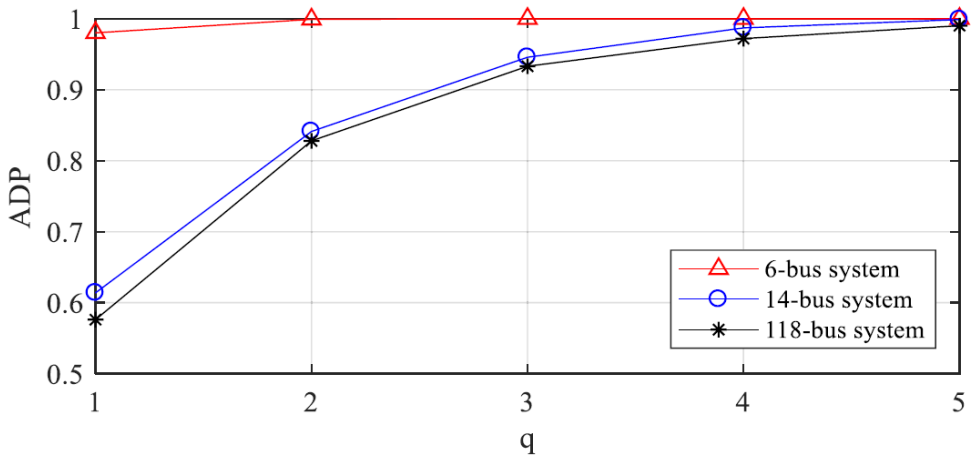


Figure 3.9: *ADP under the proposed planning with measurement noises in AC-SE.*

3.5.2 Numerical Results in Graph-theory-based MTD Planning Algorithm

We perform the proposed MTD planning method on the IEEE 14-bus system, the IEEE 118-bus system, the ACTIVSg 500-bus system, the 2746-bus system, and the 3012-bus system [101]. The proposed MTD planning method is programmed in the Java programming language. The MTD, FDI attacks, and SE are implemented using MATLAB.

Graph-theory-based MTD Planning Solutions

The proposed MTD planning method for the IEEE 14-bus system is shown in Figure 3.10. It is seen in the system there are no non-D-FACTS buses except an end bus, i.e., Bus 8. The minimum number of non-D-FACTS buses contributes to an improvement in the ADP upper bound. Furthermore, both G_{DF} and $G_{\overline{DF}}$ are loopless such that any MTDs under this planning solution remain the maximum rank of the composite matrix. Compared with the max-rank planning [70], the proposed planning installs only two extra D-FACTS devices on Lines 17 and 18 to transform four non-D-FACTS buses (i.e., Buses 9, 10, 11, and 14) in the max-rank planning to D-FACTS buses according to Rule 1. Although Lines 2, 5, and 6 have large PLIS values, placing D-FACTS devices on any of them would form a loop and, in turn, reduce the rank of the composite matrix. Therefore, no D-FACTS device is installed

on these three lines.

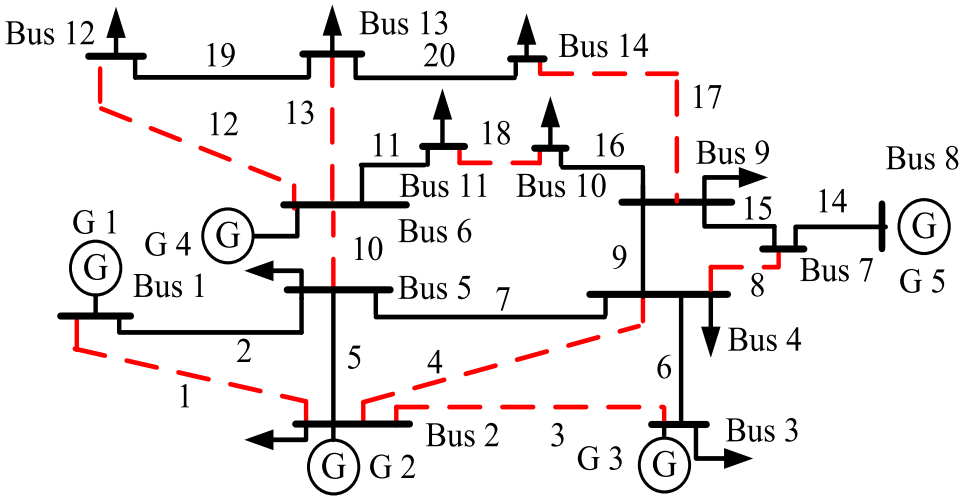


Figure 3.10: *The proposed MTD planning solution of the IEEE 14-bus system.*

For the IEEE 118-bus system, D-FACTS devices in the proposed MTD planning are installed on 97 branches indexed by L_{DF}={2, 3, 4, 5, 7, 8, 12, 13, 15, 18, 19, 21, 22, 23, 25, 27, 28, 29, 31, 35, 36, 38, 39, 41, 42, 43, 47, 48, 49, 50, 51, 52, 56, 59, 61, 63, 66, 67, 68, 69, 70, 72, 73, 76, 79, 81, 83, 84, 86, 87, 91, 93, 94, 95, 96, 99, 100, 101, 102, 103, 104, 106, 107, 110, 112, 114, 118, 119, 120, 122, 124, 126, 127, 128, 131, 132, 133, 135, 136, 139, 140, 141, 144, 148, 151, 152, 154, 155, 157, 160, 161, 163, 166, 167, 171, 175, 178}. In the max-rank planning, there are total 46 non-D-FACTS buses, including seven end buses. In the proposed MTD planning, additional 35 D-FACTS devices are used to cover these 39 non-D-FACTS buses and all seven end buses are left as non-D-FACTS buses. No loops exist in either G_{DF} or $G_{\overline{DF}}$, which guarantees the maximum rank of the composite matrix.

We test the computational efficiency of the proposed MTD planning in medium- to large-scale power systems. The CPU time of the proposed algorithm in each system is shown in Table 3.3. The results here show the proposed MTD planning is computationally efficient even on large-scale power systems.

Table 3.3: *CPU time of proposed algorithm on different systems*

Power system	118-bus system	500-bus system	2746-bus system	3012-bus system
CPU time (ms)	12	69	515	402

Comparison of Planning Methods

We compare the proposed planning method with four other planning methods on a medium-scale system and a large-scale system in Table 3.4. There are seven end buses and 179 lines in the IEEE 118-bus system, and 550 end buses and 3566 lines in the 3012-bus system. Thus, the maximum rank of the composite matrix in the IEEE 118-bus and the 3012-bus systems are 179 and 3566, respectively. Furthermore, the ADP ranges in Table 3.4 are derived based on Theorem 3.4.1 and Corollaries 3.4.1–3.4.3.

As seen, the proposed planning method has zero non-D-FACTS buses, excluding end buses in both systems. Thus, the ADP upper bound of the proposed planning method is much higher than that of the max-rank planning. In addition, the proposed planning has the same ADP lower bound as the max-rank planning. Compared with the arbitrary planning, the proposed planning has a higher ADP range and a larger PLIS sum, resulting in better performances in detecting FDI attacks and minimizing system losses. The proposed planning has the same ADP range as that of the full planning and spanning-tree planning. However, the proposed planning uses a much smaller number of D-FACTS devices than full planning and spanning-tree planning, especially in the 3012-bus system. Even though the price of a D-FACTS device is lower than that of a conventional FACTS device, the cost of D-FACTS devices is still not negligible. Therefore, the proposed planning can significantly reduce the cost of deploying D-FACTS devices, especially in such a large-scale system with thousands of lines.

Comparison of ADP in both DC- and AC-SE

In this section, we compare the ADP of four MTD planning methods in both DC- and AC-SE in IEEE 118-bus system, including the proposed planning, the arbitrary planning, the full

Table 3.4: *Planning method comparison in medium- and large-scale systems*

Power system	Planning method	PLIS sum	p_1	$r(\mathbf{M})$	$n_{2/e}$	ADP range (%)
118-bus system	Proposed	8.62	97	179	0	[53.0, 94.0]
	Max-rank	7.89	62	179	39	[53.0, 60.7]
	Full	10.27	179	179	0	[53.0, 94.0]
	Arbitrary	5.4	97	173	11	[47.9, 84.6]
	Spanning-tree	2.38	117	179	0	[53.0, 94.0]
3012-bus system	Proposed	207.5	1844	3566	0	[18.4, 81.7]
	Max-rank	185.5	555	3566	1695	[18.4, 25.4]
	Full	227.2	3566	3566	0	[18.4, 81.7]
	Arbitrary	53.2	1844	3555	438	[18.1, 67.2]
	Spanning-tree	41.7	3011	3566	0	[18.4, 81.7]

planning, and the max-rank planning.

To simulate FDI attacks, we randomly generate a voltage angle increment vector $\Delta\theta$ with a fixed number of attack buses, i.e.,$||\Delta\theta||_0 = q, q = 1, 2, ...5$. For each q, we simulate 1,000 distinct voltage angle increment vectors. Using these vectors, we construct both a DC-FDI attack pool and an AC-FDI attack pool, each of which has five thousand FDI attacks. For the DC-FDI attacks, we inject $a = H_0\Delta\theta$ in the real measurement vector. For the AC-FDI attacks, we assume that attackers know the actual system states after the MTD through data infiltration and use the original line parameters before the MTD, i.e., $\mathbf{x}_0$ to calculate the manipulated measurements, i.e., $z_a = h(V, \theta + \Delta\theta, x_0)$. RMTD operation method, the simplest MTD operation method, is adopted in all four planning methods to determine the setpoints of D-FACTS devices. In RMTD, we set $\eta = 0.2$ and $\tau = 0.05$, consistent with the settings in reference[79]. For each planning method, one hundred different RMTDs are constructed as a defense pool. Note that we generate 100 different D-FACTS random placement solutions for the arbitrary planning method.

For each planning method, all attacks generated in the attack pool are sequentially launched on each MTD in the corresponding defense pool. Then, the average ADP is calculated as an indicator of the detection effectiveness of this planning method. In the noisy condition, the measurement noise is assumed to be Gaussian distributed with zero mean and

the standard deviation as 3% of the actual measurement. The SE and BDD are used to detect the FDI attacks, the threshold of which is set to have a 0.5% false-positive rate.

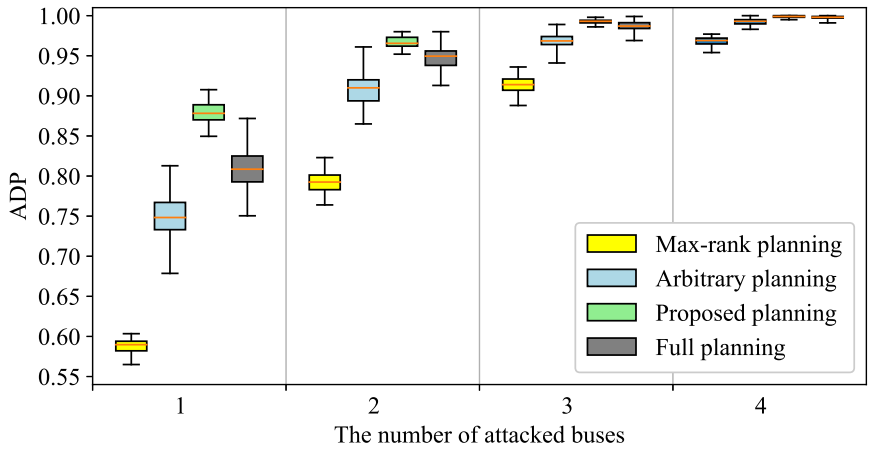


Figure 3.11: *The ADP of the four MTD planning algorithms versus q in DC-SE.*

The ADP of the four planning methods against FDI attacks with q varying from 1 to 4 in the DC noisy condition is demonstrated in Figure 3.11. As seen, the max-rank planning method has the lowest ADP, consistent with the range derived in Table 3.4. This is because there are 39 non-D-FACTS buses excluding end buses, which seriously reduces the MTD detection effectiveness. In Figure 3.11, the proposed planning has the best detection effectiveness, as it ensures the maximum rank of the composite matrix and efficiently eliminates all non-D-FACTS buses at the same time. The comparison between the max-rank planning and proposed planning highlights the importance of covering buses with D-FACTS devices.

It is interesting to compare the ADPs between the full and the proposed planning. Even though the full planning has zero non-D-FACTS buses, its detection performance is not as good as the proposed planning. This is because if an RMTD happens to changes the reactance of all lines connected to one bus using a unity factor, any single-bus FDI attack on this bus is undetectable, according to Lemma 3.4.3. Since D-FACTS devices are placed on all lines connected to each bus in the full planning, this special case happens many times in the IEEE 118-bus system under the one hundred RMTDs. However, it rarely occurs in the RMTDs under the proposed planning due to Rule 1, where there are only ten buses whose

connecting lines are all D-FACTS lines.

With the same number of D-FACTS devices as the proposed planning, the one hundred arbitrary planning solutions have 18 non-D-FACTS buses on average, which limits the performance of detecting single-bus FDI attacks. Furthermore, there are 21 buses whose connecting lines are all D-FACTS lines on average in the one hundred arbitrary planning. The ADP of the arbitrary planning method further decreases when an RMTD happens to modify the reactance of all lines connected to one of these buses using a unity factor. In addition, the ADP range of the four planning methods under the single-bus FDI attacks in Figure 3.11 is consistent with the analytically derived ADP range in Table 3.4. The consistency shows the validity of Theorem 3.4.2 and Corollaries 3.4.1–3.4.3.

It is necessary to investigate the MTD detection effectiveness under each MTD planning method against FDI attacks with different magnitudes of injected false data. Here, we introduce voltage angle injection magnitudes (VAIM) to measure the magnitude of injected false data. For an FDI attack with a given VAIM, injected voltage angle increment $\Delta\theta$ is randomly generated in the range $\Delta\theta \in [0.9, 1.1] \cdot \bar{\theta} \cdot VAIM$, where $\bar{\theta}$ is the average voltage angle. According to the attack vector equation $a = H \cdot \Delta\theta$, the injected false data are proportional to VAIM. We generate FDI attack pools under 13 different VAIMs, ranging from 0.001 to 1.5. For each VAIM, we simulate 1170 single-bus FDI attacks, in which each bus, except the slack bus, in the IEEE 118-bus system is attacked ten times. We adopt the defense pool of each planning method generated previously. The standard deviation of Gaussian noise is 1% of the actual measurement.

Comparative results are shown in Figure 3.12. As seen, MTD detection effectiveness increases with the increase of VAIM. MTDs under all planning methods have very limited capability to detect the FDI attacks with VAIM less than 0.02. This is because when VAIM is extremely small, the injected false data in FDI have the same order of magnitude as measurement noises. When VAIM is less than 0.1, MTDs fail to detect part of these FDI attacks. MTDs increase the estimation residual in these attacks compared with the situation free of attacks, but the increased residual is not large enough to alert the BDD threshold. When the VAIM is more than 0.25, MTDs under each planning method reach their best

detection effectiveness. In addition, simulation results demonstrate that the proposed graph-theory-based planning method has the best detection performance under each magnitude of injected false data magnitude.

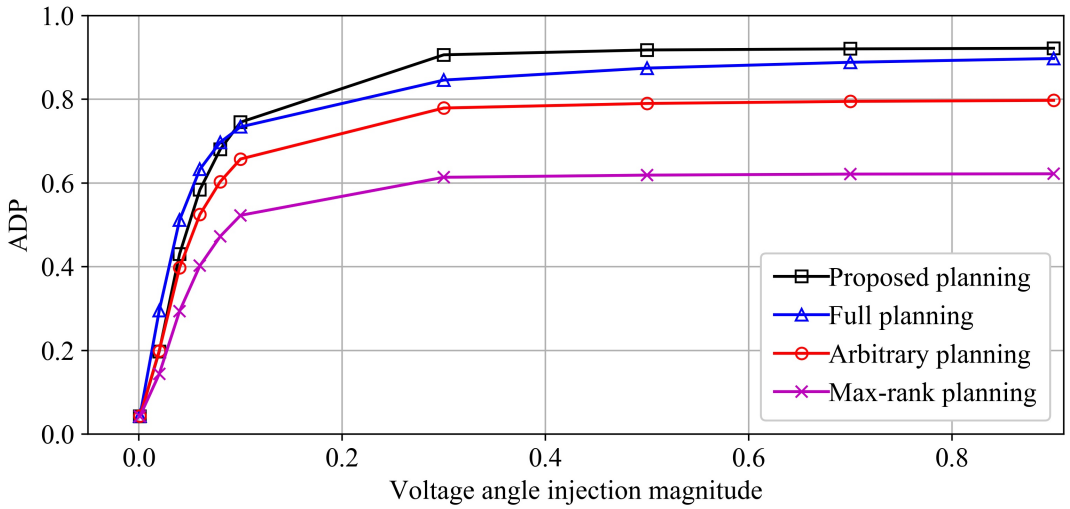


Figure 3.12: *The impact of VAIM on the MTD detection effectiveness under each planning algorithm.*

We further investigate the MTD detection effectiveness under each MTD planning method under different noise magnitudes. The standard deviation of Gaussian noise increases from 1% to 4% of the actual measurements. Here, we use the attack pool composed of 1170 single-bus FDI attacks with 0.5 VAIM. As seen, simulation results demonstrate that low measurement noises contribute to improving the MTD detection effectiveness. This is because a higher noise level causes the BDD to tolerate more deviations between the measured and estimated power flows. However, this tolerance reduces the defenders' capability in detecting FDI attacks. In Figure 3.13, even though the standard deviation of measurement noise reaches 4%, the ADP of the proposed planning method is still more than 0.80. The proposed graph-theory-based planning method has the best detection performance under each level of measurement noise.

We evaluate the detection effectiveness of MTD under the proposed MTD planning under AC-FDI attacks. The average ADP of the four planning methods in the AC noisy condition is illustrated in Figure 3.14. The ADP is calculated by sequentially launching attacks in the

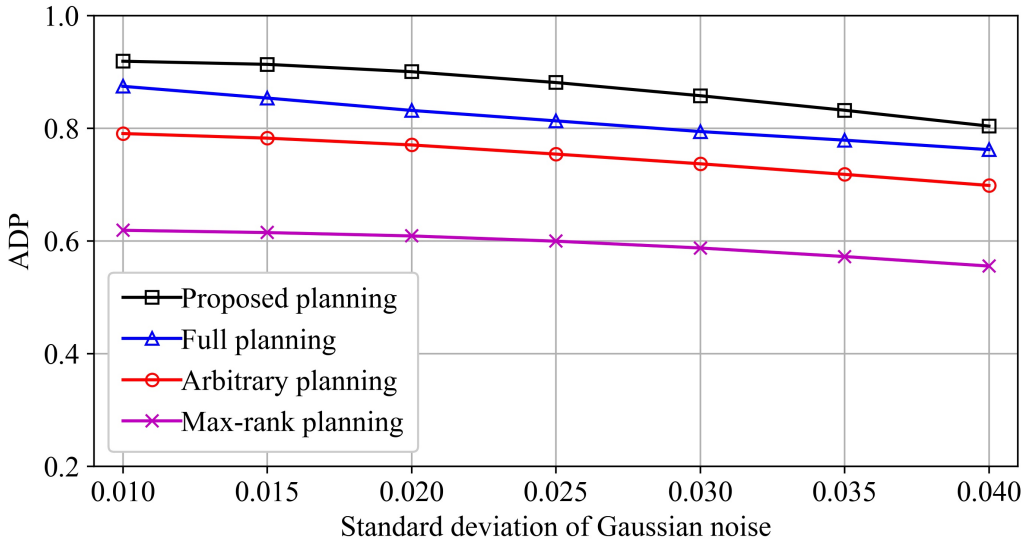


Figure 3.13: *The impact of measurement noises on the MTD detection effectiveness under each planning algorithm.*

AC-FDI attack pool against AC-SE and BDD, when the system operator uses the same D-FACTS setpoints in the generated defense pool of each planning. For the sensor deployment, we adopt a 2.5 redundant factor (i.e., a ratio between the number of measurements and the number of system states) to guarantee the observability in AC-SE. It is seen in Figure 3.14 that the ADP of the max-rank planning and the proposed planning in AC-SE are consistent with those in DC-SE in Figure 3.11. The ADP of the full planning in AC-SE is higher than that in DC-SE. This is because the end bus belonging to the D-FACTS bus is no longer an unprotected bus in AC-SE due to the non-linear relation between the system state and measurements in AC-SE.

3.6 Summary

In this chapter, we address the MTD planning issue considering both MTD detection effectiveness and MTD planning cost. We derive analytical conditions that exhibit distinct MTD requirements on the sensor deployment and the D-FACTS placement. Furthermore, we propose novel max-rank MTD planning algorithms for complete and incomplete MTDs

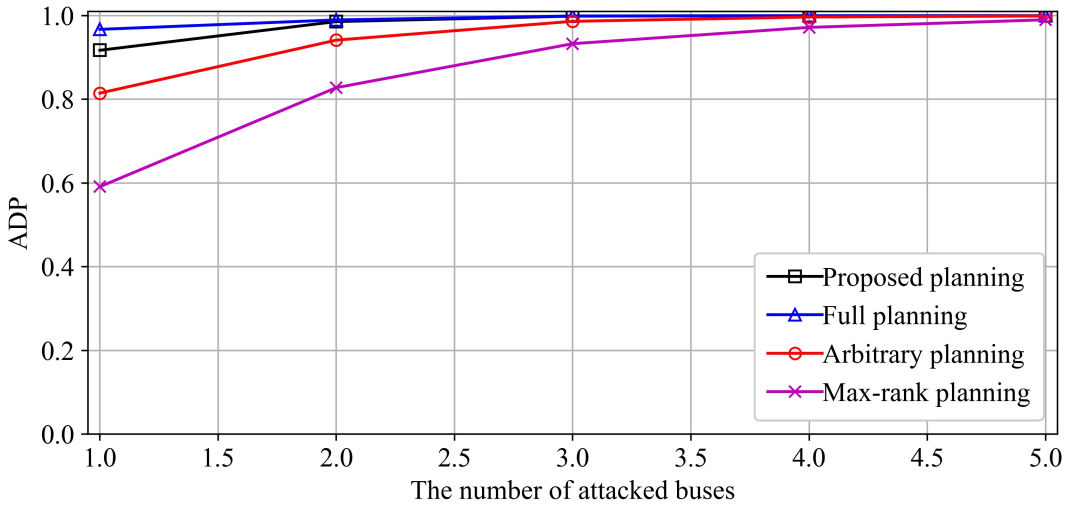


Figure 3.14: *The comparison of ADPs under the four planning algorithms in AC-SE.*

with the goal of maximizing the rank of the composite matrix, which is indicative of the MTD effectiveness. Additionally, we derive an equation to analytically determine the minimum number of D-FACTS devices required to achieve the maximum rank of the composite matrix. Numerical results show that system operators can install D-FACTS devices on less than 40% of transmission lines to reach the maximum rank of the composite matrix. Unlike the existing studies, MTD under the proposed max-rank planning approach leads to a high utilization rate of installed D-FACTS devices. In addition, the proposed max-rank planning solutions guarantee the detection effectiveness of the simplest MTD as long as the D-FACTS devices work in non-idle states.

