Transactions on Computer Systems and Networks

Amulya Sreejith
K. Shanti Swarup

# Cyber-Security for Smart Grid Control

## Vulnerability Assessment, Attack Detection, and Mitigation

Springer

# Transactions on Computer Systems and Networks

Transactions on Computer Systems and Networks is a unique series that aims to capture advances in evolution of computer hardware and software systems and progress in computer networks. Computing Systems in present world span from miniature IoT nodes and embedded computing systems to large-scale cloud infrastructures, which necessitates developing systems architecture, storage infrastructure and process management to work at various scales. Present day networking technologies provide pervasive global coverage on a scale and enable multitude of transformative technologies. The new landscape of computing comprises of self-aware autonomous systems, which are built upon a software-hardware collaborative framework. These systems are designed to execute critical and non-critical tasks involving a variety of processing resources like multi-core CPUs, reconfigurable hardware, GPUs and TPUs which are managed through virtualisation, real-time process management and fault-tolerance. While AI, Machine Learning and Deep Learning tasks are predominantly increasing in the application space the computing system research aim towards efficient means of data processing, memory management, real-time task scheduling, scalable, secured and energy aware computing. The paradigm of computer networks also extends it support to this evolving application scenario through various advanced protocols, architectures and services. This series aims to present leading works on advances in theory, design, behaviour and applications in computing systems and networks. The Series accepts research monographs, introductory and advanced textbooks, professional books, reference works, and select conference proceedings.

Amulya Sreejith · K. Shanti Swarup

# Cyber-Security for Smart Grid Control

## Vulnerability Assessment, Attack Detection, and Mitigation

Springer

Amulya Sreejith
Department of Electrical Science
Kansas State University
Manhattan, Kansas, USA

K. Shanti Swarup
Department of Electrical Engineering
Indian Institute of Technology—Madras
Chennai, Tamil Nadu, India

ISSN 2730-7484 ISSN 2730-7492 (electronic)
Transactions on Computer Systems and Networks
ISBN 978-981-97-1301-1 ISBN 978-981-97-1302-8 (eBook)
https://doi.org/10.1007/978-981-97-1302-8

# Preface

The book focuses on a very important area of Smart Grids—Cyber-Security. It deals in particular with the tools and techniques for cyber-security analysis of the Smart Grid Control systems. This includes the standards and guidelines, detailed vulnerability assessment framework, attack detection strategies, and attack mitigation methods.

The book is divided into three parts. The smart grid cyber-physical system is discussed in Part I. Part II introduces the attacks in the grid system and a vulnerability assessment framework followed by a tool that can be used to analyze the grid control systems using existing cyber-security standards. In Part III different forms of attack detection methods are discussed along with Python-based implementations for the same. Finally, attack mitigation methods are discussed with implementation.

Detailed illustrations and tables are provided in each part. The book also includes case studies based on standard test systems thus helping students to implement the discussed methods. The case studies are based on MATLAB and Python implementations thus catering to a wide range of audience. Outputs and programs are included so students can compare their results and improve upon the discussed methods.

The book can be useful to a wide variety of audiences. The primary audience will be students and researchers in Smart Grids. Students can gain in-depth knowledge about various areas of smart grid cyber-security and use the tools and methods to build secure cyber systems. They can also use the program and methods as a base to build upon their research. In addition to students, power system operators can use the book as a reference guide for analyzing the security of their systems using the discussed vulnerability assessment tools and methods. Device manufacturers can use the detection methods to build devices with cyber-security capabilities. The suggested mitigation and response can be used to build a framework for the systems to recover back to normal state in the event of an attack.

Manhattan, USA
Chennai, India

Amulya Sreejith
K. Shanti Swarup

# Contents

# About the Authors

**Dr. Amulya Sreejith** is presently a Postdoc Fellow in the Department of Electrical and Computer Engineering, the Kansas State University, Manhattan, KS. She has completed her Ph.D. from the Indian Institute of Technology Madras, India, in the area of Cyber-Security in Power System Control. She has done her M.Tech. from VNIT Nagpur. Before joining as a postdoc fellow she worked as an R&D scientist at Gridsentry, a power cyber-security startup. Her area of interest is in Electrical Power Systems, Automation and Protection, Cyber-Systems Security, and Artificial Intelligence for power systems applications.

**K. Shanti Swarup** is a faculty with the Department of Electrical Engineering, Indian Institute of Technology (IIT) Madras, India. Before joining the department as a visiting faculty member, he held positions at the Mitsubishi Electric Corporation, Osaka, Japan, and Kitami Institute of Technology, Hokkaido, Japan, serving as a visiting research scientist and visiting professor, respectively, from 1992 to 1999. Since 2000, he has been a professor at IIT Madras. His research areas include power systems, smart grids, artificial intelligence, knowledge-based systems, computational intelligence, soft computing, Energy Management Systems (EMS), Supervisory Control and Data Acquisition (SCADA), power system automation, and network protection. He has done research projects with various industries like BHEL, Hitachi, Easun-MR, etc.

# Nomenclature

## Constants

| | |
|---|---|
| $\beta$ | Bias factor |
| $D$ | Load constant |
| $H$ | Inertia constant |
| $K$ | Number of columns of $T$ |
| $L$ | Number of rows of $T$ |
| $M$ | Number of sensors |
| $N$ | Number of initial training samples |

## Indices

| | |
|---|---|
| $a$ | Attack variables |
| $i$ | Area under consideration |
| $j$ | Other areas |
| $k$ | Time sample |
| $m$ | Sensor |
| $s$ | Signal subspace |

## Matrices

| | |
|---|---|
| $\mathbf{A}$ | State Matrix |
| $\mathbf{B}$ | Input Matrix |
| $\mathbf{C}$ | Output Matrix |

| | |
|---|---|
| **D** | Feedthrough Matrix |
| **P** | Projection Matrix |
| **V** | Measurement Noise Matrix |
| **W** | Process Noise Matrix |

## Variables

| | |
|---|---|
| $\beta$ | Theoretical false positive rate |
| $\beta_p$ | FDIA Bernoulli variable |
| $\beta_r$ | Ramp attack Bernoulli variable |
| $\beta_s$ | Scaling attack Bernoulli variable |
| $\Delta f$ | Frequency change |
| $\Delta P_{tie}$ | Tie-line power change |
| $\hat{y}$ | Output estimate |
| $\Lambda$ | Monitor input |
| $\nu$ | Measurement noise |
| $\omega$ | Process noise |
| $\psi$ | Monitor output |
| $\tau$ | Threshold |
| **d** | Disturbance input vector |
| **T** | Trajectory matrix |
| **U** | Eigenvectors |
| **u** | Controlled input vector |
| **w** | Equivalent test vector of all sensor values |
| **x** | State vector |
| **y** | Output vector |
| **z** | Column of $T$ |
| ACE | Area Control Error |
| Cov | Covariance matrix |
| $c$ | Centroid of cluster of projected data |
| $D$ | Distance for detection |
| $P_g$ | Generation |
| $P_l$ | Load |
| $P_m$ | Mechanical power output |
| $r$ | Residue |
| $r_{LO}$ | Line outage index |
| $r_{LS}$ | Load shedding index |
| $S^s$ | Signal subspace |
| $T_{ij}$ | Synchronizing torque coefficient |
| $V$ | Bus voltage |
| $z$ | Sensor sample value |

# Part I
# Cyber-Physical Smart Grid Systems

# Chapter 1
# Smart Grid Cyber-Physical System: An Overview


**Abstract** This chapter gives an introduction to the Smart Grid cyber-physical system and the various attacks that can be injected into the grid at various levels. The introduction of automation and control improves the grid stability and performance but also gives rise to various attack points. The power system is built to transmit data at a very fast rate and due to this speed requirement the security features such as encryption and key management are often not available for grid protocols such as DNP3, MODBUS, and IEC-61850.

**Keywords** Cyber Physical Power Systems (CPPS) · Defense in depth · Smart grids · SCADA

## 1.1 Introduction

The electrical power system came into existence around the early 1900s, and automation has been in the grid system since the late 1960s. There has been a constant increase in the demand for energy and electricity, mainly due to the rise in industries and ever-changing lifestyles. Energy Management Systems (EMS) are used to manage the daily operations of the grid. In the late 1960s, digital computers and software were developed to replace the analog EMS thus giving rise to today's digital grids. The digital grids differ from traditional Supervisory Control and Data Acquisition (SCADA) EMS or Demand Management Systems (DMS). For example, we have the process bus in digital substations, connecting directly to the primary equipment like the optic CTs and the merging unit. Some other digital concepts getting into the grid are centrally located disturbance recorders, intelligent digital assets, and analytical and machine learning (ML) tools for asset management, demand forecasting, and generation scheduling. Thus, we see that today's power system involves many data exchanges between different systems, and all these systems are widely distributed. There is also significant Information Technology-Operations Technology (IT-OT) convergence in the present day's power grid wherein the control system environment is exposed to the consumer's IT system environment.



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_1

Hence, the considerable development in the grid and the overlay of IT systems over OT systems have created more access points, leading to an explosion in the number of attack points. According to the French think tank Institut Français des Relations Internationals (IFRI), the power sector has become a prime target for cybercriminals in the last decade, with cyber attacks surging by 380% during 2014–15. The various motives behind these include geopolitics, sabotage, and financial reasons. Thus, for efficient and undisrupted grid operation, addressing the effects of these cyber-vulnerabilities becomes critical.

## 1.2 Smart Grid Cyber-Physical System

### *1.2.1 Smart Power Grids*

The conventional grid systems have undergone tremendous changes and have moved into a more intelligent paradigm where it uses state-of-the-art technologies to drive the grid system. This intelligence built into the power system gives rise to the term *Smart Grids*. The U.S. Department of Energy (DoE) had established a Federal Smart Grid Task Force in 2007 which involved a vision-2030 to construct a self-sufficient and smart electric system to provide affordable, clean, efficient, and reliable electric power. In addition to improving the reliability and quality, the introduction of renewables is also taken up as a major objective.

The traditional grid systems only consisted of generators, transmission lines, various loads, and transformers along with controllers that coordinated between these components. On the other hand, the smart grid paradigm introduces new technologies such as wide-area monitoring and control, grid optimization, and real-time protection. These technologies can be termed as the cyber system of the smart grid. The smart grid evolves through various interactions among its components. These interactions can be divided mainly into the following three levels:

1. **Level 1**: Interaction between power system components (generator, transmission lines, loads, transformers) and the grid controller. The controller gets grid data, calculates the actuator signals, and sends them back to the grid to maintain grid operation.
2. **Level 2**: Interaction between power system controller and communication. Communication provides a link between different subsystems and coordinates the EMS functions.
3. **Level 3**: Interaction between communication and cyber system.

To ensure an efficient and secure power grid operation, both power flow and information flow play an equal role. Information and Communication Technology (ICT) is introduced in the system for the safe and secure operation of the power grid. However, failure or maloperation of the ICT system has also been a major reason for grid blackouts.