In addition, we demonstrate the rank of the composite matrix merely determines the lower bound of ADP, and the number of unprotected buses decides the upper bound of ADP. In addition, we rigorously derive the ADP range of several MTD planning methods and propose a graph-theory-based planning method to achieve maximal detection effectiveness. The proposed method eliminates non-D-FACTS buses to increase the ADP upper bound and simultaneously remain a high ADP lower bound. Numerical results show the ADP range in the proposed graph-theory-based planning is consistent with the range we mathematically derive. The ADP of the proposed planning is better than that of the arbitrary planning, the

max-rank planning, and the full planning in DC-SE. But it is slightly less than that of the full planning in AC-SE.

Chapter 4

AC Optimal Power Flow-based MTD Operation Model

After the allocation of distributed flexible AC transmission system (D-FACTS) devices is determined, D-FACTS setpoints need to be determined by the system operator in the moving target defense (MTD) operation. This chapter proposes an AC optimal power flow (ACOPF)-based MTD operation algorithm, in which generation costs and system losses are minimized by introducing the reactance of lines equipped with D-FACTS devices as the extra decision variables in the traditional ACOPF. Furthermore, an interior-point solver is developed for efficiently resolving the proposed ACOPF-based MTD model.

4.1 Introduction

As MTDs on the proposed MTD planning have maximal detection effectiveness, the MTD operation algorithm can focus on exploiting the economic benefits of D-FACTS devices. In addition, as the D-FACTS devices are installed on the lines sensitive to system losses in the MTD planning, D-FACTS devices ought to be utilized to reduce system losses. As we know, D-FACTS devices are traditionally used to manage power flows and minimize system losses in the power system operation[96]. In the current grid operation, system operators install D-

FACTS devices mainly for economic considerations. Cyber-defense benefits from D-FACTS devices in MTD can be viewed as an important by-product in the future smart grid. To fully utilize D-FACTS devices, integration of D-FACTS devices into the mathematical model of real-time operations is necessary. As significant tools in real-time power system operation and control, OPF models can determine the minimum operating cost and system losses, as well as retain the control variables in secure boundaries.

In the literature, work has been done on the incorporation of D-FACTS devices in DC optimal power flow (DCOPF) to study the impact of MTD on generation costs[79]. However, the DCOPF model cannot be used to minimize system losses, which is one of the main functions of D-FACTS devices. A rectangular representation of FACTS devices such as Phase Shift Transformer and the Unified Power Flow Controller were integrated into ACOPF[102]. It is worth mentioning that there are mainly three types of D-FACTS devices, namely, distributed series static compensator (DSSC), distributed series reactor (DSR), and distributed series impedance (DSI). DSSC is similar to a phase shifter studied in reference[102], while DSR and DSI are designed to adjust the impedance of power lines, i.e., the D-FACTS model used in MTD. Although AC power flow models are widely used in practical power systems, ACOPF with the model of DSR and DSI is still missing in the literature. To fill this gap, we propose an ACOPF model considering the D-FACTS devices, in which the reactance of lines equipped with D-FACTS devices are introduced as decision variables. The proposed ACOPF model can be applied in the control center to achieve the minimum system losses and generation costs in real-time while determining the setpoints of D-FACTS devices in MTD simultaneously.

An important reason for the existence of this gap is that the development of an efficient solver to solve the ACOPF model considering D-FACTS devices remains challenging, as the impedance variables introduced substantially complicate the solution of the ACOPF problem. Even though intelligent computational algorithms, such as particle swarm optimization (PSO), genetic algorithm (GA), simulated annealing (SA), and differential evolution (DE), can be used to resolve ACOPF model and ACOPF model considering the FACTS device without deriving the gradient and Hessian matrices[103–105], low computational efficiency of

these algorithms exclude themselves to be used in real-time. On the other hand, it has been proven that interior-point methods are efficient tools to resolve the ACOPF problem[106;107]. Therefore, we first derive the gradient and Hessian matrices of the objective function and the constraints in the proposed ACOPF model with respect to branch impedance. Then, we develop an interior-point solver to resolve the proposed ACOPF by modifying and extending Matlab Interior-Point Solver (MIPS) in MATPOWER developed for the conventional ACOPF[108].

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

1) In MTD operation, how can system operators maximize the economic benefit of D-FACTS devices? How can system operators integrate MTD operation into real-time EMS function? Is there any trade-off between the economic benefit and the MTD detection effectiveness?

2) How can system operators solve the proposed ACOPF model considering D-FACTS devices efficiently?

The first question is solved in Section 4.2, and the second question is answered in Section 4.3.

4.2 ACOPF-based MTD Operation Model

After the installation of D-FACTS devices, system operators can change the setpoints of D-FACTS devices for 1) reducing the system loss through power flow control provided by D-FACTS devices; 2) managing power congestion and reducing the generation cost; and 3) detecting false data injection (FDI) attacks by actively changing the setpoints of D-FACTS devices to prevent attackers from knowing true system configurations. Since D-FACTS devices are installed on the lines with high power loss to impedance sensitivity (PLIS) values under the proposed MTD planning to guarantee the MTD detection effectiveness, a natural idea is that we can also take advantage of these D-FACTS devices to effectively manage system losses. We propose an ACOPF-based MTD operation model, in which the reactance of D-FACTS lines is introduced as decision variables in the traditional ACOPF. The ACOPF-

based MTD model with the objective of minimizing system losses and generation costs is formulated as:

$$\min_{\mathbf{Y}} \ \omega_1 L^s(\mathbf{X}) + \omega_2 \sum_{i=1}^{n_g} f^i(p_g^i) \tag{4.1}$$

$$s.t.\ L^s(\mathbf{X}) = \sum_{i=1}^{n_l} S_i^f + S_i^t \tag{4.1a}$$

$$g_P(\theta, \mathbf{V}, \mathbf{P_g}, \mathbf{x}) = 0 \tag{4.1b}$$

$$g_Q(\theta, \mathbf{V}, \mathbf{Q_g}, \mathbf{x}) = 0 \tag{4.1c}$$

$$h_f(\theta, \mathbf{V}, \mathbf{x}) \leq 0 \tag{4.1d}$$

$$h_t(\theta, \mathbf{V}, \mathbf{x}) \leq 0 \tag{4.1e}$$

$$\theta_{ref} \leq \theta_0 \leq \theta_{ref} \tag{4.1f}$$

$$v_i^{\min} \leq v_i \leq v_i^{\max},\ i = 1, ..., n_b \tag{4.1g}$$

$$p_i^{\min} \leq p_i \leq p_i^{\max},\ i \in K_g \tag{4.1h}$$

$$q_i^{\min} \leq q_i \leq q_i^{\max},\ i \in K_g \tag{4.1i}$$

$$\tau x_i^0 \leq |x_i - x_i^0| \leq \eta x_i^0,\ i \in K_{DF} \tag{4.1j}$$

where $\mathbf{X} = \begin{bmatrix} \theta & \mathbf{V} & \mathbf{P_g} & \mathbf{Q_g} & \mathbf{x} \end{bmatrix}$ are decision variables corresponding to voltage angle, voltage magnitude, generator active generation, generator reactive generation, and reactance of D-FACTS lines, respectively; ω_1 and ω_2 are weight parameters; n_b, n_l, n_g and n_{DF} are the number of buses, lines, generators, and D-FACTS lines, respectively; K_{DF} is the line index set of D-FACTS lines, and K_g is the bus index set of generators; $L^s(\mathbf{X})$ is the system loss; f^i is the active power generation cost of the i-th generator; S_i^f and S_i^t are complex power flows at the from-end and to-end of the i-th line; (4.1b) and (4.1c) are nonlinear equality constraints of the nodal active and reactive power balance, respectively; (4.1d) and (4.1e) are nonlinear inequality constraints of line power flow limits corresponding to lines starting

from from-end and to-end, respectively; (4.1f) and (4.1g) are voltage angle and magnitude constraints; (4.1h) and (4.1i) are generator constraints; in (4.1j), η in % reflects the physical capacity of D-FACTS devices and τ in % is introduced to decide if D-FACTS devices can be operated in an idle state. It is worth mentioning that, in order to maximize the utilization of installed D-FACTS devices (i.e., to minimize the number of required D-FACTS devices), none of D-FACTS devices ought to work in the idle state (i.e.,$\tau \neq 0$) under the proposed MTD planning. Nevertheless, when the D-FACTS device is allowed to operate in an idle state, the proposed ACOPF-based MTD model (4.1) is also applicable by setting $\tau = 0$, whereby its operating point can be optimally determined.

The proposed operating approaches can be seamlessly integrated into the existing energy management system (EMS) of a power system in the control room. In the control room, the SE, usually launched every 5-10 minutes in real-time, provides the current system states for ACOPF[109]. The conventional ACOPF is launched every 5-10 minutes in the real-time operation of the transmission system[110]. The proposed ACOPF-based MTD can have the same periodicity as ACOPF. Thus, system operators can apply the proposed ACOPF-based MTD to manage the system operation in real-time and determine the setpoints of D-FACTS devices in their control room. In this way, the maximum rank of the composite matrix is guaranteed, translating to effective MTD. The proposed MTD planning and operational approaches enable the system operator to detect FDI attacks by directly using the existing state estimation (SE) and bad data detection (BDD).

We demonstrate the time-sequence diagram of SE and MTD in Figure 4.1, where T_1 is the time period of SE and MTD, and T_2 is the detection time for FDI attacks. During each time period, system operators perform SE by using updated line parameters containing new setpoints of D-FACTS devices, followed by BDD to detect the existence of FDI attacks. It can be seen that the maximum detection time for FDI attacks is T_1 with reference to the FDI attacks launched. In addition to running MTD periodically, MTD can be used as an event-based defense strategy. Specifically, system operators can apply MTD on an ad-hoc basis if they find abnormal situations attributed to possible cyber-attacks.

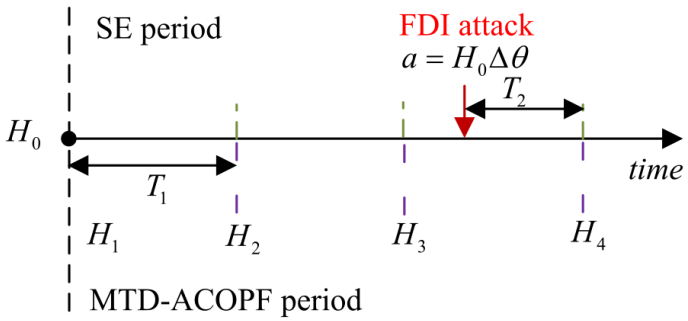


Figure 4.1: *Time-sequence diagram of FDI attacks and MTDs.*

4.3 An Interior-Point Solver for ACOPF-based MTD Model

In this section, we utilize the interior-point solver to solve the proposed ACOPF-based MTD model by modifying and extending the MIPS in MATPOWER[108]. Work[108] provides the first derivatives and Hessian matrices of objective function and constraints in the conventional ACOPF model. More specifically, voltages are in polar coordinates and nodal balance equations are expressed by complex power. Here, we follow suit and extend the interior-point solver for the proposed ACOPF-based MTD model. We derive the gradient and Hessian matrices of nonlinear equality constraints, inequality constraints, and objective function with respect to the reactance of lines equipped with D-FACTS devices.

4.3.1 Preliminaries in Derivatives in ACOPF Model

Let $\mathbf{V}$ be a vector of complex voltages of all buses. Then, the first derivatives of complex voltage with respect to voltage angle and magnitude are given as follows:

$$\mathbf{V}_\theta = \frac{\partial \mathbf{V}}{\partial \theta} = j[\mathbf{V}]$$

$$\mathbf{V}_v = \frac{\partial \mathbf{V}}{\partial v} = [\mathbf{E}]$$

where θ is the voltage angle vector; υ is the voltage magnitude vector; $\mathbf{E} = [\upsilon]^{-1}\mathbf{V}$; $[\centerdot]$ is a diagonalizable operator defined in reference[108], which converts a vector to a diagonal matrix, i.e., $[\centerdot] : \mathbb{C}^n \to \mathbb{C}^{n\times n}$. For example, when we apply the diagonalizable operator on $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, we have $[\mathbf{b}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

As independent variables in the proposed ACOPF model, the line reactance controlled by the D-FACTS devices directly determines the nodal admittance matrix. Since the nodal admittance matrix plays an important role in deriving the derivatives of objective function and constraints with respect to the reactance, we present the definition of nodal admittance in MATPOWER. MATPOWER models line parameters, transformers and shunt elements in the nodal admittance matrix, which is defined as:

$$\mathbf{Y}_{bus} = C_f^T \left([\mathbf{Y}_{ff}]C_f + [\mathbf{Y}_{ft}]C_t\right) + C_t^T \left([\mathbf{Y}_{tf}]C_f + [\mathbf{Y}_{tt}]C_t\right) + [\mathbf{Y}_{sh}] \tag{4.2}$$

where $Y_{ft}^i = -y_s^i(\tau e^{-j\theta_{sh}})^{-1}$, $Y_{tf}^i = -y_s^i(\tau e^{j\theta_{sh}})^{-1}$, $Y_{tt}^i = y_s^i + j0.5b_c$ and $Y_{ff}^i = (y_s^i + j0.5b_c)\tau^{-2}$ are equivalent admittance of the i-th line between different ends in the standard π transmission line model; y_s^i is the admittance of the i-th line, i.e., $y_s^i = (r_i + jx_i)^{-1}$, and r_i and x_i are resistance and reactance of the i-th line, respectively; τ is the transformer tap ratio magnitude, and θ_{sh} is the transformer phase shift angle; Y_{sh}^i is admittance of shunt elements of the i-th bus; C_t and C_f are connection matrices used in building the system admittance matrices, defined in reference[108]. The first and second derivatives of y_s^i with respect to the line reactance are calculated as:

$$\frac{\partial y_s^i}{\partial x_i} = \frac{-2x_i r_i + j(x_i^2 - r_i^2)}{(x_i^2 + r_i^2)^2}$$

$$\frac{\partial^2 y_s^i}{\partial x_i^2} = 2\frac{r(3x_i^2 - r_i^2) + jx_i(3r_i^2 - x_i^2)}{(x_i^2 + r_i^2)^3}$$

The gradient of $\mathbf{Y}_{ff}, \mathbf{Y}_{ft}, \mathbf{Y}_{tf}$ and $\mathbf{Y}_{tt}$ with respect to line reactance are diagonal matrices and their diagonal entries can be calculated as follows:

$$\nabla_{\mathbf{x}} \mathbf{Y}_{ff}(i,i) = \frac{\partial \mathbf{Y}_{ff}}{\partial \mathbf{x}}(i,i) = \frac{1}{\tau^2} \frac{\partial y_s^i}{\partial x_i} \tag{4.3}$$

$$\nabla_{\mathbf{x}} \mathbf{Y}_{ft}(i,i) = -\frac{1}{\tau e^{-j\theta_{sh}}} \frac{\partial y_s^i}{\partial x_i} \tag{4.4}$$

$$\nabla_{\mathbf{x}} \mathbf{Y}_{tf}(i,i) = -\frac{1}{\tau e^{j\theta_{sh}}} \frac{\partial y_s^i}{\partial x_i} \tag{4.5}$$

$$\nabla_{\mathbf{x}} \mathbf{Y}_{tt}(i,i) = \frac{\partial y_s^i}{\partial x_i} \tag{4.6}$$

Similarly, the Hessian matrices of $\mathbf{Y}_{ff}$, $\mathbf{Y}_{ft}$, $\mathbf{Y}_{tf}$, $\mathbf{Y}_{tt}$ are diagonal matrices as shown in (4.7), whose diagonal entries can be calculated as $\nabla^2_{\mathbf{xx}} \mathbf{Y}_{ff}(i,i) = \frac{1}{\tau^2} \frac{\partial^2 y_s^i}{\partial x_i^2}$, $\nabla^2_{\mathbf{xx}} \mathbf{Y}_{tt}(i,i) = \frac{\partial^2 y_s^i}{\partial x_i^2}$, $\nabla^2_{\mathbf{xx}} \mathbf{Y}_{ft}(i,i) = -\frac{1}{\tau e^{-j\theta_{sh}}} \frac{\partial^2 y_s^i}{\partial x_i^2}$, and $\nabla^2_{\mathbf{xx}} \mathbf{Y}_{tf}(i,i) = -\frac{1}{\tau e^{j\theta_{sh}}} \frac{\partial^2 y_s^i}{\partial x_i^2}$, respectively. Note that the first and second derivatives of the admittance of shunt elements with respect to reactance are zero matrices, i.e., $\nabla_{\mathbf{x}} \mathbf{Y}_{sh} = \mathbf{0}$ and $\nabla^2_{\mathbf{xx}} \mathbf{Y}_{sh} = \mathbf{0}$. For presentation simplicity, the subscripts of the gradient and Hessian matrices of all admittance matrices with respect to the reactance are omitted hereinafter, i.e., $\nabla \mathbf{Y}_{ff} = \nabla_{\mathbf{x}} \mathbf{Y}_{ff}$ and $\nabla^2 \mathbf{Y}_{ff} = \nabla^2_{\mathbf{xx}} \mathbf{Y}_{ff}$.

$$\begin{aligned}
&\nabla^2 \mathbf{Y}_{ff} = \tfrac{\partial}{\partial \mathbf{x}}((\tfrac{\partial}{\partial \mathbf{x}} \mathbf{Y}_{ff}^T) \times \mathbf{a}) = \\
&= \begin{bmatrix}
\frac{\partial}{\partial x_1}(\frac{\partial Y_{ff}^1}{\partial x_1}) & \frac{\partial}{\partial x_2}(\frac{\partial Y_{ff}^1}{\partial x_1}) & \cdots & \frac{\partial}{\partial x_{n_l}}(\frac{\partial Y_{ff}^1}{\partial x_1}) \\
\frac{\partial}{\partial x_1}(\frac{\partial Y_{ff}^2}{\partial x_2}) & \frac{\partial}{\partial x_2}(\frac{\partial Y_{ff}^2}{\partial x_2}) & \cdots & \frac{\partial}{\partial x_{n_l}}(\frac{\partial Y_{ff}^2}{\partial x_2}) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial}{\partial x_1}(\frac{\partial Y_{ff}^n}{\partial x_{n_l}}) & \frac{\partial}{\partial x_2}(\frac{\partial Y_{ff}^n}{\partial x_{n_l}}) & \cdots & \frac{\partial}{\partial x_{n_l}}(\frac{\partial Y_{ff}^n}{\partial x_{n_l}})
\end{bmatrix} = \begin{bmatrix}
\frac{\partial^2 Y_{ff}^1}{\partial x_1^2} & 0 & \cdots & 0 \\
0 & \frac{\partial^2 Y_{ff}^2}{\partial x_2^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{\partial^2 Y_{ff}^n}{\partial x_{n_l}^2}
\end{bmatrix}
\end{aligned} \tag{4.7}$$

where $\mathbf{a} = \mathbf{1} \in \mathbb{R}^{n_l \times 1}$.