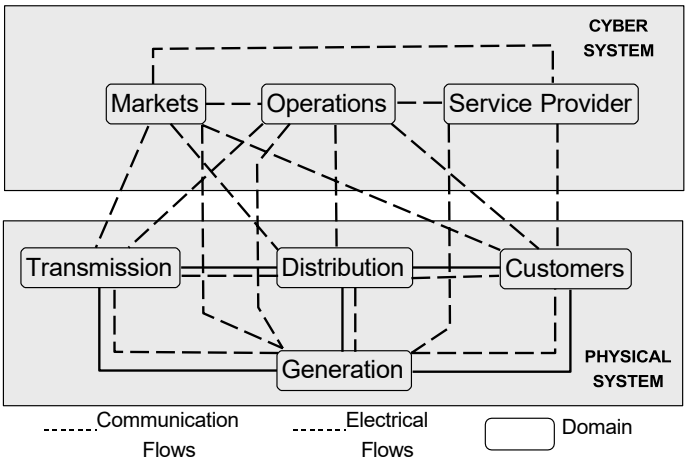



**Fig. 1.1** Cyber-physical power system

### *1.2.2 Cyber-Physical Systems*

In today's digital world, most devices have some computing capabilities within them. The electrical grid system is also a combination of physical devices with various control and computing capabilities. This combination of the physical and computing layers is called a Cyber-Physical system.

The electrical grid is one of the most complicated and wide-area cyber-physical systems. It consists of various control systems to maintain stability and load demand. Automation and control are implemented at all stages of the power system, including generation, transmission, and distribution. This combination of physical grid equipment with cyber and control systems gives rise to what is called a Cyber-Physical Power System (CPSS) (Yohanandhan et al. 2020).

The Cyber layer is typically a combination of computation and control with communication. The physical layer is made up of interconnected electrical equipment that operate on the principles of physics. The cyber layer consists of software and programs that guide the physical system operations. The two layers interact with each other using a Communication layer. Figure 1.1 shows the schematic of the two layers and their interaction. The various domains and the interactions between them are from the NIST Smart Grid Framework, (Greer et al. 2014).

The various equipment and tasks involved in each layer are given below.

#### 1.2.2.1 Physical Layer

The physical layer involves the conventional smart grid components such as generators, transmission lines, transformers, and loads. Physical sensors and actuators are

used for controls and form a part of the physical layer. The merging unit usually picks these analog signals and converts them to digital for data transmission. Actuators are used to send back control signals to the equipment so as to operate them.

#### 1.2.2.2 Communication Layer

The communication layer involves various data transmission media like cables, switches, and routers. The communication protocols commonly used in power systems are IEC-61850, MODBUS, DNP3, and IEC-60870-5. These protocols are particularly designed for efficient and fast data exchange and therefore usually lack the security features like encryption.

#### 1.2.2.3 Cyber Layer

The cyber layer involves the programs used to maintain efficient and reliable grid operation. State estimation, Volt-VAr controls, demand response programs, and other control systems are a part of the cyber layer. The cyber layer processes the data obtained from the physical layer and sends back the signal to the physical layer. This completes the loop in the smart grid cyber-physical system operation.

## 1.3 Issues in Smart Grid Cyber-Physical Systems

The introduction of control and automation into the grid has several benefits as discussed in the previous sections. However, the introduction of Smart Grid Cyber-Physical System (SG-CPS) also has several challenges that are yet to be resolved. Some of the most significant challenges of a cyber-physical power system are as follows:

1. The available cyber-physical technologies have to be tailored to suit the power system and thus cannot be directly fitted on to an existing system.
2. Distributed control faces several issues such as time delays, packet drops, and errors.
3. The messages in the smart grid are time-critical which means that the cyber-physical system and protocols should have low latency.
4. Smart grid is a market-driven supply-demand system. This gives rise to competition and game-based transfer between various market participants resulting in severe network congestion.
5. The communication channel should be improved for application to real-time dynamic situations.
6. Smart grid involves several uncertainties due to the introduction of renewables. The CPS algorithms should be adaptive to such uncertainties.

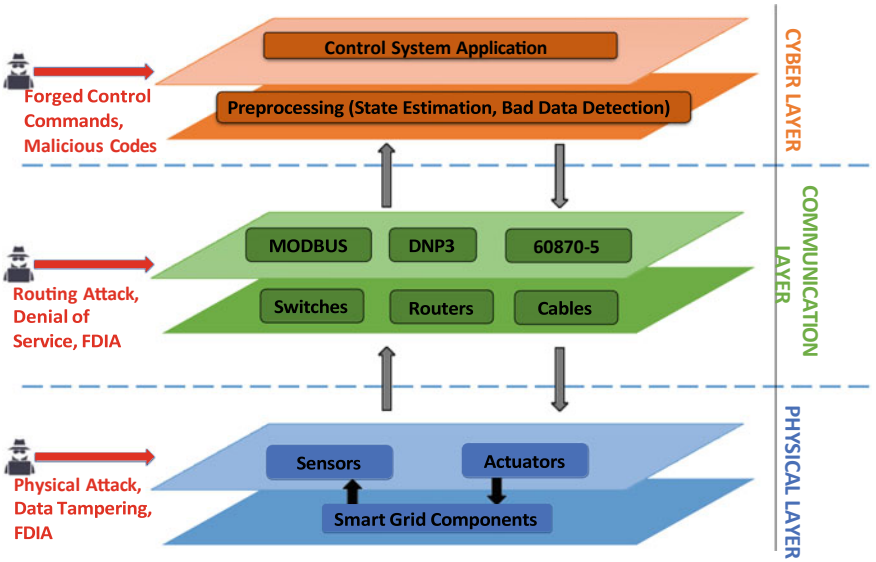



**Fig. 1.2** Smart grid cyber-physical system with attack surface

7. Smart Grid is not a standalone system since it affects other critical systems. Thus, it has to be analyzed in conjunction with other environmental and social systems.
8. Since the availability of data is more important than the security of the data to maintain operations, the smart grid CPS is prone to cyber attacks.

Many of the above challenges have been discussed in the literature and industry and solutions are still being formulated for most of the issues. The security of power system was not an important concern until recent years. However, over the past few years, it has gained importance due to the multiple demonstrated attacks on the power system.

Figure 1.2 shows the different layers of the SG-CPS. It also illustrates the various attacks at each of these layers which will be discussed in the next section.

## 1.4 Attacks on Smart Grid Systems

Increased ICT integration exposes the system to the risk of cyber-vulnerabilities. Consequently, resolving these vulnerabilities is essential for enhancing grid efficiency. Figure 1.2 depicts the possible attacks at various Smart Grid control system tiers and can be explained as follows:

1. Physical Layer: The attack targets equipment at the process level where physical settings can be adjusted (Physical Attack) or measurements can be inserted into different devices in the field (False Data Injection Attack (FDIA)/Data tampering).
2. Communication Layer: An attack enters the telecommunication channel. The channel may be saturated with incorrect data to disrupt data flow (FDIA), or the communication path may be stopped (Denial of Service (DoS)). The protocols like MODBUS and DNP3 lack security features such as encryption as the importance is on the availability and speed of data transmission.
3. Cyber layer: An attack is launched into the power grid control system to interrupt overall system processing. The settings of relays or algorithms can be altered (Forged control commands), or malicious codes may be introduced into the control systems to disrupt the normal course of action.

An attacker can have a variety of objectives and effects while launching an attack. These can be divided into the following two broad categories:

1. Economic impact: This may result in monetary losses for the grid system operators and the utility, while the hacker may profit monetarily from the attack.
2. Stability impact: These attacks can cause frequency variations, generation-load imbalances, and sequential outages.

Until recent years, the Operational Technology systems, especially the power grid, were considered immune to attacks due to the existence of dedicated communication channels. However, the cyber-attack events over the past few years have revealed that the power system is also highly prone to attacks. A few sample attacks on power systems are described in Table 1.1.

There is extensive research in the area of cyber attacks on power systems. These researches mainly focused on State Estimation (Liu et al. 2009). Such research is necessary because even a minor attack can traverse the grid system and create considerable consequences that lead to blackouts. Stability attacks can result in substantial system damage. As they are intended to maintain the grid's stability, Grid control systems might be attractive attack targets. High excursions in grid parameters can cause generators to lose synchronization, resulting in disastrous results.

## 1.5 Defense in Depth Security Approach

Defense in Depth (DiD) is an approach involving multiple layers of security countermeasures to protect the integrity of the information or operation network as shown in Fig. 1.3. DiD reduces the probability of an intruder succeeding in penetrating the system. It can also be used to identify attackers attempting to tamper with the system. Hence, if an attacker gains access to a system, the DiD security approach can considerably delay the intended harm and give system operators enough time to employ countermeasures. This will help in either preventing or at least minimizing the impact of attacks (David and Mark 2006; Smith et al. 2018).

**Table 1.1** History of attacks on power systems

| Incident | Attack type | Details | Impact |
|---|---|---|---|
| Iran Nuclear plant (Kushner 2013) | Stuxnet | Initiated by a worker's USB drive. Stuxnet worm targeted the PLC systems in Iran's nuclear program, causing centrifuges to spin out of control without triggering alarms | Destroyed 984 uranium enriching centrifuges |
| Elexon (Winder 2020) | Ransomware | Attack on IT system. Since Elexon is a power administrator, the attack did not impact the power supply | Has potential to damage if it penetrates to power lines |
| Ukraine Grid attack (Liang et al. 2017) | BlackEnergy (Spear-phishing) | Spear-phishing emails with **BlackEnergy malware**. SCADA signaling and remotely switching off substations. Disabling infrastructure components. File destruction **KillDisk malware**. **Denial-of-Service** attack on call-center | 230000 people went without electricity for 6 h |
| US Power Grid attack (INL 2016) | Dragonfly Trojan | Operators at a power control center started losing communication with "multiple remote power generation sites" for minutes at a time | Immediately detected and hence no losses |
| Enel, Italian Energy Company | EKANS Ransomware attack | Attack on ICS kill all processes and encrypt all files. It is a new strain from MEGACORTEX | Not disclosed |

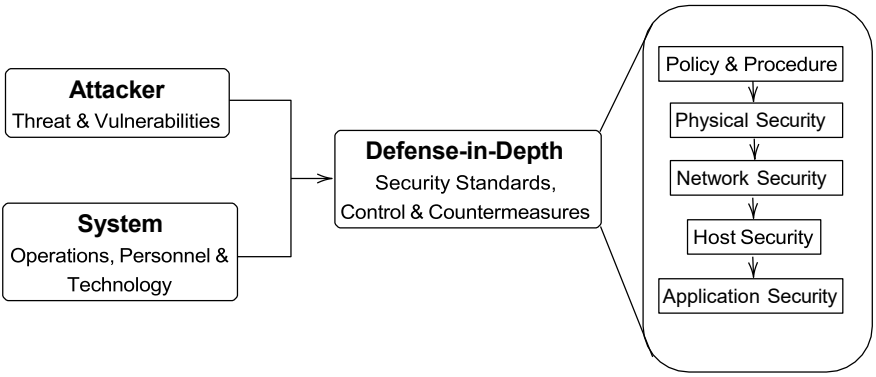



**Fig. 1.3** Defense in depth planning

**Table 1.2** Comparison of IT, OT, and electrical systems

| Parameter | IT system | OT system | Electrical system |
|---|---|---|---|
| Primary risk | Information and finance | Safety and health | Safety and life |
| Acceptable downtime (per year) | 18 − 4 days | 9 h − 5 mis | 30 − 5 min |
| Asset lifespan (years) | 3–5 | 15–30 | 10–20 |
| Problem response | Offline rebooting/patch management | Online recovery | Online recovery/replacement |

To effectively apply DiD security to any system, it is essential to identify the relationship between attackers (threats) and system vulnerabilities. This analysis helps to design suitable standards and employ countermeasures to protect the operations, personnel, and technologies that make up an OT system. Additionally, security countermeasures must be constantly refined to ensure protection against new attacks also called as zero-day attacks.