4.3.2 Gradient of Power Injection Constraints

The complex power balance equations can be expressed as $G^s(\mathbf{X}) = \mathbf{S}^{bus} + \mathbf{S}_d - \mathbf{C}_g\mathbf{S}_g = \mathbf{0}$, where $\mathbf{S}_d$ is a complex power load vector of all buses; $\mathbf{S}_g$ is a complex power generation vector of all buses; $\mathbf{S}^{bus}$ is a complex power injection vector of all buses, i.e., $\mathbf{S}^{bus} = [\mathbf{V}]\mathbf{I}^{bus*}$, and $\mathbf{I}^{bus}$ is a complex current injection vector, i.e., $\mathbf{I}^{bus} = \mathbf{Y}_{bus}\mathbf{V}$. The gradient of power balance equations can be expressed as follows:

$$G^s_{\mathbf{X}}(\mathbf{X}) = [G^s_\theta \quad G^s_\nu \quad G^s_{\mathbf{P_g}} \quad G^s_{\mathbf{Q_g}} \quad G^s_{\mathbf{x}}]$$

where the first four items irrelative to the line reactance are consistent with the results in reference[108]. We can calculate $G^s_{\mathbf{x}}$ by using equations (4.1) and (4.3)-(4.6) as follows:

$$\begin{aligned}
&G^s_{\mathbf{x}} = \mathbf{S}^{bus}_{\mathbf{x}} = \tfrac{\partial}{\partial \mathbf{x}}([\mathbf{V}]\mathbf{Y}^*_{bus}\mathbf{V}^*) \\
&= \tfrac{\partial}{\partial \mathbf{x}}\left\{[\mathbf{V}]\left\{C_f^T\left([\mathbf{Y}_{ff}]C_f + [\mathbf{Y}_{ft}]C_t\right) + C_t^T\left([\mathbf{Y}_{tf}]C_f + [\mathbf{Y}_{tt}]C_t\right) + [\mathbf{Y}_{sh}]\right\}\mathbf{V}^*\right\} \\
&= \tfrac{\partial}{\partial \mathbf{x}}\left\{[\mathbf{V}]\left\{C_f^T\left([C_f\mathbf{V}^*]\mathbf{Y}^*_{ff} + [C_t\mathbf{V}^*]\mathbf{Y}^*_{ft}\right) + C_t^T\left([C_f\mathbf{V}^*]\mathbf{Y}^*_{tf} + [C_t\mathbf{V}^*]\mathbf{Y}^*_{tt}\right)\right\}\right\} \\
&= [\mathbf{V}]C_f^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{ff} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{ft}\right) + [\mathbf{V}]C_t^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{tf} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{tt}\right)
\end{aligned} \tag{4.8}$$

4.3.3 Hessian Matrix of Power Injection Constraints

The Hessian matrix of complex power balance constraints in the proposed ACOPF can be expressed as:

$$G^s_{\mathbf{XX}}(\mathbf{X}) = \tfrac{\partial}{\partial \mathbf{X}}({G^s_{\mathbf{X}}}^T\lambda) = \begin{bmatrix} G^s_{\theta\theta} & G^s_{\theta\nu} & 0 & 0 & G^s_{\theta\mathbf{x}} \\ G^s_{\nu\theta} & G^s_{\nu\nu} & 0 & 0 & G^s_{\nu\mathbf{x}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ G^s_{\mathbf{x}\theta} & G^s_{\mathbf{x}\nu} & 0 & 0 & G^s_{\mathbf{xx}} \end{bmatrix} \tag{4.9}$$

where λ is a constant vector for calculating the Hessian matrix; the expressions of $G^s_{\theta\theta}$, $G^s_{\theta\nu}$, $G^s_{\nu\theta}$ and $G^s_{\nu\nu}$ can be found in reference[108]; $G^s_{\mathbf{xx}}$, $G^s_{\mathbf{x}\nu}$, $G^s_{\mathbf{x}\theta}$, $G^s_{\theta\mathbf{x}}$, and $G^s_{\nu\mathbf{x}}$ need to be derived for the proposed ACOPF model. Due to the space limit, we ignore the derivation

process and directly present the calculation results. First, we calculate $G^s_{\mathbf{xx}}$ using equation (4.8).

$$\begin{aligned} G^s_{\mathbf{xx}} &= \tfrac{\partial}{\partial \mathbf{x}}(G^T_{\mathbf{x}}\lambda) \\ &= [C_f[\mathbf{V}]\lambda]([C_f\mathbf{V}^*]\nabla^2_{\mathbf{xx}}\mathbf{Y}^*_{ff} + [C_t\mathbf{V}^*]\nabla^2\mathbf{Y}^*_{ft}) \\ &+ [C_t[\mathbf{V}]\lambda]([C_f\mathbf{V}^*]\nabla^2_{\mathbf{xx}}\mathbf{Y}^*_{tf} + [C_t\mathbf{V}^*]\nabla^2\mathbf{Y}^*_{tt}) \end{aligned}$$

Similarly, we calculate $G^s_{\mathbf{x}\nu}$, $G^s_{\nu\mathbf{x}}$, $G^s_{\mathbf{x}\theta}$, and $G^s_{\theta\mathbf{x}}$ as follows:

$$\begin{aligned} G^s_{\mathbf{x}\nu} &= ([\nabla\mathbf{Y}^*_{ff}C_f[\mathbf{V}]\lambda] + [\nabla\mathbf{Y}^*_{tf}C_t[\mathbf{V}]\lambda])C_f[\mathbf{E}]^* + ([\nabla\mathbf{Y}^*_{ft}C_f[\mathbf{V}]\lambda] + [\nabla\mathbf{Y}^*_{tt}C_t[\mathbf{V}]\lambda])C_t[\mathbf{E}]^* \\ &+ [C_f\mathbf{V}^*](\nabla\mathbf{Y}^*_{ff}C_f + \nabla\mathbf{Y}^*_{tf}C_t)[\lambda][\mathbf{E}] + [C_t\mathbf{V}^*](\nabla\mathbf{Y}^*_{ft}C_f + \nabla\mathbf{Y}^*_{tt}C_t)[\lambda][\mathbf{E}] \end{aligned}$$

$$\begin{aligned} G^s_{\theta\mathbf{x}} &= j[\mathbf{V}][\lambda]C_f^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{ff} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{ft}\right) + j[\mathbf{V}][\lambda]C_t^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{tf} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{tt}\right) \\ &- j[\mathbf{V}^*]C_f^T\left([C_f[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{ff} + [C_t[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{tf}\right) - j[\mathbf{V}^*]C_t^T\left([C_f[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{ft} + [C_t[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{tt}\right) \end{aligned}$$

$$\begin{aligned} G^s_{\mathbf{x}\theta} &= j\left\{[C_f\mathbf{V}^*](\nabla\mathbf{Y}^*_{ff}C_f + \nabla\mathbf{Y}^*_{tf}C_t) + [C_t\mathbf{V}^*](\nabla\mathbf{Y}^*_{ft}C_f + \nabla\mathbf{Y}^*_{tt}C_t)\right\}[\lambda][\mathbf{V}] \\ &- j\left\{[\nabla\mathbf{Y}^*_{ff}C_f[\mathbf{V}]\lambda] + [\nabla\mathbf{Y}^*_{tf}C_t[\mathbf{V}]\lambda]\right\}C_f[\mathbf{V}]^* - j\left\{[\nabla\mathbf{Y}^*_{ft}C_f[\mathbf{V}]\lambda] + [\nabla\mathbf{Y}^*_{tt}C_t[\mathbf{V}]\lambda]\right\}C_t[\mathbf{V}]^* \end{aligned}$$

$$\begin{aligned} G^s_{\nu\mathbf{x}} &= [\mathbf{E}][\lambda]C_f^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{ff} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{ft}\right) + [\mathbf{E}][\lambda]C_t^T\left([C_f\mathbf{V}^*]\nabla\mathbf{Y}^*_{tf} + [C_t\mathbf{V}^*]\nabla\mathbf{Y}^*_{tt}\right) \\ &+ [\mathbf{E}^*]C_f^T([C_f[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{ff} + [C_t[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{tf}) + [\mathbf{E}^*]C_t^T([C_f[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{ft} + [C_t[\mathbf{V}]\lambda]\nabla\mathbf{Y}^*_{tt}) \end{aligned}$$

4.3.4 Gradient and Hessian Matrix of Power Flow Constraints

In the power flow constraints, we derive the gradient and Hessian matrix for the complex power flow at the from-ends of the lines. The derivative results for the to-ends of the line can be identically calculated by replacing all f sub/super-scripts with t. Similar to power injection constraints, the first derivatives of the power flow with respect to voltage angle, voltage magnitude, real and reactive power generation are identical to that in reference[108].

We only derive the first derivatives of power flow with respect to reactance as follows.

$$\begin{aligned}\mathbf{S}_{\mathbf{x}}^{f} &= \nabla_{\mathbf{x}}\mathbf{S}_{f} = \tfrac{\partial}{\partial \mathbf{x}}\left\{[C_f\mathbf{V}]([\mathbf{Y}_{ff}^{*}]C_f + [\mathbf{Y}_{ft}^{*}]C_t)\mathbf{V}^{*}\right\} \\ &= [C_f\mathbf{V}]([C_f\mathbf{V}^{*}]\nabla\mathbf{Y}_{ff}^{*} + [C_t\mathbf{V}^{*}]\nabla\mathbf{Y}_{ft}^{*})\end{aligned}$$

The Hessian matrix of complex power flow constraints in the proposed ACOPF has the same form as that of power flow constraints. In the Hessian matrix, $\mathbf{S}_{\theta\theta}^{f}$, $\mathbf{S}_{\theta\nu}^{f}$, $\mathbf{S}_{\nu\theta}^{f}$ and $\mathbf{S}_{\nu\nu}^{f}$ are identical to that in reference[108]; $\mathbf{S}_{\mathbf{xx}}^{f}$, $\mathbf{S}_{\mathbf{x}\nu}^{f}$, $\mathbf{S}_{\mathbf{x}\theta}^{f}$, $\mathbf{S}_{\theta\mathbf{x}}^{f}$ and $\mathbf{S}_{\nu\mathbf{x}}^{f}$ can be derived as follows, using $\mathbf{S}_{\theta}^{f}$ and $\mathbf{S}_{\nu}^{f}$ in reference[108].

$$\mathbf{S}_{\mathbf{xx}}^{f} = \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}_{\mathbf{x}}^{T}\lambda) = [C_f\mathbf{V}][\lambda]([C_f\mathbf{V}^{*}]\nabla^2\mathbf{Y}_{ff}^{*} + [C_t\mathbf{V}^{*}]\nabla^2\mathbf{Y}_{ft}^{*}) \tag{4.10}$$

$$\mathbf{S}_{\mathbf{xv}}^{f} = \nabla\mathbf{Y}_{ff}^{*}[\lambda]([C_f\mathbf{V}]C_f[\mathbf{E}]^{*} + [C_f\mathbf{V}^{*}]C_f[\mathbf{E}]) + \nabla\mathbf{Y}_{ft}^{*}[\lambda]([C_f\mathbf{V}]C_t[\mathbf{E}]^{*} + [C_t\mathbf{V}^{*}]C_f[\mathbf{E}]) \tag{4.11}$$

$$\mathbf{S}_{\mathbf{x}\theta}^{f} = j\nabla\mathbf{Y}_{ff}^{*}[\lambda](-[C_f\mathbf{V}]C_f[\mathbf{V}]^{*} + [C_f\mathbf{V}^{*}]C_f[\mathbf{V}]) + j\nabla\mathbf{Y}_{ft}^{*}[\lambda](-[C_f\mathbf{V}]C_t[\mathbf{V}]^{*} + [C_t\mathbf{V}^{*}]C_f[\mathbf{V}]) \tag{4.12}$$

$$\mathbf{S}_{\mathbf{vx}}^{f} = [\mathbf{E}]C_f^{T}[\lambda]([C_f\mathbf{V}^{*}]\nabla\mathbf{Y}_{ff}^{*} + [C_t\mathbf{V}^{*}]\nabla\mathbf{Y}_{ft}^{*}) + [\mathbf{E}^{*}](C_f^{T}[C_f\mathbf{V}][\lambda]\nabla\mathbf{Y}_{ff}^{*} + C_t^{T}[C_f\mathbf{V}][\lambda]\nabla\mathbf{Y}_{ft}^{*}) \tag{4.13}$$

$$\mathbf{S}_{\theta\mathbf{x}}^{f} = j[\mathbf{V}]C_f^{T}[\lambda]([C_f\mathbf{V}^{*}]\nabla\mathbf{Y}_{ff}^{*} + [C_t\mathbf{V}^{*}]\nabla\mathbf{Y}_{ft}^{*}) - j[\mathbf{V}^{*}](C_f^{T}[C_f\mathbf{V}][\lambda]\nabla\mathbf{Y}_{ff}^{*} + C_t^{T}[C_f\mathbf{V}][\lambda]\nabla\mathbf{Y}_{ft}^{*}) \tag{4.14}$$

4.3.5 Gradient and Hessian Matrix of System Losses

The system complex power loss is the sum of complex power loss of each line, and the line power loss is the sum of complex power flows at the from-end and to-end of this line, as shown in (4.1a). The system loss can be expressed in matrix form, i.e., $L^{s} = a^{T}(\mathbf{S}^{f} + \mathbf{S}^{t})$, where $a = \mathbf{1} \in \mathbb{R}^{n_l \times 1}$. Therefore, the first derivative of the system loss is $L_{\mathbf{x}}^{s} = a^{T}(\mathbf{S}_{\mathbf{x}}^{f} + \mathbf{S}_{\mathbf{x}}^{t})$.

The Hessian matrix of the system loss has the same form as that of power injection

constraints in equation (4.9). Take $L^s_{\mathbf{xx}}$ for example, $L^s_{\mathbf{xx}}$ can be calculated as follows:

$$\begin{aligned} L^s_{\mathbf{xx}} &= \tfrac{\partial}{\partial \mathbf{x}}((L^s_{\mathbf{x}})^T) = \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{fT}_{\mathbf{x}} a + \mathbf{S}^{tT}_{\mathbf{x}} a) \\ &= \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{fT}_{\mathbf{x}} \lambda)\big|_{\lambda=a} + \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{tT}_{\mathbf{x}} \lambda)\big|_{\lambda=a} \\ &= \mathbf{S}^{f}_{\mathbf{xx}}\big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{xx}}\big|_{\lambda=a} \end{aligned}$$

Note that we have $\frac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{fT}_{\mathbf{x}} a) = \mathbf{S}^{f}_{\mathbf{xx}}\big|_{\lambda=a}$ according to equation (4.10).

Similarly, the remaining none-zero matrix blocks in the Hessian matrix can be calculated as follows:

$$L^s_{\theta\theta} = \tfrac{\partial}{\partial \theta}((L^s_{\theta})^T) = \tfrac{\partial}{\partial \theta}(\mathbf{S}^{fT}_{\theta} a + \mathbf{S}^{tT}_{\theta} a) = \mathbf{S}^{f}_{\theta\theta}\Big|_{\lambda=a} + \mathbf{S}^{t}_{\theta\theta}\big|_{\lambda=a}$$

$$L^s_{\mathbf{vv}} = \tfrac{\partial}{\partial \mathbf{v}}((L^s_{\mathbf{v}})^T) = \tfrac{\partial}{\partial \mathbf{v}}(\mathbf{S}^{fT}_{\mathbf{v}} a + \mathbf{S}^{tT}_{\mathbf{v}} a) = \mathbf{S}^{f}_{\mathbf{vv}}\big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{vv}}\big|_{\lambda=a}$$

$$L^s_{\mathbf{x}\theta} = \tfrac{\partial}{\partial \theta}((L^s_{\mathbf{x}})^T) = \tfrac{\partial}{\partial \theta}(\mathbf{S}^{fT}_{\mathbf{x}} a + \mathbf{S}^{tT}_{\mathbf{x}} a) = \mathbf{S}^{f}_{\mathbf{x}\theta}\Big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{x}\theta}\big|_{\lambda=a}$$

$$L^s_{\mathbf{xv}} = \tfrac{\partial}{\partial \mathbf{v}}((L^s_{\mathbf{x}})^T) = \tfrac{\partial}{\partial \mathbf{v}}(\mathbf{S}^{fT}_{\mathbf{x}} a + \mathbf{S}^{tT}_{\mathbf{x}} a) = \mathbf{S}^{f}_{\mathbf{xv}}\big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{xv}}\big|_{\lambda=a}$$

$$L^s_{\theta\nu} = \tfrac{\partial}{\partial \mathbf{v}}((L^s_{\theta})^T) = \tfrac{\partial}{\partial \mathbf{v}}(\mathbf{S}^{fT}_{\theta} a + \mathbf{S}^{tT}_{\theta} a) = \mathbf{S}^{f}_{\theta\mathbf{v}}\Big|_{\lambda=a} + \mathbf{S}^{t}_{\theta\mathbf{v}}\big|_{\lambda=a}$$

$$L^s_{\nu\theta} = \tfrac{\partial}{\partial \theta}((L^s_{\nu})^T) = \tfrac{\partial}{\partial \theta}(\mathbf{S}^{fT}_{\mathbf{v}} a + \mathbf{S}^{tT}_{\mathbf{v}} a) = \mathbf{S}^{f}_{\mathbf{v}\theta}\Big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{v}\theta}\big|_{\lambda=a}$$

$$L^s_{\theta\mathbf{x}} = \tfrac{\partial}{\partial \mathbf{x}}((L^s_{\theta})^T) = \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{fT}_{\theta} a + \mathbf{S}^{tT}_{\theta} a) = \mathbf{S}^{f}_{\theta\mathbf{x}}\Big|_{\lambda=a} + \mathbf{S}^{t}_{\theta\mathbf{x}}\big|_{\lambda=a}$$

$$L^s_{\nu\mathbf{x}} = \tfrac{\partial}{\partial \mathbf{x}}((L^s_{\nu})^T) = \tfrac{\partial}{\partial \mathbf{x}}(\mathbf{S}^{fT}_{\mathbf{v}} a + \mathbf{S}^{tT}_{\mathbf{v}} a) = \mathbf{S}^{f}_{\mathbf{vx}}\big|_{\lambda=a} + \mathbf{S}^{t}_{\mathbf{vx}}\big|_{\lambda=a}$$

This work only minimizes the active power loss. Then, the gradient and Hessian matrix of the reactive power loss can be simply obtained by taking the imaginary part of that of the complex power loss. Note that matrix blocks in the gradient and Hessian matrix of generation costs related to line reactance are zero matrices, and the remaining matrix blocks are identical to the results in reference[108].

4.4 Experiment Results

To validate the validity of the proposed ACOPF model and effectiveness of the developed interior-point solver, we conduct case studies on the IEEE 118-bus transmission system. The algorithms are performed on a laptop with Intel Core i7 processor CPU 2.90 GHz with 8 GB RAM.

4.4.1 Comparison of Traditional ACOPF and ACOPF-based MTD

We compare generation costs, system losses, and CPU time in the following three cases. Case 0: the conventional ACOPF is applied; Case 1: the proposed ACOPF model with $\omega_1 = 0$ and $\omega_2 = 1$ is used only to minimize the generation cost; Case 2: the proposed ACOPF model with $\omega_1 = 1000$ and $\omega_2 = 1$ is used to minimize the generation cost and the system loss. We identify the maximum line power flow using the conventional ACOPF under the default load in MATPOWER, denoted by $S^f_{\max}$. Then, we make the power flow limit of each line equal to $k \times S^f_{\max}$, where factor $k = \{0.4,\ 0.6,\ 0.8,\ 1\}$ in different tests. We assume that D-FACTS devices are installed on all lines, and set $\eta = 20\%$ to be consistent with the D-FACTS setting in reference[79].

The simulation results are listed in Table 4.1. The generation cost in the proposed ACOPF is always less than that in the conventional ACOPF since the dispatchable line reactance can reduce the congestion in the system. The system loss in Case 2 is always less than that in Cases 0 and 1 under different flow limit conditions, which indicates the effectiveness of the proposed ACOPF model in minimizing system losses. The CPU time for solving the proposed ACOPF is less than 15 seconds in most cases, suggesting that the proposed ACOPF model can be applied in real-time system operations using the modified interior-point solver.

Table 4.1: *Costs, losses and CPU time under different flow limit conditions*

k		0.4	0.6	0.8	1.0
Generation Cost ($)	Case 0	131,395	130,337	129,830	129,660
	Case 1	131,219	130,150	129,643	129,475
	Case 2	131,242	130,170	129,664	129,498
System Loss (MW)	Case 0	67.33	70.73	73.54	77.39
	Case 1	65.45	67.09	70.25	74.07
	Case 2	60.63	63.04	65.91	69.18
CPU Time (s)	Case 0	0.22	0.20	0.41	0.45
	Case 1	5.58	3.93	2.41	3.26
	Case 2	4.81	5.66	4.64	13.24

4.4.2 Impact of MTD Planning on System Losses

Installing D-FACTS devices on the lines with the highest PLIS values generally results in the minimum system losses[96]. However, the proposed MTD planning algorithms do not necessarily install D-FACTS devices on the entirety of those lines to maximize the MTD detection effectiveness, which may in turn increase system losses. To investigate the impact of MTD planning on system losses and the MTD detection effectiveness, simulations are also carried out in the IEEE 118-bus system at the peak hour. We construct the following six MTD planning cases, in which 62 is the number of D-FACTS devices identified by Algorithm 1, and the proposed ACOPF-based MTD with $\tau = 0$ is conducted.

Case 0: This is the base case where no D-FACTS devices are used in the system.

Case 1: D-FACTS devices are installed on 62 lines with the lowest PLIS.

Case 2: D-FACTS devices are placed on 62 randomly selected lines.

Case 3: D-FACTS devices are placed on 62 lines chosen by Algorithm 1.

Case 4: D-FACTS devices are installed on 62 lines with the highest PLIS.

Case 5: D-FACTS devices are installed on all 179 lines in the system.

The total PLIS of D-FACTS lines, system losses, the rank of the composite matrix, and the loss reduction in the above six cases are summarized in Table 4.2. It is seen that installing D-FACTS devices using the max-rank planning algorithm can reduce 1.29% of the system loss compared with that in Case 0. The system loss reduction in Case 3 is higher than that

in Cases 1 and 2, and slightly lower than that in Cases 4 and 5. However, the rank of the composite matrix resulted from the proposed max-rank planning algorithm in Case 3 is the same as that in Case 5, which is significantly higher than that in Cases 1, 2 and 4. The comparison in Table 4.2 illustrates that the proposed MTD planning solution is effective in finding a trade-off between the system loss minimization and the rank maximization of the composite matrix, translating into economic yet cyber-secure operations of the power system.

Table 4.2: *System losses under different MTD planning algorithms*

MTD planning	PLIS sum	Loss (MW)	Loss decrease (%)	Rank of composite matrix
Case 0	0	30.67	–	117
Case 1	0.10	30.59	0.25	155
Case 2	1.55	30.55	0.38	162
Case 3	6.33	30.27	1.29	179
Case 4	8.55	30.21	1.52	161
Case 5	9.19	30.08	1.93	179

4.4.3 Impact of τ on System Losses

As discussed in Section 4.2, the value of parameter τ in constraint (4.1j) imposes constraints on whether the D-FACTS devices can operate in an idle state. Simulations are further conducted to study the impact of τ on system losses in the IEEE 118-bus system. Figure 4.2 shows the system loss under the proposed MTD planning with τ varying from 0 to 15%. As seen, the system loss slightly increases with an increase in τ. Nonetheless, such an increase is trivial. For instance, the system loss at $\tau = 10\%$ only increases by 0.02% compared with that at $\tau = 0$. Compared with the system loss under other MTD plannings at $\tau = 0$ in Table 4.2, the system loss under the max-rank planning at $\tau = 0.10$ is lower than that in Cases 1 and 2, and slightly higher than that in Cases 4 and 5. From the perspective of the MTD effectiveness, each idle-state D-FACTS device under the proposed MTD planning will decrease the rank of the composite matrix by one. By setting $\tau > 0$, it ensures the

maximum MTD effectiveness with a slight increase in the system loss, whereas setting $\tau = 0$ could further reduce the system loss, but may decrease the MTD effectiveness. The results here show that the proposed MTD planning and operational approaches can yield a good trade-off between the system loss and the MTD effectiveness.

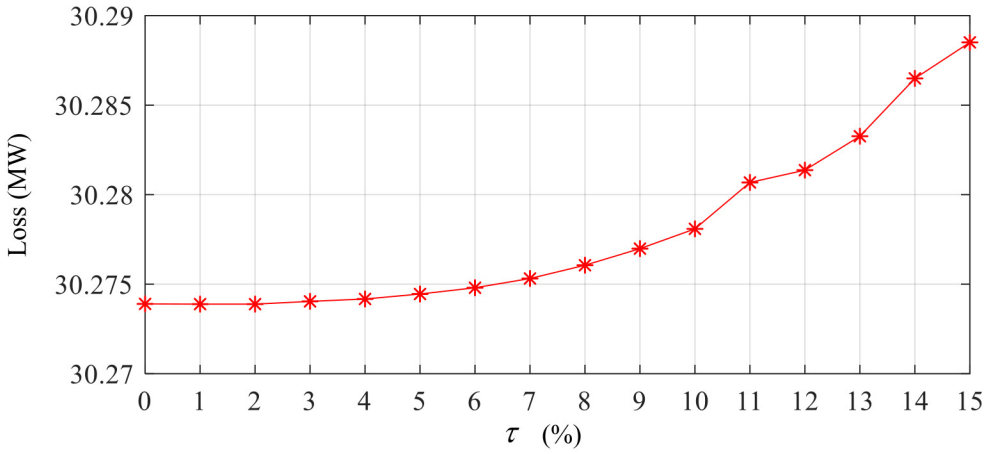


Figure 4.2: *System losses versus* τ.

4.5 Summary

We propose an ACOPF model considering D-FACTS devices as an MTD operating model, in which generation costs and system losses are minimized. The proposed model introduces the reactance of lines equipped with D-FACTS devices as the decision variables, which provides the ACOPF model with the extra capability to manage power flow. The proposed model can be seamlessly integrated into the existing energy management system of a power system in the control room. System operators can apply the proposed ACOPF model to manage the real-time system operation and determine the setpoints of D-FACTS devices. Furthermore, the setpoints of D-FACTS devices can be adopted by MTD to safeguard a cyber-secure power system. Therefore, economic benefits and additional cyber-defense benefits can be simultaneously obtained from D-FACTS devices.

In order to efficiently solve the ACOPF-based MTD model, we build an interior-point solver of the proposed ACOPF. We derive the gradient and Hessian matrices of the objective

function and constraints with respect to the line reactance. As the derivations adopt the same voltage coordinate and complex power expression as MIPS in MATPOWER, the derived gradient and Hessian matrices can be simply integrated into MIPS. The case study compares the proposed ACOPF with the conventional ACOPF regarding generation costs, system losses, and CPU time under different power flow limits. The results show that D-FACTS devices can effectively reduce the system loss, and the CPU time of solving the proposed ACOPF is generally less than 15 seconds in the IEEE 118-bus system. The results verify the effectiveness of the proposed ACOPF model and the interior-point solver.

We further investigate the impact of the proposed MTD planning on the system loss. The proposed MTD planning and operational approaches present a trade-off between the system loss and the MTD effectiveness. Numerical results show that, at a slight increase in the system loss, these new approaches ensure the maximum rank of the composite matrix while using the minimum number of D-FACTS devices, which in turn reduce the cost of deploying D-FACTS devices.

Chapter 5

Hidden MTD Planning and Operation Algorithm

In previous sections, the proposed moving target defense (MTD) planning algorithms and MTD operation models maximize the MTD detection effectiveness and minimize the MTD application costs. However, MTD hiddenness is not considered in the previous methods. This chapter derives a novel hidden MTD (HMTD) operation condition, proposes a hidden MTD planning method, and hidden MTD operation models in both the DC and AC models.

5.1 Introduction

A random MTD (RMTD) approach was proposed in reference[84], in which the reactance of distributed flexible AC transmission system (D-FACTS) equipped lines was randomly changed without considering the detection effectiveness. However, one inherent drawback of the RMTD is that a strong adversary can easily detect whether an MTD is in place by eavesdropping measurements. As shown in the dotted red box of Figure 5.1, an alert and sophisticated attacker can detect the existence of MTDs, if the attacker conducts the residual-based bad data detection (BDD) based on the eavesdropped measurements and his knowledge about the system parameters, i.e., the original line susceptance b_0. The detection

of the existence of MTD can drive the attacker to postpone the planned attacks, invest more resources to gain updated system knowledge through topology learning (TL), and potentially intrude into more critical parts. Consequently, a power grid may face a higher level of cyber threats. To overcome this drawback, hidden MTD (HMTD) operation approaches were initially presented in the transmission system[72] and the distribution system[91], in which setpoints of the D-FACTS devices were delicately changed to make system measurements unchanged after the HMTD.

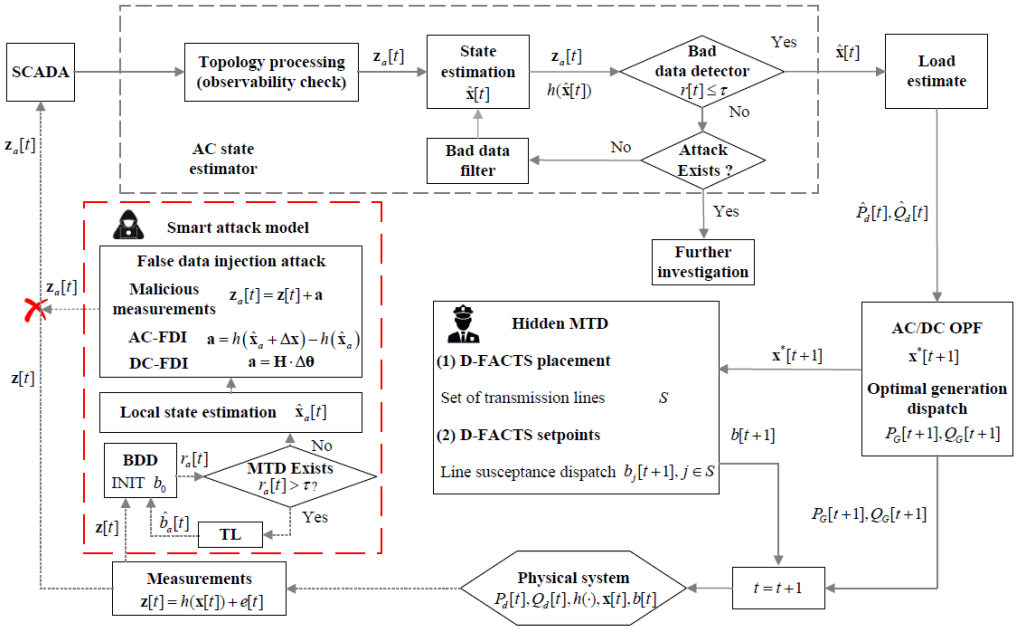


Figure 5.1: *Illustration of HMTD and smart attack model in the smart grid.*

In the construction of an HMTD, the MTD hiddenness and detection effectiveness are two primary objectives that are closely related and mostly conflicting. Specifically, the hiddenness is not achievable in a system with the highest detection effectiveness, i.e., a complete MTD system, which can detect all FDI attacks[72]. On the other hand, while incomplete MTD systems have limited detection effectiveness, their incompleteness provides viability for the MTD hiddenness. Fortunately, HMTDs can be constructed in the majority of power systems since most systems belong to incomplete MTD systems owning to the restrictive requirements of a complete MTD[70;79]. It is worth noting that some HMTDs are ineffective in detecting FDI attacks, even though they are hidden to attackers[72;86]. Consequently, the main concern

in the construction of HMTDs becomes how to maximize detection effectiveness.

D-FACTS placement in the context of MTDs has been recently studied to improve the detecting effectiveness. In Chapter 3, we proved that the rank of the composite matrix, i.e., one metric on the detection effectiveness, could be determined by the D-FACTS placement regardless of the D-FACTS setpoints. Max-rank planning algorithms were proposed in Chapter 3 to achieve the maximum rank of the composite matrix using the minimum number of D-FACTS devices. Zhang et al.[88] proposed a heuristic-based D-FACTS placement algorithm to maximize the rank of the composite matrix and cover the largest number of buses. Tian et al.[72] showed that the rank of the composite matrix in HMTD is related to D-FACTS placements, but no solution was further proposed to construct an HMTD with the maximum rank of the composite matrix. Zhang et al.[86] proposed a joint HMTD algorithm by combining D-FACTS placement with protected meters placement. More specifically, the joint algorithm places a protected meter in each loop to achieve an HMTD with the maximum rank of the composite matrix. They concluded that an MTD is hidden only if the reactance of branches in a loop is modified by a unity factor. However, this is an overly strong condition for an HMTD. In this chapter, we will show, for the first time, that HMTDs are achievable and their detection effectiveness is guaranteed without using a unity factor or protected meters.

Towards practical applications of HMTD, the optimal planning and operation of D-FACTS devices that ensure the MTD hiddenness and maximal detection effectiveness are challenging yet unresolved issues. In this chapter, we aim at addressing these two intertwined issues by establishing a systematic planning and operation approach for HMTDs. In the planning stage, our objective is to identify a D-FACTS placement, which ensures HMTDs can always be constructed under different load conditions and D-FACTS setpoints. During the operation stage, our objective is to achieve the hiddenness operation condition efficiently. Additionally, the proposed planning and operation together ought to guarantee the maximum detection effectiveness of HMTDs and reduce the MTD operation cost.

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

1) The hidden MTD operation condition in reference[72] is difficult to be used in the construction of HMTDs. Can we derive a novel hidden MTD operation condition that can

be easily integrated into HMTD operation methods?

2) How can system operators guarantee the existence of HMTD in a given system? How can MTD planning guarantee the MTD hiddenness and MTD detection effectiveness at the same time? Are HMTDs achievable without using a unity factor or protected meters?

3) The existing DC-HMTD operation[72] is based on random weight searching. What are the drawbacks of this DC-HMTD operation? How can the proposed operation model overcome these drawbacks? Specifically, how can the proposed operation model obtain the D-FACTS setpoints in HMTD more efficiently and avoid the D-FACTS devices working in the idle state?

4) How can system operators construct an HMTD in the AC model? How can AC-HMTD operation model explore the economic benefit from D-FACTS devices?

5) Does HMTD need more generation costs compared with that in the system without MTD, in the RMTD, and in the OPF-based MTD? Can we present a cost-benefit analysis of HMTD in both the DC and AC models?

6) Is the proposed HMTD planning better than other planning methods regarding the detection effectiveness and the MTD hiddenness? Does the proposed DC-HMTD operation model outperform the existing HMTD operation method[72]? Is there any trade-off between the MTD hiddenness and MTD detection effectiveness?

Question 1 is solved in Section 5.3, Question 2 is answered in Section 5.4, Questions 3, 4, and 5 are answered in Section 5.5, and the answer for Question 6 is presented in Section 5.6.

5.2 Preliminaries

5.2.1 MTD Hiddenness and Detection Effectiveness

Encrypted commands from the system operator's control room can be securely transmitted to the D-FACTS gateway for changing the setpoint in D-FACTS devices through DNP3, IEC 61850, and 60870-5-104 protocol[81]. MTD takes advantage of D-FACTS devices to

create uncertainties for attackers. The incremental reactance of line i–j can be periodically modified by the D-FACTS device within $|x_{ij} - x_{ij}^0| \leq |\eta x_{ij}^0|$, where the upper bound $\eta = 20\%$, generally referred to as the MTD magnitude, reflects the physical capacity of D-FACTS devices[70;72;79]. Consequently, the measurement matrix $\mathbf{H}$ used in the state estimation (SE) becomes a time-variant matrix. If attackers construct FDI attacks based on an outdated knowledge of $\mathbf{H}$, the estimation residual in the defender's BDD is no longer zero.

HMTD, a superior MTD, is stealthy to alert attackers who use the well-known, residual-based BDD to detect the existence of MTD[72;86;91]. The key idea of HMTD is to create little to no changes in system measurements after HMTD is applied, i.e., $\mathbf{H}_0\theta_0 = \mathbf{H}\theta$, such that the estimation residual in the attacker's BDD verification remains the same after HMTD. The defense stealthiness probability (DSP) is a widely used metric to quantify the MTD hiddenness from the perspective of attackers, which is defined as:

$$DSP = \frac{\text{Number of MTDs hidden to attackers}}{\text{Total number of MTDs}}$$

Since HMTD cannot be constructed in all power systems[72], whether or not the hiddenness of an MTD can be attained in a particular power system becomes a primary concern. A sufficient and necessary condition for the existence of HMTD was proposed in reference[72]. Let us denote the immutable part of $\mathbf{H}$ by $\widetilde{\mathbf{H}}$, consisting of the $\mathbf{H}$'s rows corresponding to all non-D-FACTS lines. An HMTD exists if and only if $r(\widetilde{\mathbf{H}}) < r(\mathbf{H})$[72]. This condition is beneficial for checking the existence of HMTD when the locations and setpoints of D-FACTS devices are all given. Nevertheless, this condition provides no guidance on how to optimally place and operate D-FACTS devices in a particular system. This chapter bridges this gap by deriving the existence requirements of HMTD for the D-FACTS placement and operation.

Once the existence of HMTDs is guaranteed, the detection effectiveness of HMTDs becomes a prominent concern. The rank of the composite matrix $\mathbf{M} = [\mathbf{H}_0\ \mathbf{H}]$ is a good metric to quantify the MTD detection effectiveness[79]. The HMTD with the maximum rank of the composite matrix, which is referred to as a max-rank MTD, is desirable. Besides, the attack detection probability (ADP), another widely utilized metric to measure the detection

effectiveness of an MTD, is defined as:

$$ADP = \frac{\text{Number of FDIs detected by the MTD}}{\text{Total number of FDIs}}$$

A max-rank MTD with both high ADP and high DSP is a more desired MTD. An MTD with high ADP and low DSP is good at detecting FDI attacks but can be easily detected by alert attackers. An MTD with low ADP is least desirable as detection capability is the primary concern of an MTD.

5.2.2 Max-rank MTD Planning Algorithms

The power system topology can be treated as an edge-weighted graph with buses as nodes and lines as edges. The rank of the composite matrix of an MTD is determined by the MTD planning[70]. We summarize below the relation between the MTD planning and the rank of the composite matrix. Suppose all D-FACTS devices work in non-idle (compensating) states and G_{DF} is loopless, the rank of the composite matrix in MTDs is determined by the number of loops in $G_{\overline{DF}}$ as follows:

$$r(\mathbf{M}) = p - lp_{\overline{DF}} \tag{5.1}$$

where $lp_{\overline{DF}}$ is the number of loops in $G_{\overline{DF}}$. Note that equation (5.1) does not hold if there exists any loop in G_{DF} and each loop in $G_{\overline{DF}}$ decreases $r(\mathbf{M})$ by one. Thus, an MTD is a max-rank MTD if the MTD planning ensures either G_{DF} or $G_{\overline{DF}}$ is loopless[70]. It is worth mentioning that neither the number of connected components in G_{DF} nor that in $G_{\overline{DF}}$ influence the rank of the composite matrix. In this chapter, we utilize the connected components to construct HMTDs and apply the relationship (5.1) to achieve a max-rank MTD.

5.3 Novel HMTD Operation Condition

In DC-SE, the objective of HMTD is to remain all measurements on active power flow and active power injection unchanged after the setpoint changes of D-FACTS devices. After the control signal is sent to D-FACTS devices from the control room, the line reactance can be changed within seconds. During the activation of the MTD, the change in nodal active power injection measurements is minute. Thus, it is reasonable to assume that the system loads are constant during the activation of the MTD for analysis. This assumption was also made in other HMTD analyses[72;86]. In cases that this assumption does not hold, the influence of variant loads on the hiddenness of the proposed HMTD will be evaluated in the subsequent case studies of this chapter. We focus on the power flow measurements in the HMTD model to facilitate the analysis.

Suppose power flow measurements are arranged in the following order, i.e., $\mathbf{z} = [\ \overline{\mathbf{z}}^T \quad \tilde{\mathbf{z}}^T\ ]^T$, where $\overline{\mathbf{z}}$ and $\tilde{\mathbf{z}}$ are power flow measurements of the D-FACTS lines and non-D-FACTS lines, respectively. Accordingly, the measurement matrices before and after the HMTD in the DC model are expressed as $\mathbf{H}_0 = \begin{bmatrix} \overline{\mathbf{H}}_0 \\ \widetilde{\mathbf{H}} \end{bmatrix}$ and $\mathbf{H} = \begin{bmatrix} \overline{\mathbf{H}} \\ \widetilde{\mathbf{H}} \end{bmatrix}$, where $\overline{\mathbf{H}}_0$ and $\overline{\mathbf{H}}$ are the submatrices of the measurement matrix and represent power flow measurements of the D-FACTS lines before and after the HMTD, respectively; and $\widetilde{\mathbf{H}}$ corresponds to power flow measurements of non-D-FACTS lines. Thus, the power flow measurements before and after the HMTD are $\mathbf{z}_0 = \begin{bmatrix} \overline{\mathbf{H}}_0 \\ \widetilde{\mathbf{H}} \end{bmatrix} \theta_0 + \mathbf{e}$ and $\mathbf{z} = \begin{bmatrix} \overline{\mathbf{H}} \\ \widetilde{\mathbf{H}} \end{bmatrix} (\theta_0 + \Delta\theta) + \mathbf{e}$, respectively, where $\Delta\theta$ is the incremental voltage angle introduced by the HMTD. Since all measurements remain unchanged after the HMTD, i.e., $\mathbf{z}_0 = \mathbf{z}$, we can derive the hiddenness condition in the noiseless condition as follows:

$$\overline{\mathbf{H}}_0\theta_0 = \overline{\mathbf{H}}(\theta_0 + \Delta\theta) \tag{5.2}$$

$$\widetilde{\mathbf{H}}\Delta\theta = 0 \tag{5.3}$$

As $\widetilde{\mathbf{H}}$ is a fixed matrix, equation (5.3) indicates that $\Delta\theta$ determined by the system operator (defender) must belong to the null space of $\widetilde{\mathbf{H}}$:

$$\Delta\theta = \mathbf{U}\mathbf{W} \tag{5.4}$$

where $\mathbf{U} = [u_1, u_2, ..., u_s] \in \mathbb{R}^{p_1 \times s}$ is the matrix of kernel bases of $\widetilde{\mathbf{H}}$; $\mathbf{W} = [w_1, w_2, ..., w_s]^T \in \mathbb{R}^{s \times 1}$ is the weight determined by the system operator; and s is the dimension of kernel bases. In addition, D-FACTS setpoints ought to be delicately chosen to make $\overline{\mathbf{H}}$ satisfying equation (5.2). The hiddenness condition demonstrates that the D-FACTS setpoints are closely related to incremental voltage angle.

5.4 HMTD Planning Algorithm

5.4.1 Requirements of MTD Planning for HMTD

We utilize the topology analysis to derive a sufficient condition for D-FACTS placement to achieve the MTD hiddenness. The decomposition of $\mathbf{H}$ in reference[79] can be applied on $\overline{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$, respectively, as follows:

$$\overline{\mathbf{H}} = \mathbf{D}_1 \cdot \mathbf{X}_1 \cdot \mathbf{A}_1^T \tag{5.5}$$

$$\widetilde{\mathbf{H}} = \mathbf{D}_2 \cdot \mathbf{X}_2 \cdot \mathbf{A}_2^T \tag{5.6}$$

where $\mathbf{X}_1 \in \mathbb{R}^{p_1 \times p_1}$ and $\mathbf{X}_2 \in \mathbb{R}^{p_2 \times p_2}$ are the diagonal reactance matrix of the D-FACTS lines and non-D-FACTS lines, respectively; $\mathbf{A}_1 \in \mathbb{R}^{n-1 \times p_1}$ and $\mathbf{A}_2 \in \mathbb{R}^{n-1 \times p_2}$ are the reduced bus-branch incidence matrix of G_{DF} and $G_{\overline{DF}}$, respectively, in which the row of the slack bus is removed; In a power flow fully measured power system, $\mathbf{D}_1$ and $\mathbf{D}_2$ are of full column rank since $\mathbf{D}_1 = [\mathbf{I}_1 \quad -\mathbf{I}_1]^T$ and $\mathbf{D}_2 = [\mathbf{I}_2 \quad -\mathbf{I}_2]^T$, where $\mathbf{I}_1 \in \mathbb{R}^{p_1 \times p_1}$ and $\mathbf{I}_2 \in \mathbb{R}^{p_2 \times p_2}$ are identity matrices. Note that if the system is not power flow fully measured, the decomposition in equations (5.5) and (5.6) become $\overline{\mathbf{H}} = \mathbf{C}_1 \cdot \mathbf{D}_1 \cdot \mathbf{X}_1 \cdot \mathbf{A}_1^T$ and $\widetilde{\mathbf{H}} = \mathbf{C}_2 \cdot \mathbf{D}_2 \cdot \mathbf{X}_2 \cdot \mathbf{A}_2^T$, where $\mathbf{C}_1$ and $\mathbf{C}_2$ are meter selection matrices defined in reference[79]. As long as $\mathbf{C}_1 \cdot \mathbf{D}_1$ and $\mathbf{C}_2 \cdot \mathbf{D}_2$ are of full column rank, the conclusions in a fully measured system can be extended to a partially measured system. More specifically, $\mathbf{C}_1 \cdot \mathbf{D}_1$ with a full column rank indicates the power flow of each D-FACTS line is at least measured either at the from-bus or the to-bus. The same requirements apply for non-D-FACTS lines to achieve the full column rank of

$\mathbf{C}_2 \cdot \mathbf{D}_2$.

A sufficient condition for the existence of HMTD is given by the following lemma from the perspective of D-FACTS placement.

Lemma 5.4.1. an HMTD exists if no D-FACTS devices work in the idle state and $G_{\overline{DF}}$ is a disconnected graph, i.e., $t_{\overline{DF}} > 1$.