The IT and OT systems are quite different from each other in the cyber-security aspect, as explained in Table 1.2.

The bulk transmission system is a type of OT system but it is more critical compared to most of the other OT systems. Any outage in the electrical facilities impacts multiple other systems and can be hazardous to life and safety. All voltage security violations are usually given a resolution time of 30 min, which is shown as the allowable downtime. However, for critical systems, the downtime allowed is further low. Most of the equipment in the power sector are costly due to which replacement is not an option leading to the use of old equipment. In the event of a vulnerability being found, the resolution has to be done online or equipment replacement is required due to which it is often ignored leading to a larger attack surface.

Thus, it is important to devise tailor-made approaches specific to each system, considering the time of response of an entire process. As shown in Fig. 1.3, several layers of defense need to be established.

1. Policy and Procedure: This involves that the system, organization, and personnel follow certain rules set forth using cyber-security standards and protocols. Table 1.3 compares various cyber-security standards dealing with the power system.
2. Physical Security: Role-based access passwords and physical zone separation can be used for protecting physical field devices.

**Table 1.3** SAS relevant cyber-security documents

| Authority | Document/tool | Domain | Details |
|---|---|---|---|
| National Institute of Standards and Technology (NIST) | NIST-special publication-800-30 guide for conducting risk assessments (Division 2012) | Critical infrastructure | Risk assessments of federal information systems and organizations |
| | NIST-Special Publication-800-82 Guide to Operational Technology (OT) Security (Stouffer et al. 2022) | OT | Industrial control systems (ICS; risk management; security controls; SCADA systems |
| North American Electric Reliability Corporation (NERC) | NERC-critical infrastructure protection (NERC-CIP) CIP-002 to CIP-014 | Bulk electric systems | Set of cyber-security standards to reduce the risk of compromise to electrical resources |
| Cyber-security and Infrastructure Security Agency (CISA) | CISA Cyber-Security Evaluation Tool (CSET) CISA ICS alerts, advisories, and reports | ICS, OT | Tools and reports for vulnerability and security assessments. |
| MITRE | ATT&CK for ICS (Alexander et al. 2020) | ICS | Threat modeling tool |
| Electric Power Research Institute (EPRI) | National Electric Sector Cyber-security Resource (NESCOR) (NESCOR 2015; Searle et al. 2016) | Electric sector | Tools and guidelines for electric sector Penetration testing guidelines, attack, failure scenarios and mitigation for electric sector |
| International electrotechnical commission | IEC TS 62443-1 to 62443-4 (IEC 2019) | Industrial automation systems | Secure industrial automation and control systems (IACS) throughout their lifecycle |

3. Network Security: This can be ensured using logical segmentation using demilitarized zones, deception technology to hide actual network, and network monitoring tools. Firewalls can also protect against external attacks.
4. Host Security: Host can be protected using application dynamic safelists, memory protection, and read/write protection.
5. Application Security: Application security can be ensured using Intrusion Detection System (IDS) and cryptographic techniques.

Figure 1.4 shows the different layers of a grid system with the implementation of DiD security measures.

The increased number of attacks over the past few years has shown that the attackers are highly sophisticated and technologically advanced. Thus, the operators and the grid system should also handle such attacks and be prepared to respond and recover fast enough to combat these sophisticated attacks. Several standards have been developed to deal with the cyber-security issues in the power system, which can effectively build a defense in depth approach to system cyber-security. However, the attackers' capabilities are also improving. Thus, it is necessary to have detection techniques at the application level. Application level detection can handle situations where an attack can bypass undetected through the various defense levels.

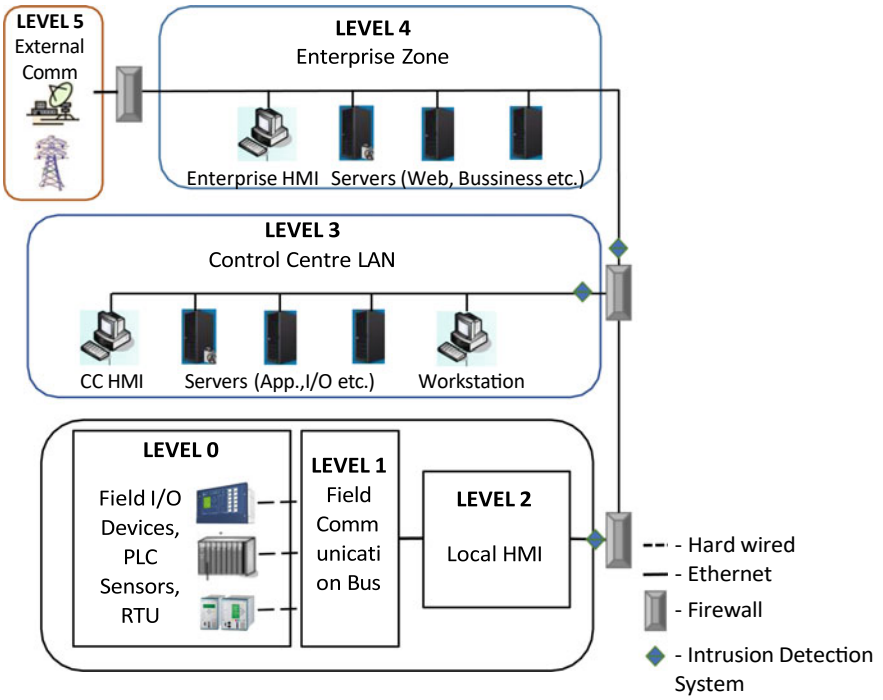



**Fig. 1.4** Defense in depth security of grid system

## 1.6 Cyber-Security in Smart Grid Control

There are a number of control systems built into the grid system in order to maintain the stability and power requirements of the grid. Additionally, protection equipment are also used to protect the system in case of high excursions of voltage, line flow, frequency, etc. While these control systems help in the fast recovery of the grid, they can also be attractive targets for attackers.

Security controls have been built into the control grid systems at various levels based on standards and guidelines proposed by standard organizations. In many cases, guidelines published by an organization are considered comparable in significance to published standards and these are used by utilities to secure their control systems.

The power grid is made up of various control systems such as voltage control, frequency control, reactive power control, and protection systems. Since the analysis of all the control systems is not feasible, the thesis focuses on one of the control systems, the Load Frequency Control (LFC). LFC is used as the representative control system as any attack into this control system affects the stability of the entire power grid and could also lead to blackouts. Therefore, it can be a suitable target for the attackers. The results presented in the thesis can also be extended to other control systems.

## References

International Electrotechnical Commission (2019) Industrial communication networks-Network and system security

Alexander O, Belisle M, Steele J (2020) Mitre att&ck for industrial control systems: design and philosophy. MITRE

David K, Mark F (2006) Control systems cyber security: defense in depth strategies. US Department of Homeland Security

Division CS (2012) Guide for conducting risk assessments. National Institute of Standards and Technology (NIST), pp I1

Greer C, Wollman D, Prochaska D, Boynton P, Mazer J, Nguyen C, FitzPatrick G, Nelson T, Koepke G, Jr Hefner A, Pillitteri V, Brewer T, Golmie N, Su D, Eustis A, Holmberg D, Bushby S (2014) Nist framework and roadmap for smart grid interoperability standards, release 3.0

INL (2016) Cyber threat and vulnerability analysis of the U.S. Electric Sector. Idaho National Laboratory

Kushner D (2013) The real story of stuxnet. IEEE Spectr 50(3):48–53

Liang G, Weller SR, Zhao J, Luo F, Dong ZY (2017) The 2015 Ukraine blackout: implications for false data injection attacks. IEEE Trans Power Syst 32(4):3317–3318

Liu Y, Ning P, Reiter MK (2009) False data injection attacks against state estimation in electric power grids. In: Proceedings of the 16th ACM conference on computer and communications security, CCS'09, pp 21–32

NESCOR (2015) Electric sector failure scenarios and impact analyses-version 3.0. National Electric Sector Cybersecurity Organization Resource, p 8

Searle J, Rasche G, Wright A, Dinnage S (2016) Nescor guide to penetration testing for electric utilities. National Electric Sector Cybersecurity Organization Resource, p 8

Smith J, Kipp N, Gammel D, Watkins T (2018) Defense-in-depth security for industrial control systems. Sensible cybersecurity for power systems, A collection of technical papers representing modern solutions, p 2

Stouffer K, Michael Pease CT, Zimmerman T, Pillitteri V, Lightman S (2022) Guide to operational technology (ot) security. National Institute of Standards and Technology (NIST)

Winder D (2020) Cyber attack on u.k. electricity market confirmed: national grid investigates

Yohanandhan RV, Elavarasan RM, Manoharan P, Mihet-Popa L (2020) Cyber-physical power system (cpps): a review on modeling, simulation, and analysis with cyber security applications. IEEE Access 8:151019–151064

# Chapter 2
# Smart Grid Control


**Abstract** This chapter offers a comprehensive overview of the progressive developments in power system control within the dynamic landscape of smart grids. It discusses in detail the various control systems present in a power system and the importance of frequency control. Load Frequency Control (LFC) is used as the representative control system to analyze the various proposed algorithms. A detailed analysis of the Multi-Area LFC system model is derived based on the machine and system equations, and a state-space model is developed to represent the MA-LFC system. Since the system is a linearized one, we also introduce the nonlinearities. The model derived in this chapter will further be used as a representative system to build and analyze various algorithms in future chapters.

**Keywords** Power system control · Load frequency control · Power system modelling · MODBUS · ICCP

## 2.1 Introduction

A grid system consists of various control systems to maintain stability and demand. This combination of physical grid equipment with cyber and control systems gives rise to a Cyber-Physical Power System (CPSS) (Yohanandhan et al. 2020). A grid system consists of physical and cyber layers that interact using a Communication layer. Some of the primary control systems are as follows:

1. Frequency Control: Imbalance between the generation and load directly affects the electric oscillation frequency of the grid.
2. Voltage Control: Voltage Control is indirectly affected by the reactive power in the grid. The automatic voltage regulator (AVR) controls the voltage at the generators.
3. Protection Control: Protective relays and protection controls prevent system damages and other incidents while considering shorter time scales to preserve system stability.



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_2

4. Demand Response Control: Demand response is effective when there is insufficient power to meet demand or in special cases where it is economically impractical to generate more energy.
5. Microgrid Control: It refers to the control component that responds to concerns other than those listed previously. Traditional decentralized controls are ineffective in Microgrids.

All of the above control systems depend on each other and work in unison at different time scales in order to maintain the stability of the power system.

This chapter discusses the grid control systems and the motivation for choosing Load Frequency Control (LFC) as the case study for research. A detailed analysis of the Multi-Area LFC (MA-LFC) will be discussed.