Proof: Since $\mathbf{D}_2$ and $\mathbf{X}_2$ are of full column rank, the rank of $\widetilde{\mathbf{H}}$ equals to $r(\mathbf{A}_2)$, i.e., $r(\widetilde{\mathbf{H}}) = r(\mathbf{A}_2)$. According to graph theory, the rank of incidence matrix $\mathbf{A}$ in a planar graph with n nodes and t components is $n - t$, i.e., $r(\mathbf{A}) = n - t$[95]. Thus, in $G_{\overline{DF}}$, the following equation holds:

$$r(\widetilde{\mathbf{H}}) = rank(\mathbf{D}_2 \cdot \mathbf{X}_2 \cdot \mathbf{A}_2^T) = rank(\mathbf{A}_2^T) = n - t_{\overline{DF}} \tag{5.7}$$

Suppose $G_{\overline{DF}}$ is a disconnected graph, i.e., $t_{\overline{DF}} > 1$, $r(\widetilde{\mathbf{H}}) < n - 1$ holds. Thus, $r(\widetilde{\mathbf{H}}) < r(\mathbf{H}) = n - 1$ holds. Thus, an HMTD exists according to the sufficient and necessary condition of HMTD mentioned in Section 5.2.1. ■

As mentioned earlier, not all HMTDs are effective in detecting FDI attacks. The hiddenness and the detection effectiveness must be simultaneously considered in the D-FACTS placement. To ensure the detection effectiveness, HMTDs constructed on a D-FACTS placement ought to have the maximum rank of the composite matrix. Here, we propose Lemma 5.4.2 to present the requirements of D-FACTS placement for constructing max-rank HMTDs.

Lemma 5.4.2. a max-rank HMTD exists if the following conditions are satisfied: 1) all D-FACTS devices work in the non-idle states; 2) G_{DF} is loopless; and 3) $G_{\overline{DF}}$ is a disconnected loopless graph.

Proof: According to equation (5.1), if both G_{DF} and $G_{\overline{DF}}$ are loopless, any MTD under this topology is a max-rank MTD, i.e., $r(M) = p$[70]. According to Lemma 5.4.1, if $G_{\overline{DF}}$ is a disconnected and loopless graph, an HMTD exists. Therefore, a max-rank HMTD exists in the D-FACTS placement. ■

In addition to maximizing the rank of the composite matrix, it is important to cover all

necessary buses using D-FACTS lines concluded in Section 3.3. If a bus is not in any loop, an FDI attack on this bus is undetectable regardless of D-FACTS placement and setpoints[86]. Thus, there is no need to cover these buses, whereas the other buses need to be covered.

As no D-FACTS devices work in the idle state is the prerequisite of Lemma 5.4.2, it is necessary to consider this MTD operation requirement during the MTD planning. According to the hiddenness condition (5.2), setpoints of D-FACTS devices are closely related to the nodal incremental voltage angle in HMTD. Understanding how voltage angles change is the key to constructed HMTD. Here, we propose Lemma 5.4.3 to demonstrate how the nodal incremental voltage angle is related to the MTD planning.

Lemma 5.4.3. In an HMTD, all buses in a connected component in $G_{\overline{DF}}$ must have the same nodal incremental voltage angle.

Proof: Assume Bus i and Bus j are two neighbor nodes in the same connected component in $G_{\overline{DF}}$, and their voltage angle before the HMTD is θ_i and θ_j, respectively. Before the HMTD, the power flow on branch i-j is $p_{ij}^0 = (\theta_i - \theta_j)/x_{ij}$, where x_{ij} is the reactance of branch i-j. Note that x_{ij} cannot be modified by the D-FACTS device, as branch i-j is a non-D-FACTS line. Assume Buses i and j have different incremental voltage angles after the HMTD, i.e., $\Delta\theta_i$ and $\Delta\theta_j$ ($\Delta\theta_i \neq \Delta\theta_j$). The power flow on branch i-j becomes $p_{ij} = (\theta_i + \Delta\theta_i - \theta_j - \Delta\theta_j)/x_{ij}$ after the HMTD. It is obvious that $p_{ij}^0 \neq p_{ij}$, which conflicts with the fact that power flow remains the same before and after the HMTD. Therefore, any pair of neighbor nodes in $G_{\overline{DF}}$ has the same nodal incremental voltage angle in an HMTD. It infers all nodes in the same connected component have the same nodal incremental voltage angle in the HMTD. ■

We can further explain the HMTD operation characteristics by combining Lemma 5.4.3 and equation (5.4). The i^{th} kernel base in equation (5.4), i.e., i^{th} column in $\mathbf{U}$, identifies all buses in the i^{th} connected component in $G_{\overline{DF}}$. Weight w_i in equation (5.4) indicates that all buses in the i^{th} connected component have the same incremental voltage angle, which is equal tow_i. For example, in Figure 5.2(a), Buses 1 and 6 are in the same connected component in the $G_{\overline{DF}}$, and they need to have the same incremental voltage angle in an HMTD. Let an isolated node refers to a bus whose branches are all D-FACTS lines. For

each isolated node in $G_{\overline{DF}}$, it has its own kernel base and weight. For example, in Figure 5.2(a), Buses 2 and 5 are two isolated nodes in the $G_{\overline{DF}}$.

We further propose Corollaries 5.4.1 and 5.4.2 to identify two special cases in which D-FACTS devices must work in the idle state.

Corollary 5.4.1. In an HMTD, if a D-FACTS line's two nodes belong to the same connected component in $G_{\overline{DF}}$, the D-FACTS device associated with this line must work in the idle state.

Corollary 5.4.2. In an HMTD, if a D-FACTS line's two nodes are two isolated nodes in $G_{\overline{DF}}$ and have the same incremental nodal voltage angle, i.e., $w_i = w_j$, the D-FACTS device associated with this line must work in the idle state.

Note that Corollary 5.4.1 is in the context of the HMTD planning, whereas Corollary 5.4.2 is on the HMTD operation. Based on the above theoretical foundations, the requirements of MTD planning to construct a max-rank HMTD covering all necessary buses is summarized as follows: 1) $G_{\overline{DF}}$ is a disconnected and loopless graph; 2) G_{DF} is a loopless graph, and its links should cover all buses except for the buses not in any loops; and 3) no D-FACTS devices should work in the idle state. In the following subsection, we will design D-FACTS placement rules and an algorithm to achieve these requirements.

5.4.2 Hidden MTD Planning Algorithm

We design the following D-FACTS placement rules in each loop of power system topology. Rule 1 is proposed for two purposes. Firstly, it can effectively prevent D-FACTS devices from working in the idle state identified in Corollary 5.4.1. In a loop (with more than two links), if two end-nodes of a D-FACTS line belong to the same connected component in $G_{\overline{DF}}$, there must be at least two successive non-D-FACTS lines in the loop, which is forbidden according to Rule 1. Secondly, it makes D-FACTS lines to cover all nodes in the loop. Since the degree of each node in the loop is no less than two, any end-node of a non-D-FACTS line has to connect to another D-FACTS line due to Rule 1. By extending Rule 1 from a single loop to all loops in the entire system, all buses in all loops of the system are covered by D-FACTS

lines.

Rule 1. In each loop of system topology, two or more than two successive non-D-FACTS lines are not allowed.

Further, we design Rule 2 to avoid the appearance of idle D-FACTS devices identified in Corollary 5.4.2. This is because that if three or more than three successive D-FACTS lines may generate two or more isolated nodes in $G_{\overline{DF}}$. As two successive D-FACTS lines generate no more than one isolated node in $G_{\overline{DF}}$, the scenario described in Corollary 5.4.2 is excluded in the MTD planning.

Rule 2. In each loop of the system, more than two successive D-FACTS lines are not allowed.

Note that Rules 1 and 2 propose requirements on the topology of non-D-FACTS and D-FACTS lines in each loop, respectively. Thus, they provide essential guidance on the MTD planning. We take a loop with six transmission lines as an example. Figure 5.2 demonstrates all five feasible solutions subject to Rules 1 and 2. It is seen that all these solutions effectively cover all buses in the loop and avoid idle D-FACTS devices identified in Corollaries 5.4.1 and 5.4.2.

Based on these two rules, we propose a depth-first-search (DFS)-enabled, hidden MTD planning algorithm, which is illustrated in Algorithms 4 and 5. Since the existence of HMTD directly relies on the topology of $G_{\overline{DF}}$ as demonstrated in Lemma 5.4.1, here we focus on finding the location of non-D-FACTS lines. In Algorithm 4, we initialize all lines as D-FACTS lines by using the system graph G as G_{DF} and utilize the DFS algorithm to place non-D-FACTS lines. Note that we use set E_{DF} and E_{NDF} to store the D-FACTS and non-D-FACTS lines determined by DFS, respectively.

The proposed DFS algorithm (Algorithm 5) traverses all loops in the order of a stack (first-in, last-out) based on recursion. In each iteration, we first check whether the following stopping criteria are simultaneously met: 1) G_{DF} is are loopless; 2) $G_{\overline{DF}}$ is a loopless and disconnected graph, and 3) the placement satisfies Rules 1 and 2. If they are satisfied, the

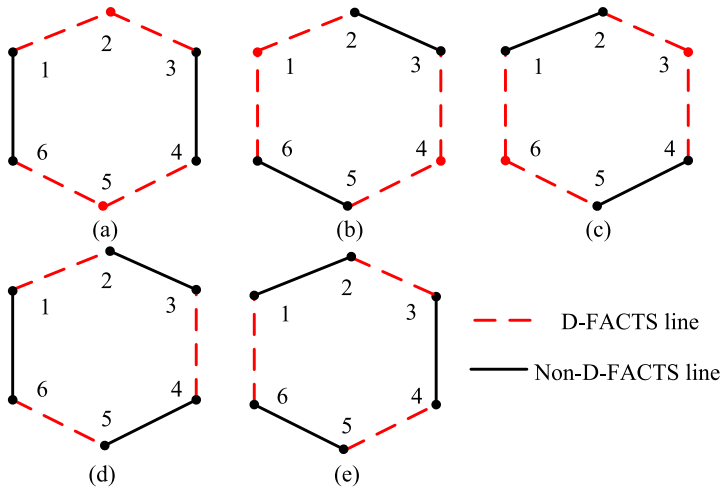


Figure 5.2: *An illustration of D-FACTS placement solution in HMTD.*

algorithm returns the D-FACTS placement solution and stops searching. Otherwise, the algorithm continues to deal with the next loop in G_{DF}, where the D-FACTS lines already placed are identified. Then, all solutions in the loop subject to Rules 1 and 2 are found and saved in a set called *solutionsInSingleLoop*. For each of the stored solutions, we make the recursive call to search in the next loop by updating the latest D-FACTS and non-D-FACTS lines in the system. Algorithm 4 stops after finding a feasible solution or traversing all loops.

Algorithm 4: Hidden MTD Planning Algorithm

Input: The edge-weighted graph $G(V, E)$ of a power grid topology
Output: *Results:* set of non-D-FACTS line placement solution

1: **Initialization:** Suppose all lines are D-FACTS lines, i.e., G_{DF}= G. $E_{DF} = \emptyset$ $E_{NDF} = \emptyset$, Global variable *Results* $= \emptyset$
2: dfs(G_{DF}, E_{DF}, E_{NDF}, V)
3: **return** *Results*

When a system operator has a constrained budget for D-FACTS devices, the number of D-FACTS devices can be reduced by removing D-FACTS devices from the hidden MTD planning solution until the budget is met. Additionally, if D-FACTS devices are also used to minimize the power losses in the system operation, D-FACTS devices on lines with lower power loss to impedance sensitivity (PLIS) are suggested to be removed. However, one must take into consideration the impact of removing D-FACTS devices on the MTD hiddenness

Algorithm 5: Depth-First Search (DFS)

Input: D-FACTS graph G_{DF}, set of placed D-FACTS line E_{DF}, set of placed non-D-FACTS line E_{NDF}, System buses V

dfs(G_{DF}, E_{DF}, E_{NDF}, V)
Generate a non-D-FACTS graph G_{NDF} composed of E_{NDF} and V
if (G_{DF} has no loops and G_{NDF} is a disconnected graph without loops && placement subjects to Rules 1 and 2)
    Add E_{NDF} to *Results*
    return
end if
Select a loop in G_{DF} and add all lines in the loop to a set, denoted by E_i
$E_{DF0} = E_i \cap E_{DF}$ // DF lines already placed in the loop
Find all feasible solutions in the loop subject to Rules 1 and 2, and save in *solutionsInSingleLoop*
if (*PlaceSet* == $\emptyset$)
    return
end if
for each feasible *placement* solution in *solutionsInSingleLoop*
    ΔE_{DF} = DF lines in *placement*$-E_{DF0}$ // new DF lines placed
    ΔE_{NDF} =NDF lines in *placement* // new NDF lines placed
    $G_{DF1} = G_{DF}$, and update G_{DF1} by removing new NDF lines ΔE_{NDF}
    dfs(G_{DF1}, $E_{DF} + \Delta E_{DF}$, $E_{NDF} + \Delta E_{NDF}$, V)
end for
end dfs

and detection effectiveness. Hiddenness can still exist after removing D-FACTS devices as long as $t_{\overline{DF}} > 1$ according to Lemma 5.4.1. This is because the removal of D-FACTS devices, equivalent to adding links to $G_{\overline{DF}}$, doesn't necessarily reduce $t_{\overline{DF}}$ to one. However, the MTD planning ought to simultaneously guarantee the max rank of the composite matrix and cover all buses in loops to achieve the maximal detection effectiveness. Removing D-FACTS devices can result in uncovered buses and forming loops in $G_{\overline{DF}}$. Consequently, the rank of the composite matrix will decrease by the number of loops in $G_{\overline{DF}}$ and MTD cannot detect FDI attacks on uncovered buses.

The differences between the proposed hidden MTD planning and the max-rank planning established in Chapter 3 are summarized into the following two aspects. From the aspect of hiddenness, the proposed hidden MTD planning requires the number of connected components in a non-D-FACTS graph to be greater than one to guarantee the existence of HMTD,

whereas the max-rank planning has no such requirement. From the aspect of detection capability, both MTD planning methods ensure the max-rank MTDs. The max-rank planning focuses on using the minimum number of D-FACTS devices such that certain buses may thus be uncovered. In contrast, the proposed hidden MTD planning in this chapter places D-FACTS lines and non-D-FACTS lines alternately in each loop such that all buses in loops are covered, contributing to much improved detection effectiveness.

In summary, the proposed hidden MTD planning algorithm ensures 1) the maximum rank of the composite matrix; 2) the coverage of all necessary buses; and 3) the existence of the HMTD. To further achieve the HMTD with maximal detection effectiveness, we propose an HMTD operation model to determine setpoints of D-FACTS devices in the following subsections.

5.5 HMTD Operation Model

5.5.1 DC-HMTD Operation Model

The non-idle setpoints of D-FACTS devices ought to be delicately chosen in the HMTD operation. We propose a non-convex, nonlinear, optimization-based DC-HMTD operation model in (5.8), which maximizes the susceptance changes of D-FACTS lines and utilizes the hiddenness condition as constraints.

$$\max_{\mathbf{b},\mathbf{W}} \|\mathbf{b}-\mathbf{b}_0\|_2 \tag{5.8}$$

$$s.t. \qquad \overline{\mathbf{H}}_0\hat{\theta}_0 = \overline{\mathbf{H}}(\mathbf{b})(\hat{\theta}_0+\Delta\theta) \tag{5.8a}$$

$$\Delta\theta = \mathbf{UW} \tag{5.8b}$$

$$\mathbf{b}_0^{\min} \le \mathbf{b} \le \mathbf{b}_0^{\max} \tag{5.8c}$$

where $\mathbf{b}$ is the susceptance of each D-FACTS line, which is the reciprocal of reactance x; $\mathbf{b}_0^{\min}$ and $\mathbf{b}_0^{\max}$ are the vector of lower and upper bound of susceptance for D-FACTS lines

due to the physical capacity of D-FACTS devices; $\mathbf{W} = [w_1, w_2, ..., w_s]^T \in \mathbb{R}^{s\times 1}$ is the vector of voltage angle incremental in each connected component of $G_{\overline{DF}}$. Note that we replace θ_0 in (5.2) with estimated nodal voltage angle $\hat{\theta}_0$ in SE as θ_0 is unknown to system operators. Constraint (5.8a) aims to remain measurements of non-D-FACTS lines unchanged in HMTD. As measurements contain noises, we use the estimated measurements $\overline{\mathbf{H}}_0\hat{\theta}_0$ instead to reduce the impact of noise. If significant measurement errors occur, BDD can detect and identify the erroneous measurements before running the proposed HMTD operation model.

The proposed model can be seamlessly integrated into the existing energy management system (EMS) in the system control room. Specifically, after determining the optimal generation and power flow using DC optimal power flow (OPF), the system operator can calculate the setpoints of the D-FACTS devices by solving model (5.8), and then send the calculated setpoints to the field devices for implementation. Note that while model (5.8) retains the power flow unchanged, it maximizes the susceptance changes for two purposes: 1) further deviating D-FACTS devices from their idle states; and 2) allowing sufficient changes to accommodate measurement noises. The proposed method is more robust and efficient in calculating the setpoints of D-FACTS devices due to its optimization-based model, compared with the random-weight-based HMTD method[72], referred to as RW-HMTD hereafter.

5.5.2 AC-HMTD Operation Model

In the construction of an AC-HMTD, a set of system measurements before HMTD is needed as a reference. This reference operating point is usually obtained by running ACOPF before HMTD, where the system operation cost and/or system losses are minimized. In a transmission system, it is reasonable to assume the voltage magnitude of each bus and the active power flow of each transmission line are measured in AC-SE[72]. An AC-HMTD operation model needs to reduce the measurement changes as much as possible to achieve MTD hiddenness. Additionally, this model ought to consider the economic benefits of D-FACTS devices in the power system operation[96]. Therefore, we propose an ACOPF-based HMTD operation model in (5.9), in which the reactance of each D-FACTS line is introduced as a decision

variable in the traditional ACOPF. The proposed model minimizes a weighted sum of 1) the generation cost; 2) the negative of reactance changes, which is consistent with the proposed DC-HMTD model; and 3) the normalized difference in active power measurements before and after HMTD by relaxing the AC counterpart of the DC hiddenness equality constraint (5.8a).

$$\min_{\mathbf{Y}} \ \lambda_0 cost(\mathbf{Y}) + \lambda_1 distP(\mathbf{Y}) - \lambda_2 distX(\mathbf{x}) \tag{5.9}$$

$$s.t. \quad cost(\mathbf{Y}) = \sum_{i=1}^{n_g} f^i(p_g^i) \tag{5.9a}$$

$$distP(\mathbf{Y}) = \sum_{i=1}^{n_l} (P_i^f - P_{i,0}^f)^2 / {P_{i,0}^f}^2 \tag{5.9b}$$

$$distX(\mathbf{x}) = \sum_{i \in E_{DF}} (x_i - x_i^0)^2 \tag{5.9c}$$

$$g_P(\theta, \mathbf{V}, \mathbf{P_g}, \mathbf{x}) = 0 \tag{5.9d}$$

$$g_Q(\theta, \mathbf{V}, \mathbf{P_g}, \mathbf{x}) = 0 \tag{5.9e}$$

$$h_f(\theta, \mathbf{V}, \mathbf{x}) \leq 0 \tag{5.9f}$$

$$h_t(\theta, \mathbf{V}, \mathbf{x}) \leq 0 \tag{5.9g}$$

$$\theta_{ref} \leq \theta_i \leq \theta_{ref}, \qquad i = 1 \tag{5.9h}$$

$$v_i^{\min} \leq v_i \leq v_i^{\max}, \quad i = 1, ..., n_b \tag{5.9i}$$

$$p_i^{\min} \leq p_i \leq p_i^{\max}, \quad i = 1, ..., n_g \tag{5.9j}$$

$$q_i^{\min} \leq q_i \leq q_i^{\max}, \quad i = 1, ..., n_g \tag{5.9k}$$

$$|x_i - x_i^0| \leq \eta x_i^0, \quad i \in E_{DF} \tag{5.9l}$$

In (5.9), $\mathbf{Y} = [\theta \ \ \mathbf{V} \ \ \mathbf{P_g} \ \ \mathbf{Q_g} \ \ \mathbf{x}]$ are decision variables corresponding to voltage angle, voltage magnitude, generator active generation, generator reactive generation, and the re-

actance of D-FACTS lines, respectively; $cost(\mathbf{Y})$ is the system generation cost; $distP(\mathbf{Y})$ is the squared Euclidean distance between the normalized active power flow measurements before and after HMTD; $distX(\mathbf{x})$ is the squared Euclidean distance between the reactance before and after HMTD; λ_0, λ_1 and λ_2 are finely tuned weight parameters. In (5.9), active power flows and voltage magnitudes with subscript 0 are the measurements before HMTD; E_{DF} is the index set of D-FACTS lines; x_i^0 is the original reactance of i-th transmission line equipped with a D-FACTS device before HMTD; n_b, n_l, and n_g are the number of buses, lines, and generators, respectively. Constraints (5.9d) and (5.9e) are nonlinear equality constraints of the nodal active and reactive power balance, respectively. Constraints (5.9f) and (5.9g) are nonlinear inequality of line power flow limits corresponding to lines starting from from-end and to-end, respectively. Constraints (5.9h) and (5.9i) are voltage angle and magnitude constraints. Constraints (5.9j) and (5.9k) are generator constraints. In (5.9l), η in % reflects the physical capacity of D-FACTS devices and $\eta = 20\%$ is generally used in MTD[70;72;79;84;86–88;91]. In (5.9i), the per unit voltage magnitude boundary of Bus i is set as $v_i^{\max} = \min\{(1+\tau)v_{i,0},\ 1.05\}$ and $v_i^{\min} = \max\{(1-\tau)v_{i,0},\ 0.95\}$, where τ is the voltage perturbation magnitude. Note that a small τ (τ<0.5%) is suggested to ensure the voltage stability and MTD hiddenness. We solve the proposed AC-HMTD operation model (5.9) by using a modified MATLAB Interior Point Solver based on our prior work[71]. We highlight that the proposed AC-HMTD model can be easily integrated into the existing ACOPF function in EMS. It simultaneously dispatches the optimal power generations and D-FACTS setpoints.