## 2.2 Smart Grid Control and Cyber-Security

Stability is that aspect of a system that helps it to reach or maintain the desired value despite disturbances. Consequently, power system stability can be defined as the capacity of a power system to reestablish operating balance following a physical disruption. Depending on the variables to be monitored and the type of disruption, various definitions of stability exist, such as transient stability, small signal stability, and voltage collapse. Cyber attacks that target these aspects can lead to instability or even the collapse of an entire grid system.

### *2.2.1 Smart Grid Control*

Different control algorithms are implemented in the power system at different time scales for maintaining performance and stability, as shown in Fig. 2.1.

In addition to these control systems, there exist other controls for market operation, Distributed control, Wide Area Control, and Cyber-Physical security and control (Annaswamy and Amin 2013).

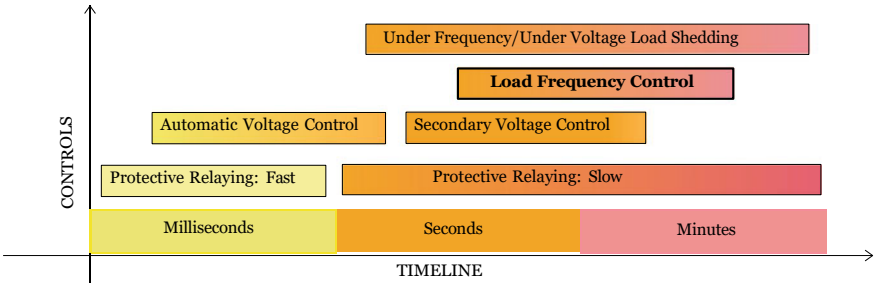



**Fig. 2.1** Smart grid controls and timescales

Most continuous control loops, such as prime mover and excitation controls, are installed within power plants and act locally on the generating unit. The continuous online controls consist of generator excitation controls (Power System Stabilizer-PSS and automatic voltage regulator-AVR), prime mover controls, reactive power controls, and HVDC controls. Typically, each control is linear, constantly operational, and employs localized sensors.

Excitation control regulates the governor voltage and reactive power output in a power plant, while prime mover controls manage energy supply system characteristics and speed. Automatic generation control (AGC) balances total generation and load (including losses) to achieve the nominal grid frequency and the planned power exchange with adjoining networks.

The discontinuous controls stabilize the system whenever there are significant disturbances and are appropriate under high-stress conditions. They execute generator and load tripping, capacitor and reactor switching, and additional protection strategies. Such electric grid controls may be localized at power plants and substations, or they may span a large geographical area. Typically, these controls guarantee a post-disturbance balance with an adequate region of attraction. Discontinuous controls give rise to additional controls, precise stability controls, and emergency control/protection schemes.

### *2.2.2 Cyber-Security in Smart Grid Control*

Smart Grid Control system can be attractive choice for the attackers as it tampers the system operational security. Distributed control systems operate based on the data from each individual controller and also neighboring control centers. This data transfer makes these systems vulnerable to cyber attacks.

In voltage controllers, the attacker can change the voltage measurement causing the tap settings to change inadvertently. If the voltage is higher than usual, the system operates at a higher voltage which is unnecessary. Voltage drops could also cause voltage collapse in the system.

Other types of attacks could change the voltage and reactive power values sent to the Volt-VAr controllers. This could lead to wrong commands to the FACTS and other controlled reactors. Such attacks could lead the system to function at very low power factors and also could impact the ancillary service actions in the system.

Attacks to the load frequency control and automatic generation control impact the grid frequency and also the generation-load balances in the system. It can also impact the economy due to wrong generation schedules.

The frequency control system is usually centrally controlled. Any disturbance in one part of the control system can propagate through the entire grid system and cause widespread impacts, including blackouts. In this work, we use the Load Frequency Control as the representative control system to build and analyze our algorithm.

## 2.3 Frequency Control

Severe network stress resulting in inequality among sources and loads significantly impairs the stability of any power system. Such a form of a typically slow phenomenon is to be studied in relation to frequency control issues in the power system.

Frequency deviation is a direct outcome of an imbalance between the electrical load and the power supplied by the linked generators; therefore, it serves as a helpful indicator of the imbalance between generation and load. Prolonged frequency deviations can impact a power system's operation, security, dependability, and efficiency by causing equipment damage, reducing load performance, overloading transmission lines, and triggering protection devices.

Frequency Deviation has the following effects on the system operation:

1. Since frequency is a function of the generator rotational speed, frequency control is indirectly a generator-turbine speed-control problem.
2. Large deviations in frequency can degrade load performance, damage equipment, and impair protection mechanisms.
3. The overall system stability is affected.

Since frequency deviations have multiple effects, there are multiple frequency control loops in the system, as shown in Fig. 2.2.

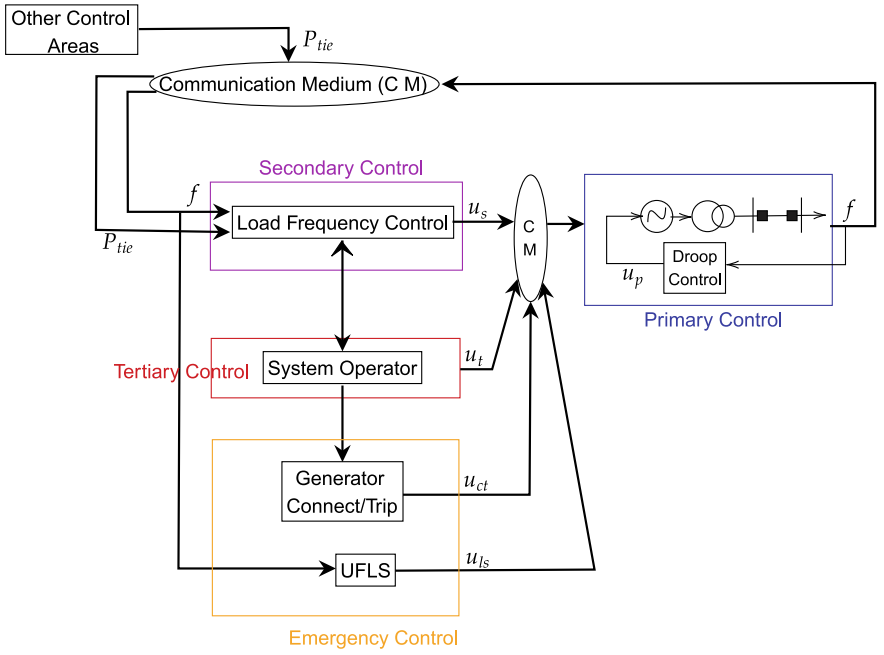



**Fig. 2.2** Power system frequency control loops

Minor deviations in frequency can be adjusted by adjusting the turbine input using the governor droop control called the *Primary Control*. The *Secondary Control* comes into the picture when the generation is adjusted for an entire control area based on the available reserve. The secondary control, called the Load Frequency Control, maintains the frequency and power exchanges at rated values. The *tertiary control* is at the system operator level and includes the market variables to adjust the generator setpoints, participation factor, and dispatches. Tertiary controls are also activated in case of large frequency deviations that the LFC cannot handle. *Emergency Controls* such as generator tripping or Underfrequency Load Shedding (UFLS) are introduced when there is a substantial frequency deviation due to faults in the system. Such emergency controls prevent the system from moving into a state of blackout.

The primary and secondary controls are the fundamental frequency control systems, which can be modeled using the system and machine parameters. Such a detailed model can be further used to analyze the system dynamics and effects of attacks in these systems.

## 2.4 Load Frequency Control Modeling

The LFC model is developed in the below section with reference to Bevrani (2014), Wood et al. (2013). As discussed in the previous sections, the generation ($P_m$)-load ($P_l$) imbalance ($\Delta P_m(t) - \Delta P_l(t)$) has a direct impact on the frequency,

$$\Delta P_m(t) - \Delta P_l(t) = 2H\frac{d\Delta f(t)}{dt} + D\Delta f(t) \tag{2.1}$$

Since only positive time values are considered and the deviations have an initial value of zero, a unilateral Laplace transform of (2.1) can be obtained as shown below,

$$\Delta P_m(s) - \Delta P_l(s) = 2Hs\Delta f(s) + D\Delta f(s) \tag{2.2}$$

where $H$ is the inertia constant, $D$ is the load damping coefficient, and $\Delta f$ is the frequency deviation from nominal value. According to Laplace transform notations, $\mathscr{L}\{f(t)\} = F(s)$. However, to avoid inconsistencies with power system notations, the Laplace transformed signals are also considered with the same notations as the time domain signal.

From (2.1), the transfer function between the generation-load imbalance and frequency is obtained as

$$\frac{\Delta f(s)}{\Delta P_m(s) - \Delta P_l(s)} = \frac{1}{2H + D} \tag{2.3}$$

There are usually multiple generating units within a balancing area, and the load generation characteristic can be lumped together to represent a single block for the area.

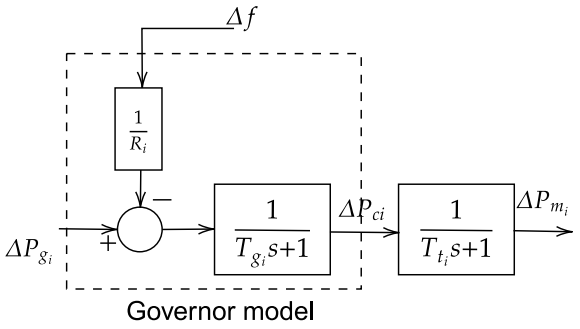



**Fig. 2.3** Steam reheat governor-turbine model

The turbine and generator dynamics are different for each machine type like a steam turbine, hydro-turbine, etc. Other generating units like batteries and distributed generators are also encapsulated using different types of turbine generator transfer functions. An example block diagram for a steam reheat governor-turbine set is shown in Fig. 2.3.

The above blocks constitute the primary frequency control.

For secondary frequency control, the change in tie-line flows also comes into the picture. A combination of the frequency deviation and the tie-line deviation called the Area Control Error is used in the MA-LFC.

The power flow between two areas is obtained as

$$P_{tie,\,ij} = \frac{V_i V_j}{X_{ij}} sin(\delta_i - \delta_j) \tag{2.4}$$

Let $T_{ij} = \frac{|V_i||V_j|}{X_{ij}}\, cos(\delta_i^0 - \delta_j^0)$ be the synchronizing torque coefficient, then (2.4) can be linearized about an equilibrium point $(\delta_i^0,\ \delta_j^0)$ as

$$P_{tie,\,ij} = T_{ij}(\delta_i - \delta_j) \tag{2.5}$$

Power angle and frequency are related as $\delta = 2\pi \int \Delta f$. Thus, (2.5) may be written in terms of frequency deviation as

$$\begin{aligned} P_{tie,ij}(t) &= 2\pi T_{ij} \left( \int \Delta f_i(t) - \int \Delta f_j(t) \right) \\ P_{tie,ij}(s) &= \frac{2\pi}{s} T_{ij}(\Delta f_i(s) - \Delta f_j(s)) \end{aligned} \tag{2.6}$$

Net tie-line flow considering all $N$ areas would then be

$$P_{tie,i} = \frac{2\pi}{s}\left[\sum_{\substack{j=1\\ j\neq i}}^{N} T_{ij}\Delta f_i(s) - \sum_{\substack{j=1\\ j\neq i}}^{N} T_{ij}\Delta f_j(s)\right] \tag{2.7}$$

Using the frequency deviation and the tie-line power deviation, the Area Control Error ($ACE$) may be calculated as

$$ACE_i = \Delta P_{tie,i}(s) + \beta_i \Delta f_i(s) \tag{2.8}$$

where $\beta_i = \frac{1}{R_i} + D_i$ is called the bias factor.

The ACE is finally sent to a PI controller that sends the necessary generation changes to the generating units based on the value received. The PI controller may be further modified to include advanced and robust control algorithms like H-inf control, sliding mode control, etc.

## 2.5 State-Space Representations

The overall state-space representation for the MA-LFC can be derived using the above equations.

The linearized state-space model of the LFC system is given by (2.9),

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} + F\mathbf{d} \\ \mathbf{y} &= C\mathbf{x} \end{aligned} \tag{2.9}$$

where

$$\begin{aligned} \mathbf{x} \in \mathbb{R}^{5n_a} &= \text{State vector} = [x_1 ... x_i ... x_n] \\ \mathbf{u} \in \mathbb{R}^{n_a} &= \text{Controlled Input} = [\Delta P_{c_1} ... \ \Delta P_{c_n}] \\ \mathbf{d} \in \mathbb{R}^{n_a} &= \text{Disturbance Input} = [\Delta P_{l_1} ... \Delta P_{l_n}] \\ \mathbf{y} \in \mathbb{R}^{n_a} &= \text{Output Vector} = [ACE_1 ... ACE_n] \\ x_i &= [\Delta f_i \ \Delta P_{tie,i} \ \Delta P_{m_i} \ \Delta P_{g_i}] \end{aligned}$$

The frequency and tie-line power are transmitted to the central or area control center, which executes the LFC procedure and returns the change in generation scheduling to the generating units based on the Area Control Error (ACE) calculation.

The linearized system matrices for each area are

$$A_i = \begin{bmatrix} \frac{-D_i}{2H_i} & \frac{-1}{2H_i} & \frac{1}{2H_i} & 0 \\ 2\pi \sum T_{ij} & 0 & 0 & 0 \\ 0 & 0 & \frac{-1}{T_{ti}} & \frac{1}{T_{ti}} \\ \frac{-1}{T_{gi}R_i} & 0 & 0 & \frac{1}{T_{gi}} \end{bmatrix} B_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{pf_i}{T_{gi}} \end{bmatrix}$$

$$F_i = \begin{bmatrix} \frac{-1}{2H_i} \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad C_i = \begin{bmatrix} \beta_i & 1 & 0 & 0 \end{bmatrix}$$

### 2.5.1 Nonlinearities Modeling in MA-LFC

A linearized model of the MA-LFC system is given in (2.9). However, a practical LFC system consists of several nonlinearities. An important physical constraint is the power generation rate change due to the limitation of thermal and mechanical movements. This rate is termed a generation rate constraint (GRC). Effects of GRC on the performance of secondary systems are reported in Nanda et al. (1983).

Speed governor dead band is another important issue affecting LFC performance. The speed governor may not immediately react by changing the input signal until the input reaches a specified value. The effect of the governor dead band is to increase the apparent steady-state speed regulation (Concordia et al. 1957).

Communication delays introduce another significant challenge in the LFC synthesis due to the restructuring, expanding of functionality, and increased complexity of power systems. In the control systems, time delays can degrade the system's performance and even causes system instability (Bhowmik et al. 2004).

Thus these nonlinearities must be considered during the design of attacks, detection, and mitigation mechanisms.

The detailed nonlinear model of the LFC governor-turbine set is as shown in Fig. 2.4 (Bevrani 2014).

The nonlinear model with time delays can be written as below

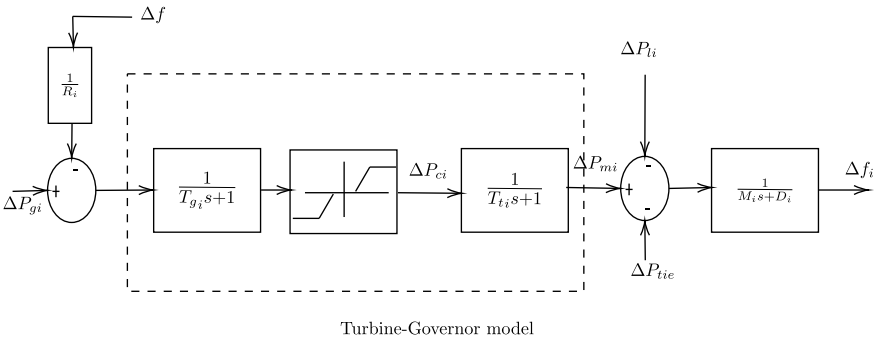



**Fig. 2.4** Detailed model of generator

$$\begin{aligned} x(k+1) &= F(x(k), x(t-\delta), d(k), d(k-\gamma) + M \begin{bmatrix} w(k) \\ v(k) \end{bmatrix} \\ y(k) &= G(x(k), d(k)) + L \begin{bmatrix} w(k) \\ v(k) \end{bmatrix} \end{aligned} \tag{2.10}$$

Here, $F$ and $G$ are nonlinear functions representing the system dynamics, and $w$ and $v$ are process and measurement zero-mean Gaussian white noise, respectively. $\delta$ and $\gamma$ are the time delays in state and input, respectively.

## 2.6 Load Frequency Control Cyber-Physical System

The mathematical modeling of the load frequency control has been discussed in the previous section and this section explains in detail the implementation of the LFC system including all the communication media and the protocols that are in place.

In a typical grid system with LFC, dedicated communication lines are used to send the tie-line and frequency values from the control area to the LFC control center using the Inter Control Center Protocol (ICCP). The SCADA/EMS system at the control center transmits this real-time data to the LFC control unit. The set-point values are then communicated by the control center to the communication equipment/substation which is nearest to the generating plants using IEC 104 or DNP3 protocol. This communication is two way. Finally, the set-point is sent to individual generating units using Open Platform Communications (OPC) or MODBUS protocol.

The detailed cyber-physical model of the LFC system with communication media and protocols used is as shown in Fig. 2.5. The model has been adopted from the proposed AGC pilot project by the Power System Operation Corporation (POSOCO). The frequency and tie-line data of each area is communicated to the Regional Load Dispatch Center (RLDC) and then to the National Load Dispatch Center (NLDC) where the LFC is located.

The LFC system communications of other large systems such as PJM, ERCOT, MISO, and Japan are also similar with small changes. Though firewall is in place at the control center, it is still possible for an attacker to spoof the measurements by attacking either the sensors or the communication channel. The protocols used, ICCP, DNP3, and MODBUS, are known to be vulnerable to cyber attacks (East et al. 2009). The data is routed at various points which could also be suitable attack points.

## 2.7 Summary

This chapter discusses in detail the various control systems present in a power system and the importance of frequency control. Frequency Control is important due to the following reasons:

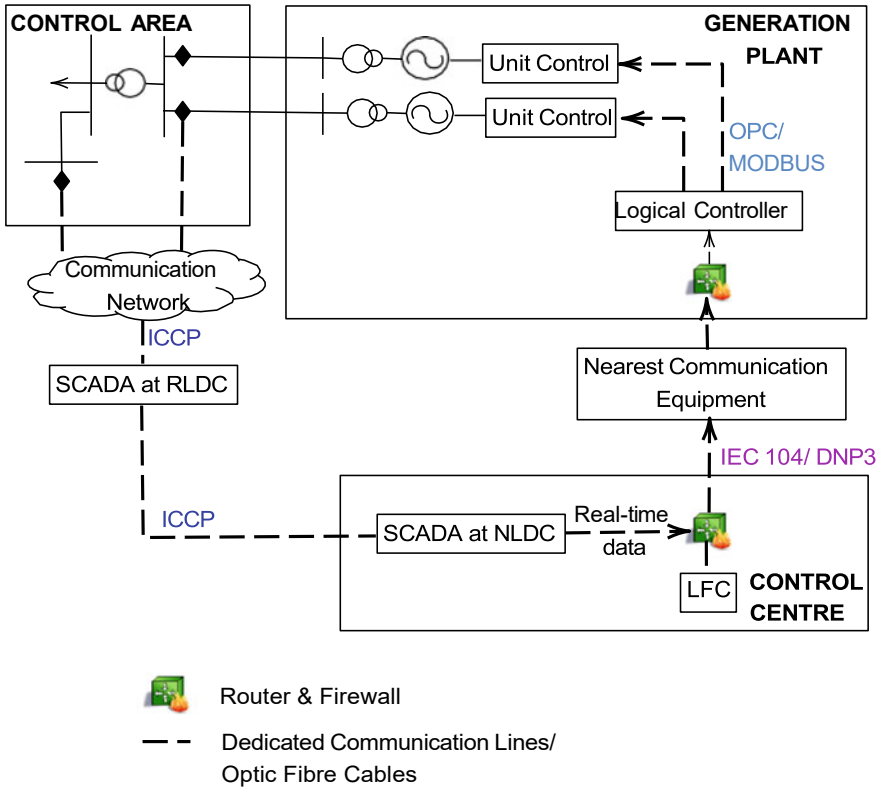



**Fig. 2.5** Load frequency control cyber-physical system

1. Frequency is a direct measure of generation-load imbalance in an entire area. This balance ensures the stability of the complete grid system.
2. Severe frequency deviations can trigger UFLS and generator trips.
3. Thus, any attack on one part of the frequency control system can propagate through the entire grid and cause stability issues, ultimately leading to blackouts.

The MA-LFC system model is derived based on the machine and system equations, and a state-space model is developed to represent the MA-LFC system. Since the system is a linearized one, we also introduce the nonlinearities. Analysis of the nonlinearities and including them in the model is essential in our study since any attack detection algorithm may fail to operate or may give false alarms if it is subjected to nonlinearities and noises.

The model derived in this chapter will further be used as the representative system to analyze the various proposed algorithms.