It is worth mentioning that this chapter focuses on constructing HMTD with maximal detection effectiveness in transmission systems traditionally equipped with the supervisory control and data acquisition (SCADA) measurements. If phasor measurement unit (PMU) devices are installed at certain buses in the transmission system, one can add specific constraints in the proposed DC-HMTD operation model (5.8) and introduce extra terms in the objective function of the proposed AC-HMTD model (5.9). For example, in the DC-HMTD operation model (5.8), we can set the elements in $\mathbf{W}$ corresponding to the buses equipped with PMU devices to zero such that the voltage angle of buses equipped with PMU devices

remains unchanged after HMTD. In the AC-HMTD operation model (5.9), an additional term regarding the difference of PMU measurements before and after HMTD can be added into the objective function. However, this is beyond the scope of this work and will be investigated in our future work.

5.5.3 Cost-benefit Analysis of DC- and AC-HMTD Model

We conduct qualitative cost-benefit analyses of HMTD in both the DC and AC models. We compare the system generation cost in the following four cases, as summarized in Table 5.1. Case 0 is the base case, where the traditional OPF is conducted without MTD. Case 1 and Case 2 are the HMTD and RMTD, respectively. Case 3 is the OPF-based MTD[71], in which only generation cost is minimized while the hiddenness is not considered. The relationship among C_0, C_1, and C_2 has been discussed in[72]. In the DC model, the relationship is $C_0{=}C_1{\leq}C_2$ in a no-congestion condition and $C_0{=}C_1{\geq}{\leq}C_2$ in a transmission congestion condition[72]. In the AC model, the relationship is $C_1{\leq}C_0{\geq}{\leq}C_2$[72]. Here, we focus on discussing the relationship between C_1 and C_3.

Table 5.1: *Generation costs in OPF and different MTD methods*

C_0	Generation costs in OPF without MTD
C_1	Generation costs in HMTD
C_2	Generation costs in RMTD
C_3	Generation costs in OPF-based MTD

Compared with HMTD, the OPF-based MTD can dispatch the line reactance through D-FACTS devices to relieve line congestion within the physical operation range of D-FACTS devices. If the congestion is relieved, generation cost will decrease, i.e., $C_3{\leq}C_0$. If the congestion is not relieved at all or there is no congestion in the system, the generation cost in the OPF-based MTD is the same as that in OPF, i.e., $C_3{=}C_0$. In summary, we have $C_3{=}C_0{=}C_1{\leq}C_2$ in a no-congestion condition and $C_3{\leq}C_0{=}C_1{\geq}{\leq}C_2$ in a congestion condition in the DC model.

Since the constraints in the AC-HMTD operation model (5.9) are a subset of the con-

straints in the OPF-based MTD[71], the optimal solution obtained from AC-HMTD must be a feasible (but may not be the optimal) solution of the OPF-based MTD, i.e., $C_3 \leq C_1$. Therefore, we have $C_3 \leq C_1 \leq C_0 \geq \leq C_2$ in the AC model.

The qualitative cost-benefit analysis in both the DC and AC models above shows that HMTD will not increase generation costs as opposed to RMTD, but it may lead to a higher generation cost than that in the OPF-based MTD. As a result, HMTD accomplishes the MTD hiddenness by compromising the maximum economic benefits that D-FACTS devices could potentially achieve, representing a trade-off between the system economic and cyber-secure operations.

5.6 Experiment Results

We perform the proposed MTD planning and operation approach in the IEEE 14-bus system and the IEEE 57-bus system[101]. We use the former to show the hidden MTD planning solution and the latter to evaluate both the hiddenness and detection effectiveness as opposed to other existing methods. In either system, we take a customary approach where multiple lines sharing the same from-bus and to-bus are merged as a single line. The hidden MTD planning algorithm is implemented using the Java programming language. We solve the HMTD operation model using fmincon function of MATLAB. In a noisy condition, the measurement noise is assumed to be Gaussian distributed with zero mean and the standard deviation as 1% of the actual measurement. The threshold of BDD used by attackers and defenders is set to have a 1% false-positive rate. The algorithms are performed on a laptop with Intel Core i5 processor CPU 2.70 GHz dual-core with 8 GB RAM.

5.6.1 HMTD Planning Solution

The hidden MTD planning solution for the IEEE 14-bus system obtained by using Algorithms 4 and 5 is shown in Figure 5.3. It is seen that both G_{DF} (the red graph) and $G_{\overline{DF}}$ (the black graph) are loopless, indicating the HMTD under this MTD planning solution is a max-rank

MTD. In addition, $G_{\overline{DF}}$ is a disconnected graph, which ensures the existence of HMTD. The D-FACTS and non-D-FACTS lines in each loop satisfy Rules 1 and 2, which prevents the D-FACTS devices from working idly. Furthermore, D-FACTS lines cover all buses except for Bus 8, which is not in any loop.

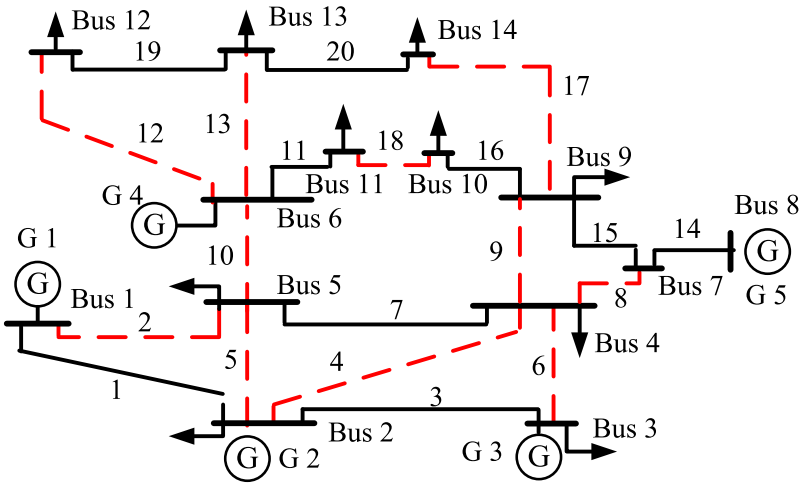


Figure 5.3: *Hidden MTD planning of the IEEE 14-bus system.*

In the IEEE 57-bus system, the proposed algorithms place D-FACTS devices on 47 lines, i.e., 60% of the transmission lines in this system, which are indexed by L_{DF}={2, 3, 4, 6, 9, 10, 11, 12, 13, 14, 15, 17, 19, 21, 22, 26, 29, 31, 32, 33, 35, 37, 39, 40, 43, 45, 46, 48, 50, 51, 52, 54, 57, 58, 59, 60, 63, 64, 67, 69, 71, 74, 75, 77, 78, 79, 80}.

5.6.2 DC-HMTD Operation Solution

In this subsection, based on the hidden MTD planning in the IEEE 57-bus system, we compare the HMTD operation model with the simplest MTD operation method, i.e., RMTD, in terms of the hiddenness and the detection effectiveness. We assume the attackers have the knowledge about the original line parameters and have read and write access to all measurements. In addition, the SE and BDD are used by attackers to detect if an MTD is in place.

We adopt a 24-hour load profile, which can be found at http://motor.ece.iit.edu/data. Under each load, we constructed 100 HMTDs and 100 RMTDs, respectively. For each MTD,

we assume the attacker launch BDD 100 times. Then, the DSP of each MTD is calculated and their mean value is treated as the DSP under that given load.

Regarding DC-FDI attacks $a = \mathbf{H}_0 \Delta\theta_a$, we generate 560 attack vectors, i.e., $\Delta\theta_a$ with a single attack target bus, i.e., $||\Delta\theta_a||_0 = 1$ as an attack pool. More specifically, we generate ten attack vectors for each bus in the IEEE 57-bus system (except for the reference bus), where the manipulated incremental voltage angle on the bus is uniformly distributed between (0.2, 0.4). For each MTD under each load, these 560 attacks, i.e., $a = \mathbf{H}_0 \Delta\theta_a$, are injected into the real measurement vector. Then, SE and BDD launched by system operators are used to detect the attacks. Similar to DSP, the ADP of each MTD is calculated, and their mean value is treated as the ADP under that given load.

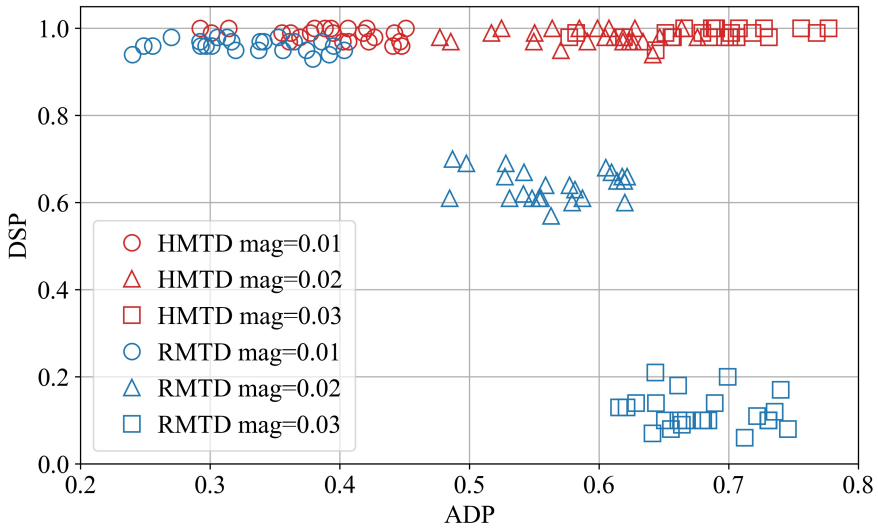


Figure 5.4: *ADP and DSP of HMTD and RMTD under different MTD magnitudes.*

Figure 5.4 demonstrates the ADP and the DSP of HMTDs and RMTDs versus different MTD magnitudes. Each node in Figure 5.4 represents the corresponding ADP and DSP under one load condition. It is seen that the HMTD operation method is always hidden to attackers regardless of MTD magnitudes, while a larger MTD magnitude contributes to an improved DSP of the HMTD. With the same MTD magnitude, the HMTD outperforms RMTD in terms of detection effectiveness since the proposed HMTD operation model maximizes the susceptance changes to introduce more uncertainties for attackers. Additionally,

when the MTD magnitude is small, the RMTD has very limited detection effectiveness, but it is hidden to attackers. This is because the susceptance changes introduced by the RMTD are too small to cause any change in power flow measurements. In RMTD, its DSP decreases while its ADP increases with an increase in the MTD magnitude.

We evaluate the influence of variant loads on the hiddenness of the proposed HMTDs through simulations. Suppose the system operator launches MTDs every T time period and setpoints of D-FACTS devices only change at the beginning of each time period by the proposed HMTD operation model, while the load of each bus can vary during this time period. Thus, we test the hiddenness of the proposed HMTD method under different levels of load changes in the IEEE 57-bus system. First, we assume the load of each bus randomly varies in the following range, i.e., $d \in [(1-\lambda)d_0, (1+\lambda)d_0]$, where λ is the load changing magnitude and d_0 is the load used to construct HMTDs. In the attackers' point of view, an average DSP of 100 HMTDs is calculated when the system loads keep changing under the given load magnitude.

The impact of variant loads on the hiddenness of the proposed HMTD method under different levels of noise standard deviation σ is shown in Figure 5.5. It is seen that the DSP decreases as the load magnitude increases. This is because the load changes result in power flow changes that deteriorate the hiddenness condition. In Figure 5.5, a higher noise level mitigates the negative impact of variant load on the hiddenness, leading to a higher DSP. This can be explained by investigating the attacker's BDD. Specifically, a higher noise level makes the attacker's BDD tolerate higher deviations between the measured and estimated power flows.

5.6.3 AC-HMTD Operation Solution

In this subsection, we compare the proposed AC-HMTD operation model with the traditional ACOPF model based on the proposed hidden MTD planning in the IEEE 14-bus system in terms of the generation cost, DSP and ADP. First, the traditional ACOPF is conducted in the IEEE 14-bus system, denoted as Case 0 (i.e., a no-MTD case), and its resultant

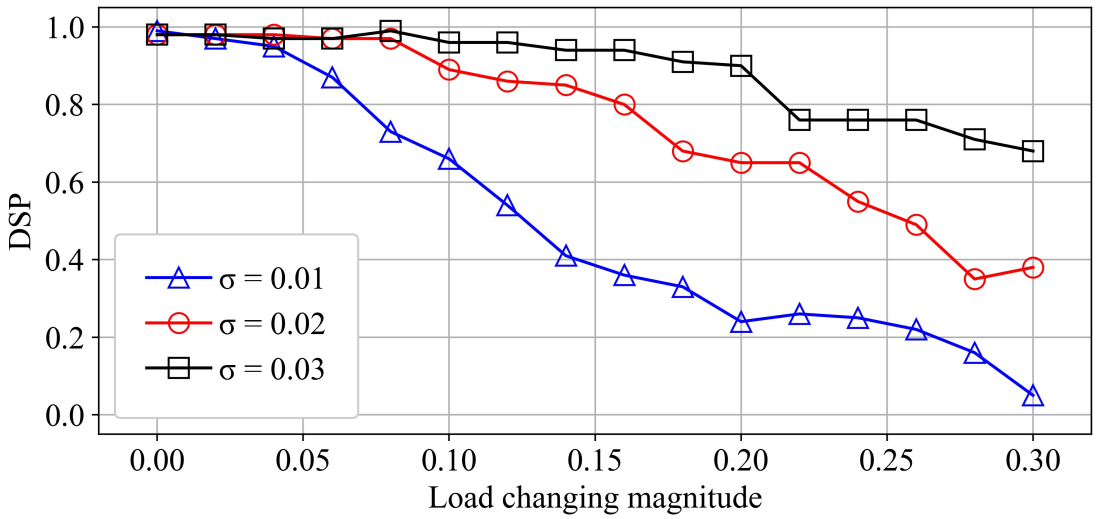


Figure 5.5: *The impact of variant loads on the hiddenness of proposed HMTD.*

measurements are adopted as the reference before HMTD. In Case 1, only the generation cost is minimized without considering the hiddenness or reactance changes. Accordingly, let λ_1 and λ_2 be zeros, and the lower and upper bounds on the voltage magnitude in constraint (5.9i) be 0.95 and 1.05, respectively. In Cases 2, 3, and 4, we apply the proposed HMTD operation approach with a decreasing value of λ_0 in model (5.9) to show the impact of λ_0 on the hiddenness.

A comparison of these five cases is summarized in Table 5.2. Here, we choose the system without MTD (Case 0) as a baseline. We calculate the generation cost savings accrued by a MTD as MTD savings, and compare this to the generation cost in the baseline. In Table 5.2, the average reactance changes in percentage (RCP) in the proposed HMTDs are more than 10%, which ensures the attack detection capability of HMTDs. It is observed that the proposed HMTD operation approach creates MTD savings compared to the baseline generation cost. Table 5.2 exhibits a trade-off between the MTD savings and its hiddenness. As seen in Table 5.2, the MTD without considering the hiddenness in Case 1 has the highest MTD savings. When DSP increases to 90% in Case 4, its MTD savings decreases to \$7.57. We further demonstrate the trade-off in Figure 5.6, where λ_0 varies from 10^{-6} to 10^{-4}. With a decreasing λ_0, the hiddenness of MTD increases but the MTD savings decreases. The

simulation results in Table 5.2 and Figure 5.6 verify the cost-benefit analysis of HMTD in Section 5.5.3.

We further evaluate the detection effectiveness of the proposed AC-HMTD operation model in the IEEE 14-bus system. We construct 130 single-bus AC-FDI attacks, in which each bus is attacked ten times except for the reference bus. Each attack is launched on each of MTDs outlined in Table 5.2. The ADP of each MTD is calculated as shown in the last column of Table 5.2. As the node degree of Bus 8 is one, attacks on Bus 8 are undetectable due to a limitation of MTD[87]. Thus, the largest ADP in the IEEE 14-bus system is 92.3%. The ADP of HMTDs in Cases 3 and 4 is lower than 92.3% since the reactance change of Line 4-7 is low under the given load. The attack detection performance of the AC-HMTD is consistent with that in the DC model.

Table 5.2: *The performance of AC-HMTD operation*

Case	λ_0	λ_1	λ_2	Cost ($)	MTD Savings ($)	DSP (%)	RCP (%)	ADP (%)
Case 0	1	N/A	N/A	8131.52	0	N/A	0	N/A
Case 1	1	0	0	8115.69	15.83	0	18.0	92.3
Case 2	1e-4	0.01	0.05	8119.36	12.16	1.0	11.0	92.3
Case 3	1e-5	0.01	0.05	8122.67	8.85	83.0	12.2	83.1
Case 4	1e-6	0.01	0.05	8123.95	7.57	90.0	12.2	85.4

5.6.4 Comparison between Proposed and Existing DC-HMTD Operations

In this section, we compare the proposed HMTD operating method (5.8) with RW-HMTD[72] under the proposed MTD planning in both the IEEE 14-bus and 57-bus systems.

In the RW-HMTD, a weight searching range must be initialized to find a feasible solution. However, strategies of searching range initialization are not provided in reference[72]. In this case, we design two simple strategies, i.e., a direct searching and an indirect searching strategy. The direct searching method sets the searching boundary using searching radius

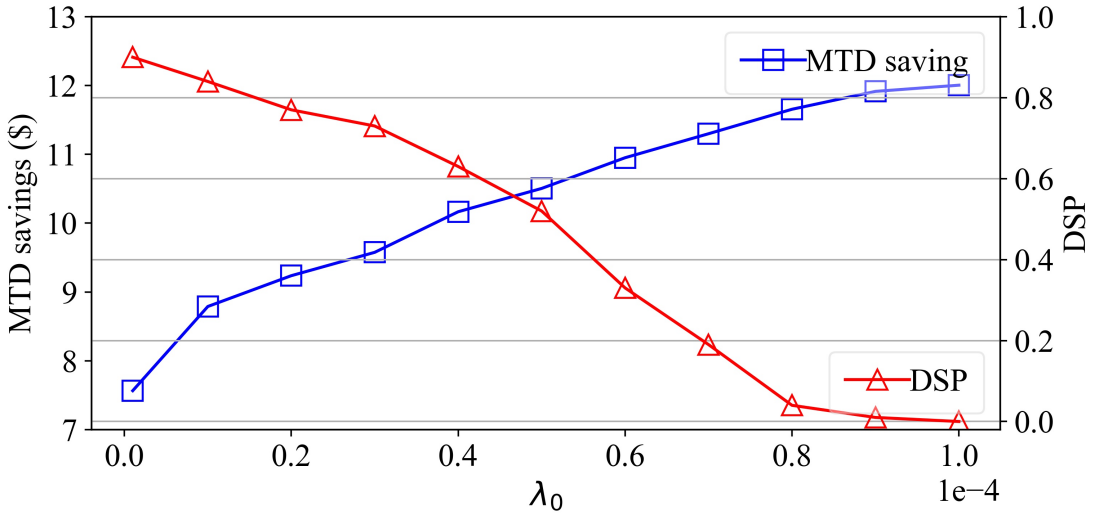


Figure 5.6: *The trade-off between MTD savings and MTD hiddenness under different* λ_0.

γ, i.e., $w \in [-\gamma, \gamma]$. The indirect searching method takes the solution w_0 from (5.8) as the center and then specifies the searching radius using a factor β, i.e., $w \in [(1-\beta)w_0, (1+\beta)w_0]$. Note that 4 weights and 25 weights need to be determined by HMTD in IEEE 14-bus and 57-bus systems, respectively. In RW-HMTD, we apply the direct searching method in the IEEE 14-bus system and the indirect searching method in the IEEE 57-bus system. This is because the direct searching method fails to find feasible solutions in the IEEE 57-bus system in a reasonable amount of time. We have to take advantage of the results in our proposed method and narrow down the searching range using the indirect searching method.

The performance of the proposed HMTD operation is summarized in Table 5.3. It is observed that the reactance changes more than 14% compared with the original line reactance in both systems. The CPU time of the proposed HMTD operation is less than 1.1 seconds in both systems.

Table 5.3: *Performance of the proposed HMTD operation*

System	RCP (%)	CPU Time (s)
14-bus System	14.50	0.31
57-bus System	14.71	1.06

The performance of the RW-HMTD in the 14-bus and 57-bus systems is summarized in

Tables 5.4 and 5.5, respectively. We obtain five feasible solutions in each searching range and then calculate the minimum, maximum, and mean of the reactance changes in percentage (RCP) as well as the CPU time. It is observed that the CPU time dramatically increases as the searching radius increases, especially in the IEEE 57-bus system. The RCP can be as low as 3.16% in the IEEE 14-bus system. In the IEEE 57-bus system, the RCP decreases accordingly when the searching radius increases. This is because the RW-HMTD solutions obtained within a larger searching range may deviate further from the optimal solution (i.e., the largest RCP point) provided by our proposed model (5.8).

Table 5.4: *Performance of RW-HMTD using direct searching in the IEEE 14-bus system*

Searching range	RCP (%)			CPU Time (s)		
	min	max	mean	min	max	mean
[-0.01,0.01]	3.16	12.10	7.35	0.001	0.021	0.007
[-0.05,0.05]	6.48	8.70	7.76	0.002	0.506	0.200
[-0.10,0.10]	3.54	11.00	7.69	0.868	3.901	1.950
[-0.15,0.15]	5.11	9.64	7.84	2.872	31.351	19.702
[-0.20,0.20]	5.90	8.97	7.51	0.348	91.923	30.409

Table 5.5: *Performance of RW-HMTD using indirect searching in the IEEE 57-bus system*

Factor β	RCP (%)			CPU Time (s)		
	min	max	mean	min	max	mean
0.05	12.62	12.84	12.73	1.4	20.1	9.1
0.10	11.45	12.76	12.10	20.0	192.2	97.5
0.15	10.79	11.57	11.31	20.4	2488.2	741.9
0.20	10.01	11.30	10.80	400.2	3185.2	2374.9

We further compare the detection effectiveness of the proposed HMTD and three RW-HMTDs under FDI attacks with different voltage angle injection magnitudes (VAIM) in the IEEE 14-bus system. Specifically, FDI attacks with $\Delta\theta_a$ are randomly generated in the range $\Delta\theta_a \in [0.8, 1.2] \cdot \bar{\theta} \cdot VAIM$, where VAIM reflects the strength of FDI attacks. Comparative results are shown in Figure 5.7. The proposed HMTD has the largest ADP. Low reactance changes in RW-HMTD decrease the detection capability of MTDs, especially under the FDI attacks with the small voltage angle injection magnitude. Note that these

three RW-HMTDs are constructed under the proposed HMTD MTD planning solution, which has maximal detection effectiveness. If an RW-HMTD is constructed under other MTD planning methods, its ADP can further decrease.