# References

Annaswamy AM, Amin M (2013) IEEE vision for smart grid controls: 2030 and beyond. In: IEEE vision for smart grid controls: 2030 and beyond, pp 1–168

Bevrani H (2014) Robust power system frequency control. Springer

Bhowmik S, Tomsovic K, Bose A (2004) Communication models for third party load frequency control. IEEE Trans Power Syst 19(1):543–548

Concordia C, Kirchmayer LK, Szymanski EA (1957) Efect of speed-governor dead band on tie-line power and frequency control performance. Trans Am Inst Electr Engineers Part III: Power Appar Syst 76:429–434

East S, Butts J, Papa M, Shenoi S (2009) A taxonomy of attacks on the dnp3 protocol. In: Palmer C, Shenoi S (eds) Critical infrastructure protection III. Springer, Berlin, pp 67–81

Nanda J, Kothari ML, PS Satsang (1983) Automatic generation control of an interconnected hydrothermal system in continuous and discrete modes considering generation rate constraints. IEEE Proc (Control Theory Appl 130(10):17–27

Wood AJ, Wollenberg BF, Sheble GB (2013) Power generation automation and control. Wiley

Yohanandhan RV, Elavarasan RM, Manoharan P, Mihet-Popa L (2020) Cyber-physical power system (CPPS): a review on modeling, simulation, and analysis with cyber security applications. IEEE Access 8:151019–151064

# Part II
# Attacks in Smart Grid Control Vulnerability Assessment

# Chapter 3
# Attack Modeling for Smart Grid Control


**Abstract** This chapter provides a comprehensive examination of Smart Grid Attack Modeling, offering insights into the intricacies of Multi-Area LFC (MA-LFC), various attack scenarios, and their potential impacts. It serves as a critical resource for researchers, practitioners, and policymakers seeking to enhance the resilience and security of smart grid systems. The chapter commences with an introduction, setting the stage for an in-depth exploration of the subject matter. It then transitions into the details of MA-LFC, highlighting the critical role it plays in grid stability. A significant portion of the chapter is devoted to Attack Modeling, shedding light on potential threats and adversarial strategies compromising the integrity of smart grid. The discussion extends to Stealth/Undetectable Attacks, emphasizing the importance of these subtle threats in real-world scenarios. The chapter also introduces the concept of Multiple-Attack Models, encompassing a variety of attack vectors such as Scaling, Ramp, False Data Injection (FDI), and Zero-Day Attacks. These models serve as a foundation for understanding and evaluating the challenges that the smart grid may face. Lastly, the chapter explores Attack Impact Analysis through a practical case study.

**Keywords** Cyber attacks · False Data Injection Attacks (FDIA) · Data Integrity Attacks (DIA) · Stealth attacks · Smart grid security

## 3.1 Introduction

A cyber-physical system security model used for analyzing the system security is a combination of the system model and the attacks that could be introduced into the system. Detailed modeling of the control system can enable better system dynamics analysis under various attacks. The designed methods are likely to fail in real grid conditions if an incorrect model is chosen for designing and testing an algorithm in a simulation environment. As discussed in Chap. 2, we focus on Load Frequency Control as a case study for analyzing various models, detection, and mitigation.

LFC is an important control system since it is responsible for maintaining grid frequency and area power exchanges (Wood et al. 2013; Kundur 1994). It is a distributed control system, and any attack on one part of the LFC can propagate throughout and



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_3

cause widespread damage. The linearised system model described by Wood et al. (2013) gives a good representation of the LFC. However, to improve the model (Nanda 1983; Concordia et al. 1957; Bhowmik et al. 2004) have introduced the generator rate constraint, governor dead band, and communication delay nonlinearities, respectively. These improvements help in analyzing better the effect of detection and response of the system because certain detection strategies may fail when nonlinearities are taken into consideration.

Attack modeling is also significant in studying system security. Data Integrity Attacks (DIA) can be defined as those which directly modify or append the sensor, and actuator measurements (Sridhar and Manimaran 2010; Wu et al. 2018a). The most famous cyber incident on the grid system, the Ukraine Cyber Attack, was primarily a False Data Injection Attack (Che et al. 2019). The attack surface usually includes Measuring units, communication networks, and control devices (Liang et al. 2017). While designing the attacks, one must consider the amount of system information available to an attacker. Chen et al. (2018a) have considered the realistic assumption of reduced network information. The attack vector has been modeled using an optimization algorithm to maximize the state deviation and minimize the attack cost. Mohajerin Esfahani et al. (2010) have developed methods for robust destabilization of a two-area power system using reachability-based data injection attack. In FDI attacks, attackers can inject erroneous data into meter measurements while maintaining the residual measurement. In generalized FDI attacks (Zhao et al. 2018), the attacker uses the standard measurement error tolerance of state estimate techniques and remains undetected and stealthy.

This chapter provides a detailed description of the LFC system hierarchy and a unified LFC and attack model. Further, we define monitors for detecting attacks using residuals and derive a condition for stealth attacks with reduced system knowledge.

## 3.2 Smart Grid Attack Modeling Overview

The attacks on the grid control systems can be broadly classified into Denial-of-Service Attacks (DoS) and Data Integrity Attacks. These attacks are as explained as follows:

1. Denial-of-Service Attack (DoS): The attacker floods the communication channels with data to prevent actual data from being unavailable to the LFC system.
    a. Distributed DoS (DDoS): Coordinated DoS over multiple distributed equipment/locations.
2. Data Integrity Attacks: Here, the attackers directly modify or append the sensor and actuator measurements.
    a. Resonance Attack: Data Modification is performed in accordance with the changes in measurement.
    b. Stealth Attack: Attacks that can surpass bad data detectors or simple detection techniques.

The above attacks can be injected into the system at various layers of the power system using different methods. Table 3.1 gives an overview of the different types of attacks and how they are implemented in each layer.

As seen from the table, the various attacks can happen at different layers of the system by using different techniques. It is therefore important to devise methods for not only detecting them at various levels but also analyzing them.

The DoS attacks cause interference to the flow of data. These attacks weakly influence the dynamics if they are launched after the system dynamics converge and the effects will be worse if the attackers launch the DoS before system convergence.

**Table 3.1** Power system attack techniques

| Sl. no. | Attack type | Process/physical layer | Application/control layer | Communication layer | References |
|---|---|---|---|---|---|
| 1 | DoS | Node Destruction/ Interference | API Attacks, Volumetric Attacks | Flooding, Wormhole | Liu et al. (2013, 2019), Cheng et al. (2020), Wu et al. (2019), Shen et al. (2017), Peng et al. (2017) |
| 2 | DDoS Attack | Sensor Overload, Path based | API, Volumetric Attacks | Resource Exhaustion | Wang et al. (2019), Girma et al. (2015) |
| 3 | FDIA | Manipulate IED SCD files through password cracking, GPS spoofing | Controller access, supply-chain compromise, attack on controller HMI | Data Network Manipulation, Attack on switches | Sridhar and Manimaran (2010), Liu et al. (2009), Tan et al. (2017), Bi et al. (2019c), Chen et al. (2018a), Bi et al. (2019a), Mohajerin Esfahani et al. (2010), Bi et al. (2019b), Chen et al. (2018b), Sarangan et al. (2018) |
| 4 | Resonance Attack | Load Manipulation, Direct switching | – | Data network manipulation | Wu et al. (2018b) |
| 5 | Stealth Attacks | Frequency and tie-line channels, direct sensor attacks | Complete system knowledge | Data network infiltration, sensor physical access/ SCD file corruption | Sridhar and Manimaran (2010) |

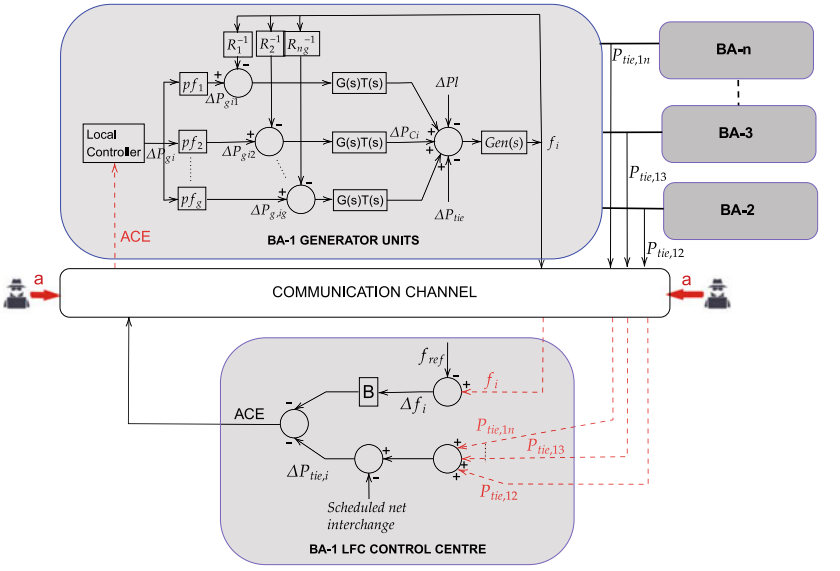
The DDoS can cause an even faster impact on the system due to the distributed and coordinated behavior. The FDIA can be tactically modeled by an attacker to bypass the bad data detection algorithms and stay undetected until it causes some impact on the system dynamics. Thus, these types of attacks, also called as stealth attacks, can cause severe impacts on the grid operation and stability. Stealth attack modeling however assumes the complete knowledge of the system to the attacker which is not a practical assumption. Thus, if an attacker has only publicly available knowledge of the system and yet launches a successful full stealth attack then the system is under a severe threat.

Analysis of the literature shows that the attacks can be modeled to remain undetected and also cause widespread impacts on the system in a very short duration of time. Thus it is important to devise fast detection and mitigation to safeguard the system from these attacks.

## 3.3 Multi-area Load Frequency Control (MA-LFC)

Figure 3.1 depicts the LFC control of one Balancing Area and the attack surface.

The Multi-Area Load Frequency Control (MA-LFC) is responsible for maintaining the frequency and tie-line flows at scheduled values. The frequency and tie-line sensors measure the physical properties and communicate these to the LFC




B: Bias Factor R: Droop factor a:attack vector $f_a, f_{ref}$: Actual & Reference frequency $P_{tie,i}$: MW Measurement of Tie-line $i$
ACE : Area Control Error $\Delta P_{g_i}$: Generation change of each generator $j\,\Delta P_{c_i}$:Net generation change in area $i$
$pf_j$: Participation factor of generator $j$G(s)T(s): Governor-Turbine function $\Delta P_l$: Change in area load $Gen(s)$: Generator function

**Fig. 3.1** Load frequency control of one area of a multi-area system

controller, which, based on the control algorithm, commands the generation units that directly adjust the generation output based on the received Area Control Error (ACE). This cycle takes place every 2–5 s.

### 3.3.1 MA-LFC Modeling

Figure 3.1 shows the detailed system model of one balancing area of a multi-area power system, including the attack surface (Wood et al. 2013). The frequency ($f_i$) and tie-line flows between areas $i$ and $j$ ($P_{tie,ij}$) are sent to the LFC control center of each area. At the control center, these values are compared with the reference values $f_{\text{ref}}$ and 'Scheduled net interchanges' to get the net change in frequency ($\Delta f_i$) and tie-line flow ($\Delta P_{tie,i}$). Based on these values, an ACE is calculated based on which the generators increase or decrease their generation values. It is assumed that the attack happens at the communication layer. Thus the sensor values to the control center and the ACE value to the generators are represented as red-dashed arrows in Fig. 3.1.

The state-space model of $i$th balancing area of the LFC system is given by (3.1),

$$\begin{aligned} x_i(k+1) &= A_i x_i(k) + B_i d_i(k) + W \begin{bmatrix} \omega(k) \\ \nu(k) \end{bmatrix} \\ y_i(k) &= C_i x_i(k) + V \begin{bmatrix} \omega(k) \\ \nu(k) \end{bmatrix} \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} &x_i \in \mathbb{R}^5 = \text{State vector} = [\Delta f_i \ \Delta P_{m_i} \ \Delta P_{g_i} \ \Delta P_{tie,i} \ \Delta P_{c_i}]' \\ &d_i \in \mathbb{R} = \text{Input Vector} = \Delta P_{l_i} \\ &y_i \in \mathbb{R} = \text{Output Vector} = ACE_i \\ &\forall i \in 1, 2, ..., \text{number of areas } (n_a) \end{aligned}$$

The system input (**d**) is the net change in the balancing area load ($\Delta P_{l_i}$) and the system output (**y**) is the ACE of each balancing area. The state vector (**x**) is a five-element vector that consists of change in frequency ($\Delta f_i$), governor output ($\Delta P_{m_i}$), turbine output ($\Delta P_{g_i}$)), tie-line power ($\Delta P_{tie,i}$), and generation ($\Delta P_{c_i}$). Here, $\omega$ and $\nu$ are process and measurement zero-mean Gaussian white noise, respectively. Matrices $A_i \in \mathbb{R}^{5\times5}$, $B_i \in \mathbb{R}^{5\times1}$, and $C_i \in \mathbb{R}^{1\times5}$ are the state-space matrices. $W$ and $V$ are matrices used to mathematically denote the process and measurement noise addition into **x** and **y**, respectively. A white Gaussian noise term is added to the state and output variables used during the simulation. The noise is assumed to have a standard deviation equal to the accuracy of the corresponding measurement and a zero mean.

The frequency and power measurements are obtained based on voltage and current measurements obtained from instrument transformers in the field. The measurement errors of instrument transformers are limited by their accuracy class. The accuracy

is in the range of $\pm 0.1\%$ to $\pm 0.3\%$ for the measurement of voltage and current magnitudes using modern Current Transformers (CTs) and Potential Transformers (PTs) as specified in IEC 60044 and IEEE C57.13. Thus, noises in the generated measurements were assumed to have a standard deviation of 0.1% (or $10^{-3}$ p.u.) for magnitude and $10^{-4}$ rad for phase/frequency (Singh and Pal 2019).

The generator nonlinearities are included in the generator-turbine-governor model by adding a limiter and saturation to the governor-turbine system model. The detailed governor-turbine model and physical meanings of the matrices in (3.1) are given in Appendix A. (3.1) is a simplified linearized model for the nonlinear LFC system that is used to generate the data for training and to test the algorithm. However, if available, historical frequency and tie-line power data can be effectively used in the proposed algorithm to get an exact representation of the system dynamics.

## 3.4 Attack Modeling for MA-LFC

Considering a worst-case attack scenario to study the proposed algorithm's effectiveness, we assume that the attacker can effectively obtain sufficient system knowledge to launch a stealth attack. This system knowledge can be obtained using an Eavesdropping Attack (EDA) to obtain the sensor data over time and use system identification methods to obtain system parameters. Additive inputs can model the FDIA attacks on sensor or actuator measurements to alter the dynamic equation. An attacker can launch such attacks on the physical, cyber, or both.

The system model with attack input $u_a(k)$ is given by (3.2)

$$\begin{aligned} x_{a,i}(k+1) =& A_i x_a(k) + B_i d_i(k) + B_a u_a(k) + W \begin{bmatrix} \omega(k) \\ \nu(k) \end{bmatrix} \\ y_{a,i}(k) =& C_i x_a(k) + D_a u_a(k) + V \begin{bmatrix} \omega(k) \\ \nu(k) \end{bmatrix} \end{aligned} \tag{3.2}$$

where $B_{a,i} \in \mathbb{R}^{5\times 3}$ and $D_{a,i} \in \mathbb{R}^{1\times 3}$ are the matrices which characterize the attack input and $u_a \in \mathbb{R}^3$ is a vector of attack inputs. $u_a \in \mathbb{R}^3$ since attack input is the frequency, tie-line, and ACE and is given by (3.3)

$$u_a = [f_{attack}\ P_{tie_{attack}}\ ACE_{attack}]' \tag{3.3}$$

The values in each row of $u_a$ will evolve according to the type of attack considered.

For a step attack, the structure of $u_a$ is

$$u_a = \begin{bmatrix} 0\ 0\ 0\ \ldots\ 0 & d_f & d_f & d_f & d_f \\ 0\ 0\ 0\ \ldots\ 0 & d_{Pt} & d_{Pt} & d_{Pt} & d_{Pt} \\ 0\ 0\ 0\ \ldots\ 0 & d_{ACE} & d_{ACE} & d_{ACE} & d_{ACE} \end{bmatrix} \tag{3.4}$$

where $d_f$, $d_{Pt}$, and $d_{ACE}$ are the step attack values for frequency, tie-line flow, and ACE, respectively. The instant of attack and value of attack can be varied according to the attacker's intent.

For random and stealth attacks, the values of $d_f$, $d_{Pt}$, and $d_{ACE}$ in the above matrix at each instant will be different.

The matrices $B_a$, $D_a$ are given by

$$B_a = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sum P_{ij} & 0 & 0 \\ 0 & 0 & -K_i \end{bmatrix} \qquad D_a = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

1. For a fixed attack, $B_a u_a$ and $D_a u_a$ remain constant.
2. For variable attacks, $B_a u_a$ and $D_a u_a$ change for every time instant $k$. These attacks can vary randomly, or an attacker with system information can generate variable attacks synchronous to the system dynamics, as explained in the next section.

## 3.5 Stealth/Undetectable Attacks

Stealth Attacks are a type of FDIA attack with the additional capability of not being detected by conventional bad data detection strategies. Such attacks are designed using available system information and injected into the system such that the system equations are satisfied. This subsection defines monitors used for bad data detection and derives a condition for the attacks to surpass these monitors.

**Definition 3.1** A **monitor** is defined as an algorithm $\phi : \Lambda \rightarrow \Psi$, where $\Lambda$ is the algorithm input that includes measurements and system data, and the output $\Psi =$ {True, False} is True if the monitor detects an attack.

Based on the criteria used, monitors can be further classified as static or dynamic. In this work, we consider only static monitors which are defined as follows.

**Definition 3.2** A **static monitor** $\phi : \Lambda \rightarrow \Psi$ is defined as

$$\begin{aligned} \Psi &= True,\ ||\mathbf{y}(k) - \hat{\mathbf{y}}(k)|| \geq \tau \\ &= False,\ otherwise \end{aligned} \tag{3.5}$$

with given input $\Lambda = \{C, \mathbf{y}(k)\, \forall k \in \mathbb{N}\}$ and predetermined threshold $\tau$.

Here, the estimate $\hat{\mathbf{y}}(k)$ depends on the measurement equation alone, i.e, $\hat{\mathbf{y}}(k) = C\, x(k)$. In the absence of an attack, the output should ideally be $\mathbf{y}(k) = C\, \mathbf{x}(k)$. An attack is detected whenever the difference between the actual and estimated

measurement using the equation is ideally any value greater than zero. We use $\tau$ instead of 0 to account for the noise.

Stealth attacks are defined as attacks that monitors cannot detect. In this section, we prove the existence of stealth attacks in the presence of a monitor. If an attack is such that the residue $r = y(k) - \hat{y}(k)$ calculated by the monitor remains the same as that without an attack, then the attack will be stealth or undetected.

**Theorem 3.1** *For the system defined in (3.2) with static monitor $\phi$ as given in (3.5), an attack $u_a$ will be undetected if $Du_a(k) \in \{0,\ Im(C)\}$.*

The proof of the above theorem is given in Appendix D.

Based on Theorem 3.1, any stealth attack $D_a u_a$ can be represented as $C\delta$ where $\delta$ is any arbitrary vector of size same as the state vector. Thus, the stealth attack is

$$D_a u_a = C\delta \tag{3.6}$$

## 3.6 Multiple-Attack Model

In the previous section, we have looked at the False Data Injection Attacks and a method for building such stealth attacks. In order to analyze a detection algorithm's performance, it is necessary to consider various attacks, their combinations, and variations with time. This work considers a random and varying combination of Scaling, Pulse, and Ramp attacks in the sensor and actuator measurements. We use Bernoulli variables to characterize the attacks. The mathematical model of the attack templates is as follows.

### *3.6.1 Scaling Attack*

In scaling attacks, the output value is scaled by a factor $\lambda_s$. The Bernoulli variable $\beta_s$ is used to characterize the attack,

$$\begin{aligned} y_{a_1}(t) &= \beta_s(t)(1+\lambda_s)y(t) + (1-\beta_s(t))y(t) \\ y_{a_1}(t) &= \beta_s(t)\lambda_s y(t) + y(t) \end{aligned} \tag{3.7}$$

### *3.6.2 Ramp Attack*

In ramp attacks, the output is increased proportional to the time progression,

$$\begin{aligned} y_{a_2}(t) &= \beta_r(t)(y_{a_1}(t) + \lambda_r \cdot t) + (1-\beta_r(t))y_{a_1}(t) \\ y_{a_2}(t) &= \beta_r(t)\lambda_r \cdot t + y_{a_1}(t) \end{aligned} \tag{3.8}$$

### 3.6.3 False Data Injection Attack (FDIA)

In a random FDIA attack, a random value is added to the output measurements,

$$\begin{aligned} y_{a_3}(t) &= \beta_p(t)(y_{a_2}(t) + \lambda_p) + (1 - \beta_p(t))y_{a_2}(t) \\ y_{a_3}(t) &= \beta_r(t)\lambda_p + y_{a_2}(t) \end{aligned} \tag{3.9}$$

Thus the overall attack model is given by

$$y_{a_3}(t) = \beta_r(t)\lambda_p + \beta_r(t)\lambda_r \cdot t + \beta_s(t)\lambda_s y(t) + y(t) \tag{3.10}$$

The model described by (3.10) gives the multiple-attack model. Based on the value of the Bernoulli variable, the attack whose $\beta$ value is 1 will be injected into the system. This gives a good representation to analyze multiple attacks on a system. There also exist some attacks that are not mathematically defined. These attacks are called zero-day attacks.

### 3.6.4 Zero-Day Attacks

Zero-day attacks are those attacks that a signature-based security software solution cannot detect at the time of the malware's release. Thus, it can evade conventional security solutions to cause the intended harm. According to multiple databases and researchers, 74% of the threats discovered in the first quarter of 2021 were zero-day attacks. It is nearly double the total for 2020 and is higher than in any other year on record.

Identifying such attacks that are not in the security or detection software database is crucial so that they can be further analyzed and mitigation steps are suitably modified.

## 3.7 Attack Impact Analysis for IEEE 39-Bus New England Test System LFC

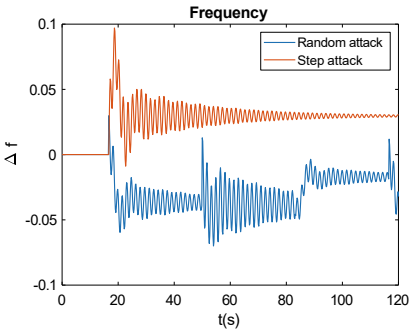
To analyze the effect of the different attacks discussed in the above sections, the attacks were simulated at the MA-LFC portion of a 39-bus 3 area New England test system. The attacks are injected into the sensor (frequency and tie-line) and actuator (ACE) measurements of balancing area-1 to analyze the impact of the attacks. The simulations are carried out on MATLAB R2018a on a core i5 system. Appendix A gives the system diagram and parameters.