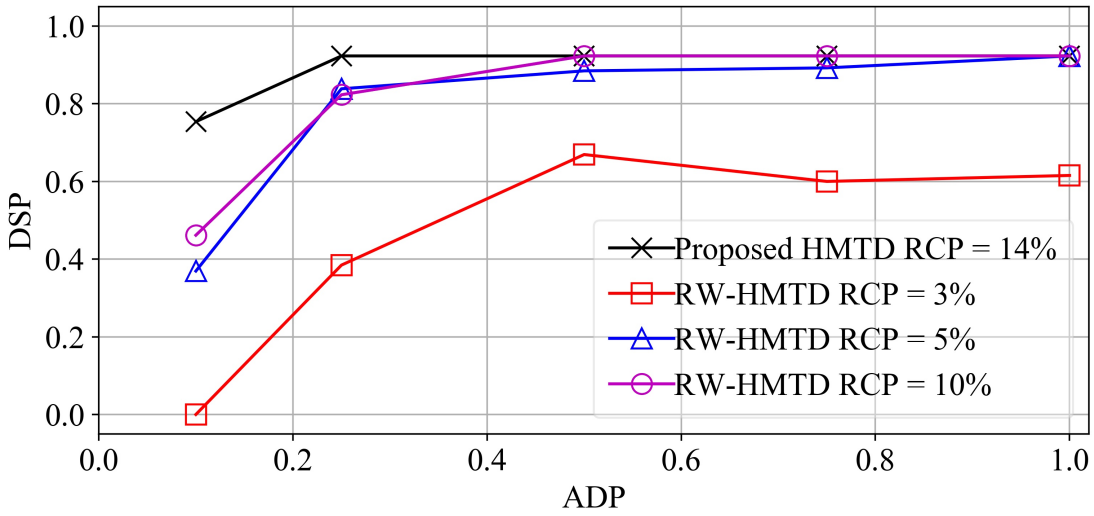


Figure 5.7: *ADP of RW-HMTD and proposed HMTD under FDI attacks with different VAIMs.*

In summary, the drawbacks of the RW-HMTD method[72] are two-fold. First, this method may generate an MTD with small reactance changes resulting in a low attack detection capability. Second, its CPU time heavily depends on the weight searching range. A larger searching radius will result in a much longer searching time, especially in large-scale systems. To make things worse, an improper searching range can cause no solution obtained. The proposed method circumvents these drawbacks by utilizing optimization to find the largest reactance changes in HMTD efficiently.

5.6.5 Comparison between Hidden and Existing MTD Planning Algorithms

In this subsection, we compare the proposed MTD planning with the other four existing planning methods, including the max-rank planning[70] in Chapter 3, the full planning, the arbitrary planning[84], and the spanning-tree planning[87]. These MTD planning methods are

summarized in Table 5.6. Except for the max-rank planning and the hidden planning, the rank of the composite matrix under other planning methods depends on the setpoints of D-FACTS devices.

Table 5.6: *Existing MTD planning algorithms*

MTD planning	Description of MTD planning	p_1
Arbitrary planning[84]	Install on the randomly selected lines	47
Full planning	Install on all transmission lines	78
Spanning-tree planning[87]	Install on lines that form a spanning tree of the system topology	56
Max-rank planning[70]	Non-D-FACTS lines form a spanning tree, and D-FACTS are placed on the remaining lines	22

Consistent with the experiment setup in the previous subsection, we apply MTDs under the above planning methods and calculate the ADP and the DSP under each load with a fixed MTD magnitude of 0.2. We run the HMTD operation model under the hidden planning and RMTDs under other planning methods. It is worthwhile to mention that the same attack pool is used to calculate the ADP. The ADP and the DSP under five MTD planning methods are shown in Figure 5.8. As seen, MTDs under the hidden planning are hidden to attackers, while MTDs under the max-rank planning can be detected by attackers. In addition, the ADP of MTDs under the hidden planning is higher than that under the max-rank planning due to the covered buses in the hidden planning. RMTDs under the other planning methods can always be detected by attackers.

Even though the rank of the composite matrix is maximized, the max-rank planning has the worst detection effectiveness due to 27 uncovered buses. Arbitrary planning uses extra 25 D-FACTS devices compared with that in the max-rank planning. The arbitrary planning used in this case study has five uncovered buses. Thus, its detection effectiveness is better than the max-rank planning but worse than either the spanning-tree planning or the full planning. Both the spanning-tree and the full plannings have similar detection effectiveness since they both cover all the buses using the D-FACTS devices. However, their ADPs are still worse than that of the hidden planning. This is because their rank of the composite matrix depends on the setpoints of D-FACTS devices. Specifically, if the reactance of all lines

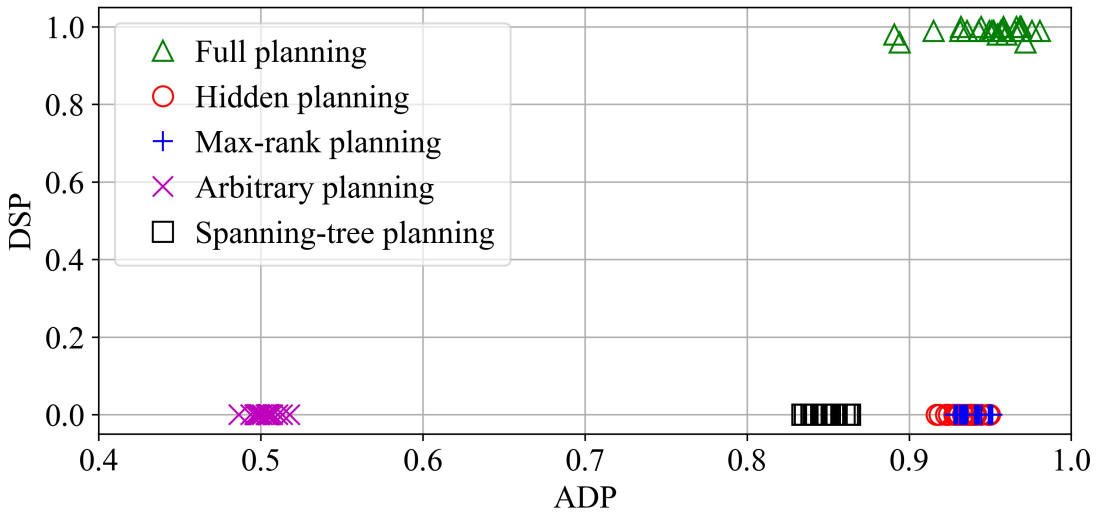


Figure 5.8: *ADP and DSP of five MTD planning algorithms under 0.2 MTD magnitude.*

connected to one bus is modified by multiplying a unity factor, their rank of the composite matrix will decrease by one. Consequently, any FDI attack on this bus is undetectable.

Figure 5.9 demonstrates a transition between the MTD hiddenness and the detection effectiveness in each MTD planning method. For each planning, we apply six discrete MTD magnitudes, i.e., 1%, 2%, 3%, 4%, 5%, and 20%, to calculate the ADP and the DSP. Note that the green arrows on each line in Figure 5.9 show the direction in which the MTD magnitude is increasing. We observe, for the first time, that the proposed MTD planning and operation method is always hidden to attackers and provides an excellent ADP under the MTD magnitude of 0.2. In comparison, when the MTD magnitude is increased, each other planning method shows a clear transition from a low ADP with a high DSP to a high ADP with a low DSP. As opposed to the MTD hiddenness, the detection effectiveness of MTDs is the fundamental requirement. Therefore, a large MTD magnitude is always desirable for the RMTD operation.

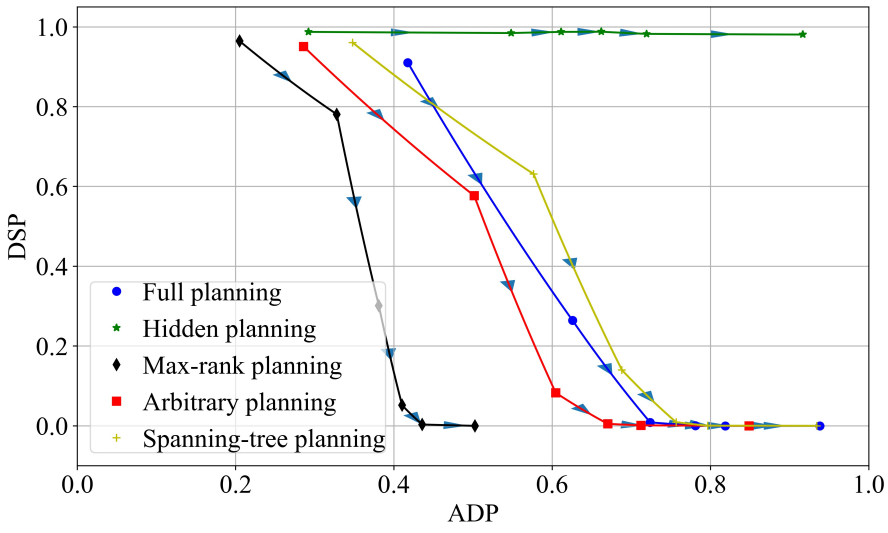


Figure 5.9: *ADP and DSP of five MTD planning algorithms with MTD magnitude varying from 1% to 20%.*

5.7 Summary

In this chapter, we propose a DFS-based hidden D-FACTS devices planning algorithm as well as the DC- and AC-HMTD operation models. We emphasize that the MTD hiddenness and detection effectiveness can be achieved simultaneously in incomplete MTDs. The proposed planning algorithm ensures the existence of HMTD and enables MTDs to have maximal detection effectiveness. The proposed hidden planning uses fewer D-FACTS devices to reach the maximal detection effectiveness compared to the full planning and spanning-tree planning.

We propose an optimization-based DC-HMTD operation model, which integrates the derived hidden operation condition as constraints. Case studies show that the proposed model is superior to the existing HMTD operation method in terms of computational time and detection effectiveness. The transition between the MTD hiddenness and the detection effectiveness versus the MTD magnitude is also presented. Additionally, we propose an ACOPF-based HMTD operation model, which minimizes the generation cost and achieves the MTD hiddenness. Simulation results show a trade-off between the generation cost savings by MTD and MTD hiddenness in the AC-HMTD operation. The results demonstrate that

the attack detection performance of AC-HMTD is consistent with that in the DC model.

Chapter 6

Conclusion and Future Work

This chapter concludes the dissertation with the main research results and provides directions for possible future work.

6.1 Conclusion

This dissertation aims to investigate the moving target defense (MTD) approaches to detect false data injection (FDI) attacks against power system state estimation (SE). This dissertation distinguishes the role that MTD planning and operation play in the MTD detection effectiveness, MTD cost, and MTD hiddenness. I prove that MTD planning determines the MTD detection effectiveness and highlight that MTD operation needs to focus on reducing the MTD operation cost. My contributions to the MTD theory in power systems are summarized from the following perspectives:

- **MTD detection effectiveness.** Chapter 3 proves that an MTD is a max-rank MTD if no distributed flexible AC transmission system (D-FACTS) devices work in idle states, and there exists no loop in either D-FACTS graph or non-D-FACTS graph. It indicates that any MTDs under the MTD planning satisfying this sufficient condition guarantee their detection effectiveness regardless of the D-FACTS setpoints, which contributes to separating MTD planning and MTD operation as two independent problems. In

addition, Chapter 3 reveals that the maximum rank of the composite matrix is not equivalent to the maximal MTD detection effectiveness, which is merely the lower bound of attack detection probability (ADP). Chapter 3 further identifies three types of unprotected buses in MTD, which determine the upper bound of ADP.

- **MTD planning cost.** Chapter 3 derives the minimum number of D-FACTS devices required to achieve the complete MTD and max-rank incomplete MTDs, respectively. It requires at least n–1 D-FACTS devices for a complete MTD and at least $p - n + 1$ D-FACTS devices for a max-rank incomplete MTD.

- **MTD operation cost.** Chapter 4 integrates D-FACTS devices into the optimal power flow (OPF) model to minimize system generation costs and system losses. Chapter 5 conducts qualitative cost-benefit analyses of MTD in both the DC and AC models. The OPF-based MTD operation model has the lowest generation costs. Hidden MTD (HMTD) will not increase generation costs as opposed to random MTD (RMTD), but it may lead to a higher generation cost than that in the OPF-based MTD. As a result, HMTD accomplishes the MTD hiddenness by compromising the maximum economic benefits that D-FACTS devices could potentially achieve, representing a trade-off between the system economic and cybersecure operations.

- **MTD hiddenness.** Chapter 5 derives a novel and explicit hiddenness condition in HMTD, which can be easily integrated into MTD operation methods. It is proved that a max-rank HMTD exists if non-D-FACTS graph is a disconnected loopless graph and D-FACTS graph is a loopless graph with no idle-state D-FACTS devices. This work also demonstrates the characteristics of voltage angle changes in HMTD, which bridge the HMTD operation and HMTD planning.

Furthermore, this dissertation proposes four MTD planning algorithms with different objectives. I summarize the characteristics of these MTD planning algorithms as follows:

- **Max-rank planning algorithm.** Chapter 3 proposes two max-rank planning algorithms for constructing complete and incomplete MTD, respectively. These algorithms

maximize the rank of the composite matrix to ensure detection effectiveness, and utilize the minimum number of D-FACTS devices to minimize the MTD planning cost. In addition, D-FACTS devices are installed on the lines with large power loss to impedance sensitivity (PLIS), which contributes to reducing system losses, equivalent to reducing MTD operation cost. However, the detection effectiveness of the max-rank planning algorithm for the incomplete MTD is limited due to the existence of non-D-FACTS buses in this planning.

- **Graph-theory-based planning algorithm.** Chapter 3 further proposes a graph-theory-based planning algorithm that ensures maximal detection effectiveness and considers the economic benefits from D-FACTS devices. The proposed algorithm eliminates non-D-FACTS buses to increase the ADP upper bound and simultaneously remain the maximum rank of the composite matrix to achieve a high ADP lower bound. Numerical results verify that the ADP of the proposed planning is better than that of the arbitrary planning, the max-rank planning, and the full planning in DC-SE.

- **Hidden planning algorithm.** Chapter 5 proposes a depth-first-search-based hidden planning algorithm, which ensures the existence of HMTD and maximal detection effectiveness without using protected meters. The proposed planning method ensures the maximum rank of the composite matrix and eliminates non-D-FACTS buses. In addition, the proposed planning method uses fewer D-FACTS devices to reach the maximal detection effectiveness compared to the full planning and spanning-tree planning.

Finally, this dissertation proposes an ACOPF-based MTD operation model and DC- and AC-HMTD operation models. All three MTD operation models can be seamlessly integrated into the existing energy management system in the system control room. In addition, this dissertation develops an interior-point solver to resolve the ACOPF-based MTD operation models. I summarize the proposed MTD operation models and the interior-point solver as follows:

- **ACOPF-based MTD operation model.** Chapter 4 proposes an ACOPF-based

MTD operation model, which minimizes the system losses and generation costs to reduce the MTD operation cost. The proposed method can save more generation costs and system losses compared with that in the traditional ACOPF model due to the D-FACTS devices. The proposed method prevents the D-FACT devices from working in the idle state at the expense of a slight increase in system losses to ensure the maximum rank of the composite matrix. System operators can simultaneously obtain the economic benefits and cyber-defense benefits from D-FACTS devices.

- **Optimization-based DC-HMTD operation model.** Chapter 5 proposes a DC-HMTD operation model, which integrates the derived hiddenness condition as constraints. Case studies show that the proposed model is superior to the existing HMTD operation method in terms of CPU time and detection effectiveness. Specifically, reactance changes in percentage (RCP) is more than 14%, and CPU time is less than 1.1 seconds in both the IEEE 14-bus and 57-bus systems.

- **ACOPF-based HMTD operation model** Chapter 5 proposes an ACOPF-based HMTD operation model, which minimizes a weighted sum of the generation cost, the negative of reactance changes, and the squared Euclidean distance between the normalized active power flow measurements before and after HMTD. Minimizing the changes in measurements contributes to achieving the MTD hiddenness, and maximizing the reactance changes ensures detection effectiveness. Simulation results verify that the attack detection performance of the AC-HMTD is consistent with that in the DC-HMTD operation.

- **Modified Matlab interior-point solver.** Chapter 4 builds an interior-point solver for the proposed ACOPF-based MTD operation model by deriving the gradient and Hessian matrices of the objective function and constraints with respect to the line reactance. Since the derivations adopt the same voltage coordinate and complex power expression as MATPOWER, the derived gradient and Hessian matrices can be easily integrated into the existing Matlab interior-point solver. The numerical results show

that CPU time of solving the ACOPF-based MTD model is generally less than 15 seconds in the IEEE 118-bus system.

6.2 Future Work

This section summarizes some potential research directions as follows:

Chapter 3 analyzes the MTD detection effectiveness in the DC model. However, the metric of the MTD detection effectiveness in the AC model is missing in the literature. Thus, it is necessary to propose a novel metric of the MTD detection effectiveness in the AC model and analyze the detection effectiveness of existing MTD operation and planning methods in AC-SE.

Chapter 4 focuses on the MTD operation model. The proposed operation models consider the MTD operation cost, and MTD hiddenness, and MTD detection effectiveness. However, the impact of MTD operation model on system stability is not considered in the literature. Investigate this impact will contribute to the application of MTD in practical power grids.

Chapter 5 focuses on constructing HMTD with maximal detection effectiveness in transmission systems traditionally equipped with the supervisory control and data acquisition (SCADA) measurements. However, phasor measurement unit (PMU) devices are the most advanced metering devices which can provide voltage and current synchrophasor measurements in real-time. It is therefore necessary to investigate the integration of PMU devices into HMTD planning and operation methods.

This dissertation studies the detection effectiveness of MTD against general stealthy FDI attacks in the case study. It is necessary to investigate the performance of MTD under more specific sophisticated adversary models in the future. In addition, this dissertation studies the application of MTD in transmission systems. It is therefore worthwhile to investigate the performance of MTD in distribution systems. The application of MTD in distribution systems needs to consider the low-observability in distribution system state estimation methods[111;112].

Bibliography

[1] D Hadziosmanovic. The process matters: cyber security in industrial control systems. 2014.

[2] Ahmed S Musleh, Guo Chen, and Zhao Yang Dong. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 11(3):2218–2234, 2019.

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

[4] Michael Swearingen, Steven Brunasso, Joe Weiss, and Dennis Huber. What you need to know (and don't) about the aurora vulnerability. *Power*, 157(9):52–52, 2013.

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

[6] ABC News. Trojan horse' bug lurking in vital us computers since 2011, 2021. URL https://abcnews.go.com/US/trojan-horse-bug-lurking-vital-us-computers-2011/story?id=26737476.

[7] Gaoqi Liang, Steven R Weller, Junhua Zhao, Fengji Luo, and Zhao Yang Dong. The 2015 ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, 32(4):3317–3318, 2016.

[8] Marie Baezner. Cyber and information warfare in the ukrainian conflict. Technical report, ETH Zurich, 2018.

[9] U.S. Homeland Security. Moving target defense, 2011. URL https://www.dhs.gov/science-and-technology/csd-mtd.

[10] Jin-Hee Cho, Dilli P Sharma, Hooman Alavizadeh, Seunghyun Yoon, Noam Ben-Asher, Terrence J Moore, Dong Seong Kim, Hyuk Lim, and Frederica F Nelson. Toward proactive, adaptive defense: A survey on moving target defense. *IEEE Communications Surveys & Tutorials*, 22(1):709–745, 2020.

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

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

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

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

[15] Stephen W Boyd, Gaurav S Kc, Michael E Locasto, Angelos D Keromytis, and Vassilis Prevelakis. On the general applicability of instruction-set randomization. *IEEE Transactions on Dependable and Secure Computing*, 7(3):255–270, 2008.

[16] Ilhami Colak. Introduction to smart grid. In *2016 International Smart Grid Workshop and Certificate Program (ISGWCP)*, pages 1–5. IEEE, 2016.

[17] Mark J Stanovich, Isaac Leonard, K Sanjeev, Mischa Steurer, Thomas P Roth, Stephen Jackson, and Matthew Bruce. Development of a smart-grid cyber-physical systems testbed. In *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6. IEEE, 2013.

[18] Khairy Sayed and Hossam A Gabbar. Scada and smart energy grid control automation. In *Smart Energy Grid Engineering*, pages 481–514. Elsevier, 2017.

[19] Ye Yan, Yi Qian, Hamid Sharif, and David Tipper. A survey on smart grid communication infrastructures: Motivations, requirements and challenges. *IEEE communications surveys & tutorials*, 15(1):5–20, 2012.

[20] Mahmoud Saleh, Yusef Esa, and Ahmed A Mohamed. Communication-based control for dc microgrids. *IEEE Transactions on Smart Grid*, 10(2):2180–2195, 2018.

[21] Sumit K Rathor and D Saxena. Energy management system for smart grid: An overview and key issues. *International Journal of Energy Research*, 44(6):4067–4109, 2020.

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

[23] Adam Hahn and Manimaran Govindarasu. Cyber attack exposure evaluation framework for the smart grid. *IEEE Transactions on Smart Grid*, 2(4):835–843, 2011.

[24] Yousif Dafalla, Bo Liu, Dalton A Hahn, Hongyu Wu, Reza Ahmadi, and Alexandru G Bardas. Prosumer nanogrids: A cybersecurity assessment. *IEEE Access*, 8:131150–131164, 2020.

[25] Victoria Y Pillitteri and Tanya L Brewer. Guidelines for smart grid cybersecurity. 2014.

[26] Electric Power Research Institute . National electric sector cybersecurity organization resource (nescor), 2014. URL https://doi.org/10.2172/1163840.

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

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

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

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

[31] Rui Tan, Varun Badrinath Krishna, David KY Yau, and Zbigniew Kalbarczyk. Impact of integrity attacks on real-time pricing in smart grids. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 439–450, 2013.

[32] Yue Zhang, VVG Krishnan, Jiaxing Pi, K Kaur, Anurag Srivastava, Adam Hahn, and S Suresh. Cyber physical security analytics for transactive energy systems. *IEEE Transactions on Smart Grid*, 11(2):931–941, 2019.

[33] Rui Tan, Hoang Hai Nguyen, Eddy YS Foo, David KY Yau, Zbigniew Kalbarczyk, Ravishankar K Iyer, and Hoay Beng Gooi. Modeling and mitigating impact of false data injection attacks on automatic generation control. *IEEE Transactions on Information Forensics and Security*, 12(7):1609–1624, 2017.

[34] Chunyu Chen, Kaifeng Zhang, Kun Yuan, Lingzhi Zhu, and Minhui Qian. Novel detection scheme design considering cyber attacks on load frequency control. *IEEE Transactions on Industrial Informatics*, 14(5):1932–1941, 2017.

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

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

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

[38] Kin Cheong Sou, Henrik Sandberg, and Karl Henrik Johansson. Electric power network security analysis via minimum cut relaxation. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4054–4059. IEEE, 2011.

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

[40] Jinsub Kim, Lang Tong, and Robert J Thomas. Subspace methods for data attack on state estimation: A data driven approach. *IEEE Transactions on Signal Processing*, 63(5):1102–1114, 2014.

[41] Zong-Han Yu and Wen-Long Chin. Blind false data injection attack using pca approximation method in smart grid. *IEEE Transactions on Smart Grid*, 6(3):1219–1226, 2015.

[42] Yanling Yuan, Zuyi Li, and Kui Ren. Modeling load redistribution attacks in power systems. *IEEE Transactions on Smart Grid*, 2(2):382–390, 2011.

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

[44] Liang Che, Xuan Liu, Zuyi Li, and Yunfeng Wen. False data injection attacks induced sequential outages in power systems. *IEEE Transactions on Power Systems*, 34(2): 1513–1523, 2018.

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

[46] Jinsub Kim and Lang Tong. On topology attack of a smart grid: Undetectable attacks and countermeasures. *IEEE Journal on Selected Areas in Communications*, 31(7): 1294–1305, 2013.

[47] Desmond Schmidt, Kenneth Radke, Seyit Camtepe, Ernest Foo, and Michał Ren. A survey and analysis of the gnss spoofing threat and countermeasures. *ACM Computing Surveys (CSUR)*, 48(4):1–31, 2016.

[48] Yi Cui, Feifei Bai, Yilu Liu, Peter L Fuhr, and Marissa E Morales-Rodriguez. Spatio-temporal characterization of synchrophasor data against spoofing attacks in smart grids. *IEEE Transactions on Smart Grid*, 10(5):5807–5818, 2019.

[49] Paresh Risbud, Nikolaos Gatsis, and Ahmad Taha. Vulnerability analysis of smart grids to gps spoofing. *IEEE Transactions on Smart Grid*, 10(4):3535–3548, 2018.

[50] Saleh Soltan, Mihalis Yannakakis, and Gil Zussman. React to cyber attacks on power grids. *IEEE Transactions on Network Science and Engineering*, 6(3):459–473, 2018.

[51] Saleh Soltan and Gil Zussman. Expose the line failures following a cyber-physical attack on the power grid. *IEEE Transactions on Control of Network Systems*, 6(1): 451–461, 2018.

[52] Saleh Soltan, Prateek Mittal, and H Vincent Poor. Line failure detection after a cyber-physical attack on the grid using bayesian regression. *IEEE Transactions on Power Systems*, 34(5):3758–3768, 2019.

[53] Jue Tian, Rui Tan, Xiaohong Guan, Zhanbo Xu, and Ting Liu. Moving target defense approach to detecting stuxnet-like attacks. *IEEE transactions on smart grid*, 11(1): 291–300, 2019.

[54] Yi Huang, Husheng Li, Kristy A Campbell, and Zhu Han. Defending false data injection attack on smart grid network using adaptive cusum test. In *2011 45th Annual Conference on Information Sciences and Systems*, pages 1–6. IEEE, 2011.

[55] Lanchao Liu, Mohammad Esmalifalak, Qifeng Ding, Valentine A Emesih, and Zhu Han. Detecting false data injection attacks on power grid by sparse optimization. *IEEE Transactions on Smart Grid*, 5(2):612–621, 2014.

[56] Junbo Zhao, Gexiang Zhang, Massimo La Scala, Zhao Yang Dong, Chen Chen, and Jianhui Wang. Short-term state forecasting-aided method for detection of smart grid general false data injection attacks. *IEEE Transactions on Smart Grid*, 8(4):1580–1590, 2015.

[57] Jacob Sakhnini, Hadis Karimipour, and Ali Dehghantanha. Smart grid cyber attacks detection using supervised learning and heuristic feature selection. In *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*, pages 108–112. IEEE, 2019.

[58] Xiangyu Niu, Jiangnan Li, Jinyuan Sun, and Kevin Tomsovic. Dynamic detection of false data injection attack in smart grid using deep learning. In *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6. IEEE, 2019.

[59] Rakesh B Bobba, Katherine M Rogers, Qiyan Wang, Himanshu Khurana, Klara Nahrstedt, and Thomas J Overbye. Detecting false data injection attacks on dc state estimation. In *Preprints of the First Workshop on Secure Control Systems, CPSWEEK*, volume 2010. Stockholm, Sweden, 2010.

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

[61] Junjian Qi, Kai Sun, and Wei Kang. Optimal pmu placement for power system dynamic state estimation by using empirical observability gramian. *IEEE Transactions on power Systems*, 30(4):2041–2054, 2014.

[62] Peng Zhuang, Talha Zamir, and Hao Liang. Blockchain for cybersecurity in smart grid: A comprehensive survey. *IEEE Transactions on Industrial Informatics*, 17(1): 3–19, 2020.

[63] Fengji Luo, Junhua Zhao, Zhao Yang Dong, Yingying Chen, Yan Xu, Xin Zhang, and Kit Po Wong. Cloud-based information infrastructure for next-generation power grid: Conception, architecture, and applications. *IEEE Transactions on Smart Grid*, 7(4): 1896–1912, 2015.

[64] Michael Mylrea and Sri Nikhil Gupta Gourisetti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23. IEEE, 2017.

[65] E Riva Sanseverino, Maria Luisa Di Silvestre, Pierluigi Gallo, Gaetano Zizzo, and Mariano Ippolito. The blockchain in microgrids for transacting energy and attributing losses. In *2017 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 925–930. IEEE, 2017.

[66] Xing Luo, Jihong Wang, Mark Dooner, and Jonathan Clarke. Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied energy*, 137:511–536, 2015.

[67] Matevž Pustišek, Andrej Kos, and Urban Sedlar. Blockchain based autonomous selection of electric vehicle charging station. In *2016 international conference on identification, information and knowledge in the Internet of Things (IIKI)*, pages 217–222. IEEE, 2016.

[68] Lawryn Edmonds, Bo Liu, Hang Zhang, Caterina Scoglio, Don Gruenbacher, and Hongyu Wu. Blockchain-enabled transactive home energy management systems in distribution networks. In *2020 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–5. IEEE, 2020.

[69] Executive Offce of the President, U.S. National electric grid security and resilience action plan, 2016. URL https://www.whitehouse.gov/sites/whitehouse.gov/fles/images/National_Electric_Grid_Action_Plan_06Dec2016.pdf.

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

[71] Bo Liu, Lawryn Edmonds, Hang Zhang, and Hongyu Wu. An interior-point solver for optimal power flow problem considering distributed facts devices. In *2020 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–5. IEEE, 2020.

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

[73] Chensheng Liu, Jing Wu, Chengnian Long, and Yebin Wang. Dynamic state recovery for cyber-physical systems under switching location attacks. *IEEE Transactions on Control of Network Systems*, 4(1):14–22, 2016.

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

[75] Suzhi Bi and Ying Jun Zhang. Graphical methods for defense against false-data injection attacks on power system state estimation. *IEEE Transactions on Smart Grid*, 5 (3):1216–1227, 2014.

[76] Suzhi Bi and Ying Jun Zhang. Using covert topological information for defense against malicious attacks on dc state estimation. *IEEE Journal on Selected Areas in Communications*, 32(7):1471–1485, 2014.

[77] Le Xie, Yilin Mo, and Bruno Sinopoli. Integrity data attacks in power market operations. *IEEE Transactions on Smart Grid*, 2(4):659–666, 2011.

[78] Liyan Jia, Jinsub Kim, Robert J Thomas, and Lang Tong. Impact of data quality on real-time locational marginal price. *IEEE Transactions on Power Systems*, 29(2): 627–636, 2013.

[79] Chensheng Liu, Jing Wu, Chengnian Long, and Deepa Kundur. Reactance perturbation for detecting and identifying fdi attacks in power system state estimation. *IEEE Journal of Selected Topics in Signal Processing*, 12(4):763–776, 2018.

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

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

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

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

[84] Mohammad Ashiqur Rahman, Ehab Al-Shaer, and Rakesh B Bobba. Moving target defense for hardening the security of the power system state estimation. In *Proceedings of the First ACM Workshop on Moving Target Defense*, pages 59–68, 2014.

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

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

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

[88] Zhenyong Zhang, Ruilong Deng, David KY Yau, Peng Cheng, and Jiming Chen. Analysis of moving target defense against false data injection attacks on power grid. *IEEE Transactions on Information Forensics and Security*, 15:2320–2335, 2019.

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

[90] Kumarsinh Jhala, Parth Pradhan, and Balasubramaniam Natarajan. Perturbation-based diagnosis of false data injection attack using distributed energy resources. *IEEE Transactions on Smart Grid*, 2020.

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

[92] Deepak Divan, William Brumsickle, Robert Schneider, Bill Kranz, Randal Gascoigne, Dale Bradshaw, Michael Ingram, and Ian Grant. A distributed static series compensator system for realizing active power flow control on existing power lines. In *IEEE PES Power Systems Conference and Exposition, 2004.*, pages 654–661. IEEE, 2004.

[93] Foad H Gandoman, Abdollah Ahmadi, Adel M Sharaf, Pierluigi Siano, Josep Pou, Branislav Hredzak, and Vassilios G Agelidis. Review of facts technologies and applications for power quality in smart grids with renewable energy systems. *Renewable and sustainable energy reviews*, 82:502–514, 2018.

[94] Lih-Hsing Hsu and Cheng-Kuan Lin. *Graph theory and interconnection networks*. CRC press, 2008.

[95] U. Agarwal and U. P. Singh. *Graph Theory*. Laxmi Publications, 2009.

[96] KM Rogers and Thomas J Overbye. Some applications of distributed flexible ac transmission system (d-facts) devices in power systems. In *2008 40th North American Power Symposium*, pages 1–8. IEEE, 2008.

[97] GR Krumpholz, KA Clements, and PW Davis. Power system observability: a practical algorithm using network topology. *IEEE Transactions on Power Apparatus and Systems*, (4):1534–1542, 1980.

[98] Adam B Birchfield, Ti Xu, Kathleen M Gegner, Komal S Shetye, and Thomas J Overbye. Grid structural characteristics as validation criteria for synthetic networks. *IEEE Transactions on power systems*, 32(4):3258–3265, 2016.

[99] Ronald L Graham and Pavol Hell. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7(1):43–57, 1985.

[100] Frank Harary. *A seminar on graph theory*. Courier Dover Publications, 2015.

[101] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems*, 26(1):12–19, 2010.

[102] Omar A Urquidez and Le Xie. Rectangular representation of facts devices in the acopf problem. In *2014 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2014.

[103] TC Subramanyam, JB Subrahmanyam, and T Ram. An adaptive chicken swarm algorithm to solve optimal power flow problem considering facts device. *Journal of Computational Mechanics, Power Syst. and Control*, 2(4):38–47, 2019.

[104] Haixiang Zhang, Jianan Liu, Dongliang Xiao, and Wei Qiao. Security-constrained optimal power flow solved with a dynamic multichain particle swarm optimizer. In *2019 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2019.

[105] Haixiang Zhang, Dongliang Xiao, Jianan Liu, and Wei Qiao. Security-constrained optimal power flow solved with a hybrid multiswarm particle swarm optimizer. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2019.

[106] Geraldo Leite Torres and Victor Hugo Quintana. An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates. *IEEE transactions on Power Systems*, 13(4):1211–1218, 1998.

[107] Geraldo L Torres and Manoel A De Carvalho. On efficient implementation of interior-point based optimal power flows in rectangular coordinates. In *2006 IEEE PES Power Systems Conference and Exposition*, pages 1747–1752. IEEE, 2006.

[108] Ray D Zimmerman. Ac power flows, generalized opf costs and their derivatives using complex matrix notation. *MATPOWER Technical Note 2*, 2010.

[109] Ren Jiangbo and Guo Zhizhong. Period-oriented state estimation approach for power system operational control. In *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*, pages 1–5. IEEE, 2005.

[110] Dan Li and Xiang Li. Decomposition-based global optimization for optimal design of power distribution systems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3265–3270. IEEE, 2016.

[111] Priya L Donti, Yajing Liu, Andreas J Schmitt, Andrey Bernstein, Rui Yang, and

Yingchen Zhang. Matrix completion for low-observability voltage estimation. *IEEE Transactions on Smart Grid*, 11(3):2520–2530, 2019.

[112] Bo Liu, Hongyu Wu, Yingchen Zhang, Rui Yang, and Andrey Bernstein. Robust matrix completion state estimation in distribution systems. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2019.

Appendix A

Graph-theory-based Topology Analysis

A.1 Proof of Proposition 3.3.1

Definition 3.2.1 suggests that $\Delta\mathbf{H}$ has full column rank in a complete MTD, i.e., $r(\Delta\mathbf{H}) = r(\mathbf{D}_1 \cdot \Delta\mathbf{X}' \cdot \mathbf{A}_1^T) = n-1$. Given the properties of matrix products, the following inequality holds:

$$\min\left\{r(\mathbf{D}_1), r(\Delta\mathbf{X}'), r(\mathbf{A}_1^T)\right\} \geq r(\mathbf{D}_1 \cdot \Delta\mathbf{X}' \cdot \mathbf{A}_1^T) = n-1$$

A.2 Proof of Lemma 3.3.1

The composite matrix is expressed as:

$$[\mathbf{H_0 H}] = \mathbf{D} \cdot [\mathbf{X_0}\ \mathbf{X}] \cdot \begin{bmatrix} \mathbf{A}_{-r}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{-r}^T \end{bmatrix} \tag{A.1}$$

Given the properties of matrix products, we have

$$r([\mathbf{H_0 H}]) \leq \min \left\{ r(\mathbf{D}), r([\mathbf{X_0}\ \mathbf{X}]), r(\begin{bmatrix} \mathbf{A}_{-r}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{-r}^T \end{bmatrix}) \right\} \tag{A.2}$$

We have $r([\mathbf{H_0 H}]) \leq \min\{p, p, 2(n-1)\}$. As p<2(n−1), it infers $\max\{r([\mathbf{H}_0\ \mathbf{H}])\} = p$.

A.3 Proof of Lemma 3.3.2

First, we prove that if G_{DF} is free of loops, there exists a non-zeros matrix $\mathbf{V}$, such that $\mathbf{H}_{DF}^0$=$\Delta\mathbf{H} \cdot \mathbf{V}$.

Since $\mathbf{H}_{DF}^0$and $\Delta\mathbf{H}$ share the same D-FACTS lines, $\mathbf{H}_{DF}^0$ can also be decomposed as $\mathbf{H}_{DF}^0 = \mathbf{D}_1 \cdot (\mathbf{X}_{DF}^0)' \cdot \mathbf{A}_1^T$, where $(\mathbf{X}_{DF}^0)' \in \mathbb{R}^{p_1 \times p_1}$ is a full rank diagonal matrix obtained from $\mathbf{X}_{DF}^0$ by removing rows and columns of all zeros. No loop in G_{DF} suggests $\mathbf{A}_1$ has full rank columns, i.e., $r(\mathbf{A}_1) = p_1$, implying that $\mathbf{A}_1^T\mathbf{A}_1$ is invertible.

Thus, there exists a matrix $\mathbf{V} \in \mathbb{R}^{n-1 \times n-1}$ defined as $\mathbf{V} = \mathbf{A}_1 \cdot (\mathbf{A}_1^T \cdot \mathbf{A}_1)^{-1} \cdot (\Delta\mathbf{X}')^{-1} \cdot (\mathbf{X}_{DF}^0)' \cdot \mathbf{A}_1^T$ satisfying:

$$\begin{aligned} \Delta\mathbf{H} \cdot \mathbf{V} &= \mathbf{D}_1 \cdot \Delta\mathbf{X}' \cdot \mathbf{A}_1^T \cdot \mathbf{A}_1 \cdot (\mathbf{A}_1^T \cdot \mathbf{A}_1)^{-1} \cdot (\Delta\mathbf{X}')^{-1} \cdot (\mathbf{X}_{DF}^0)' \cdot \mathbf{A}_1^T \\ &= \mathbf{D}_1 \cdot (\mathbf{X}_{DF}^0)' \cdot \mathbf{A}_1^T = \mathbf{H}_{DF}^0 \end{aligned} \tag{A.3}$$

It is seen that $\mathbf{H}_{DF}^0$=$\Delta\mathbf{H} \cdot \mathbf{V}$.Then, we can prove $r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r([\mathbf{H}_{\overline{DF}}^0\ \Delta\mathbf{H}])$. There exists an elementary column operation $\mathbf{T}_1 = \begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{V} & \mathbf{I} \end{bmatrix}$, which transforms $[\mathbf{H}_0\ \Delta\mathbf{H}]$ as follows:

$$[\mathbf{H}_0\ \Delta\mathbf{H}]\,\mathbf{T}_1 = \left[(\mathbf{H}_{\overline{DF}}^0 + \mathbf{H}_{DF}^0)\ \Delta\mathbf{H}\right] \begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{V} & \mathbf{I} \end{bmatrix} = [\mathbf{H}_{\overline{DF}}^0\ \Delta\mathbf{H}] \tag{A.4}$$

Since the rank of a matrix remains unchanged after an elementary column operation, $r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r([\mathbf{H}_{\overline{DF}}^0\ \Delta\mathbf{H}])$ holds.

Finally, we prove $r([\Delta\mathbf{H}\ \mathbf{H}_{\overline{DF}}^0]) = r(\mathbf{A}_1) + r(\mathbf{A}_2)$. Similar to (6), we can decompose matrix $\mathbf{H}_{\overline{DF}}^0$ as follows: $\mathbf{H}_{\overline{DF}}^0 = \mathbf{D}_2 \cdot (\mathbf{X}_{\overline{DF}}^0)' \cdot \mathbf{A}_2^T$, where $(\mathbf{X}_{\overline{DF}}^0)' \in \mathbb{R}^{p_2 \times p_2}$. Thus, $[\Delta\mathbf{H}\ \mathbf{H}_{\overline{DF}}^0]$

is expressed as:

$$[\Delta\mathbf{H}\ \mathbf{H}^0_{\overline{DF}}] = [\mathbf{D}_1\ \mathbf{D}_2]\begin{bmatrix}\Delta\mathbf{X}'\cdot\mathbf{A}_1^T & \mathbf{0}\\ \mathbf{0} & (\mathbf{X}^0_{\overline{DF}})'\cdot\mathbf{A}_2^T\end{bmatrix} \tag{A.5}$$

With $\mathbf{D}$, $\Delta\mathbf{X}'$ and $(\mathbf{X}^0_{\overline{DF}})'$ of full column rank, we have

$$\begin{aligned} r([\Delta\mathbf{H}\ \mathbf{H}^0_{\overline{DF}}]) &= r\left(\begin{bmatrix}\Delta\mathbf{X}'\cdot\mathbf{A}_1^T & \mathbf{0}\\ \mathbf{0} & (\mathbf{X}^0_{\overline{DF}})'\cdot\mathbf{A}_2^T\end{bmatrix}\right)\\ &= r(\Delta\mathbf{X}'\cdot\mathbf{A}_1^T) + r((\mathbf{X}^0_{\overline{DF}})'\cdot\mathbf{A}_2^T)\\ &= r(\mathbf{A}_1) + r(\mathbf{A}_2)\end{aligned} \tag{A.6}$$

Therefore, if there is no loop in the G_{DF}, $r([\mathbf{H}_0\ \Delta\mathbf{H}]) = r([\mathbf{H}^0_{\overline{DF}}\ \Delta\mathbf{H}]) = r(\mathbf{A}_1^T) + r(\mathbf{A}_2^T)$ holds.

A.4 Proof of Corollary 3.3.1

According to Lemma 3.3.2 and (3.14), if there is no loop in the G_{DF}, we have $r(M) = p - lp_{DF} - lp_{\overline{DF}}$. Since both the G_{DF} and the $G_{\overline{DF}}$ are free of loops, i.e., $lp_{DF} = lp_{\overline{DF}} = 0$, we have $r(M) = p$. Thus, the MTD constructed by G_{DF} is a max-rank incomplete MTD according to Lemma 3.3.1.

A.5 Proof of Corollary 3.3.2

If there is no loop in the G_{DF}, $r(M) = 2n - (t_{DF} + t_{\overline{DF}})$ holds. If both the G_{DF} and the $G_{\overline{DF}}$ are connected graphs, we have $t_{DF} = t_{\overline{DF}} = 1$. Thus, $r(M) = 2(n-1)$ holds, indicating that an MTD constructed in the G_{DF} is a complete MTD.

Appendix B

Reuse permissions from publishers