### 3.7.1 Example 3.1: Single Attack Dynamics

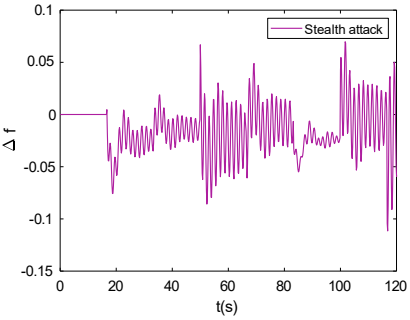
To analyze the effect of the attacks alone, no load changes were introduced into the system. Figures 3.2, 3.3, and 3.4 show the effect of different types of attacks on the frequency, tie-line, and ACE under no-load variation. Thus, the dynamics correspond only to the attack injections.

In the left figures, the red plots correspond to dynamics under a fixed step attack, and the blue plots correspond to variable attacks. The right side plots are obtained using stealth attacks. All the attack values are in the noise range.

It can be seen that in the case of fixed attacks, the impact is a change in the steady-state. However, variable attacks inject small attack values at various instants of time and are thus difficult to detect. These attacks can build up over time, as seen from the plots. Thus, any detection strategy should be fast and detect even small attack values before building up the dynamic deviations.

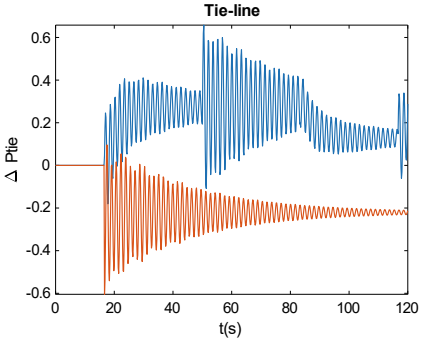



(a) $\Delta f$ : Random & Step attacks: Change in steady state but within limits

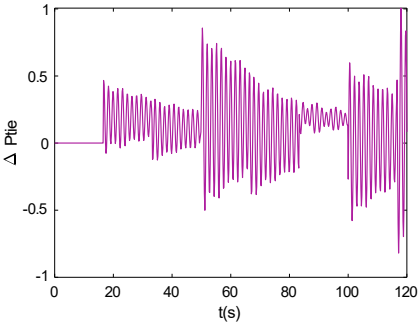



(b) $\Delta f$ : Stealth attacks: Small attacks build up the dynamics

**Fig. 3.2** Change in frequency measurement under random, step, and stealth attacks

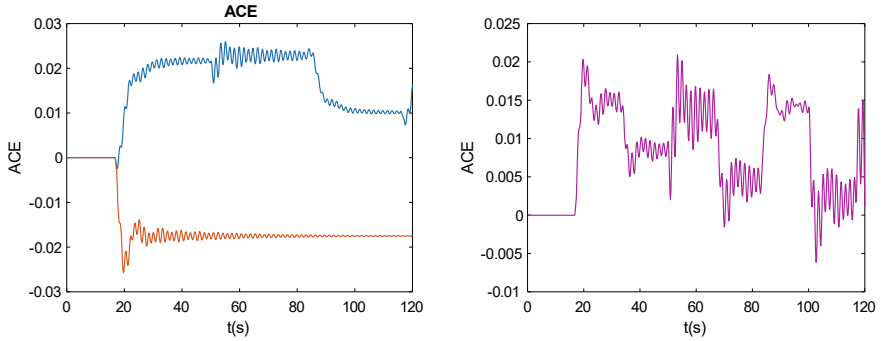



(a) $\Delta P_{tie}$: Random & Step attacks




(b) $\Delta P_{tie}$: Stealth attacks

**Fig. 3.3** Change in tie-line measurement under random, step, and stealth attacks




(a) ACE: Stealth attack (b) ACE: Random & Step attacks

**Fig. 3.4** Change in ACE under random, step, and stealth attacks

### *3.7.2 Example 3.2: Multiple-Attack Dynamics*

We next look at the system dynamics in the presence of load change and multiple-attack models. The simulation is carried out at a sampling frequency of 1 Hz. The load variations obtained from the New England ISO website simulate the normal system dynamics. The multiple-attack model is then used to inject different attacks into the system, and the dynamic variations during the attack and regular data are as shown in Fig. 3.5.

As seen from Fig. 3.5, the dynamic variations in the frequency and tie-line signals are not very prominent. Thus, it is not easy for an operator to detect them just by

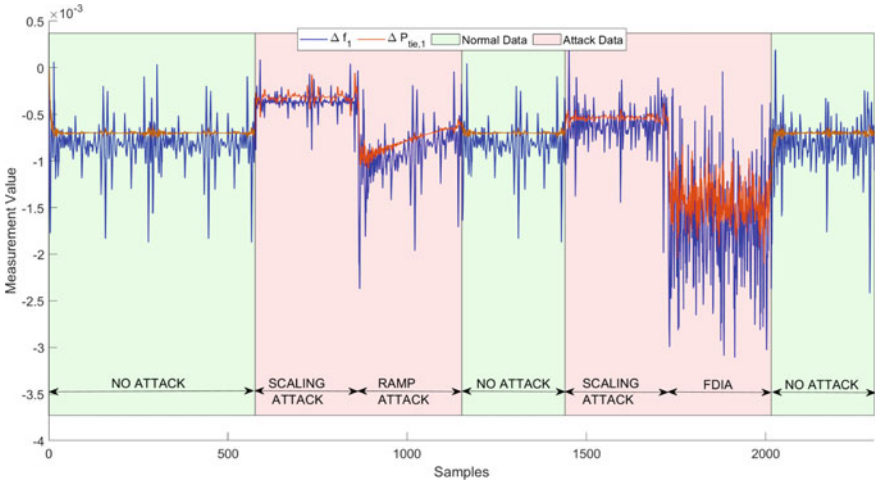



**Fig. 3.5** Dynamics of multiple-attack on LFC

observation. Monitors also cannot detect these attacks; thus, such attacks can affect the system performance and lead to instability.

Hence, the multiple-attack model is an effective tool for studying the performance of detection algorithms.

## 3.8 Future Scope

A major scope of research in attack detection includes data and models on renewable generation. Renewable penetration into the grid can cause varied dynamics which could lead to failure of the detection and can lead to a high value of False Alarm Rates. Thus, more research is required in this field. Additionally, detection algorithms that combine detection and mitigation can work in better harmony and adapt to system changes leading to a complete cyber-security solution.

### *3.8.1 Research Gap*

- Lack of comprehensive understanding of all potential attack vectors and their consequence.
- Power control systems are large cyber-physical systems and existing methods become computationally challenging.
- Limited research on the cascading effects of multiple simultaneous attacks.
- Emerging technologies and security risks.

### *3.8.2 Research Directions*

- Utilize threat models such as MITRE ATT&CK ICS and National Vulnerability Database (NVD) in combination with power system simulations.
- Natural Language Processing in combination with expert systems can be used to extract relevant threats and apply them into control strategies.
- Inductive learning-based neural network models can be used to transfer knowledge from small power systems to large systems for computational efficiency.
- Impact study of coordinated or composite cyber attacks and the creation of countermeasures for these attacks, taking communication-based aspects into account.

## 3.9 Summary

In this chapter, a detailed system and attack model is discussed. The multi-area LFC model used in further chapters is explained in detail. This system model is further modified to include the attack model, thus forming a combined system and attack model.

Static monitors are algorithms that can identify anomalies (bad data detection). The attack is called stealthy if an attacker can model the attacks based on available data and go unnoticed through the monitors. A successful stealth attack condition is derived for the attacks to surpass the detection. Thus, we successfully model a stealth attack using available partial system information.

The effect of each type of FDIA is analyzed by injecting these attacks into the 39-bus 3 area LFC model. It can be found that though the injected attack at each time step is minimal, the dynamics develop over time, and the system loses synchronicity. Finally, a multiple and time-varying attack model is proposed using Bernoulli variables. Such an attack model can be used to test the detection strategy's performance.

## References

Bhowmik S, Tomsovic K, Bose A (2004) Communication models for third party load frequency control. IEEE Trans Power Syst 19(1):543–548

Bi W, Chen C, Zhang K (2019) Optimal strategy of attack-defense interaction over load frequency control considering incomplete information. IEEE Access 7:75342–75349

Bi W, Zhang K, Li Y, Yuan K, Wang Y (2019) Detection scheme against cyber-physical attacks on load frequency control based on dynamic characteristics analysis. IEEE Syst J 13(3):2859–2868

Bi W, Zhang K, Yuan K, Wang Y, Chen C, Wang K (2019) Observer-based attack detection and mitigation for load frequency control system. 2019 IEEE Power Energy Society General Meeting (PESGM). Atlanta, GA, USA, pp 1–5

Che L, Liu X, Li Z, Wen Y (2019) False data injection attacks induced sequential outages in power systems. IEEE Trans Power Syst 34(2):1513–1523

Chen C, Cui M, Wang X, Zhang K, Yin S (2018) An investigation of coordinated attack on load frequency control. IEEE Access 6:30414–30423

Chen C, Zhang K, Yuan K, Zhu L, Qian M (2018) Novel detection scheme design considering cyber attacks on load frequency control. IEEE Trans Ind Inform 14(5):1932–1941

Cheng Z, Yue D, Hu S, Huang C, Dou C, Chen L (2020) Resilient load frequency control design: dos attacks against additional control loop. Int J Electr Power Energy Syst 115:105496

Concordia C, Kirchmayer LK, Szymanski EA (1957) Effect of speed-governor dead band on tie-line power and frequency control performance. Trans Am Inst Electr Engineers Part III: Power Apparatus Syst 76:429–434

Girma A, Garuba M, Li J, Liu C (2015) Analysis of DDoS attacks and an introduction of a hybrid statistical model to detect DDoS attacks on cloud computing environment. 2015 12th International Conference on Information Technology—New Generations. Las Vegas, NV, USA, pp 212–217

Kundur P (1994) Power System Stability and Control. McGraw-Hill, New York

Liang G, Zhao J, Luo F, Weller SR, Dong ZY (2017) A review of false data injection attacks against modern power systems. IEEE Trans Smart Grid 8(4):1630–1638

Liu J, Gu Y, Zha L, Liu Y, Cao J (2019) Event-triggered $h_\infty$ load frequency control for multiarea power systems under hybrid cyber attacks. IEEE Trans Syst Man Cybern: Syst 49(8):1665–1678

Liu S, Liu XP, El Saddik A (2013) Denial-of-service (dos) attacks on load frequency control in smart grids. In: 2013 IEEE PES innovative smart grid technologies conference (ISGT). DC, USA, Washington, pp 1–6

Liu Y, Ning P, Reiter MK (2009) False data injection attacks against state estimation in electric power grids. In: Proceedings of the 16th ACM conference on computer and communications security, CCS '09, pp 21–32

Mohajerin Esfahani P, Vrakopoulou M, Margellos K, Lygeros J, Andersson G (2010) A robust policy for automatic generation control cyber attack in two area power network. In: 49th IEEE conference on decision and control (CDC), pp 5973–5978

Nanda J, Kothari ML, Satsang PS (1983) Automatic generation control of an interconnected hydrothermal system in continuous and discrete modes considering generation rate constraints. IEEE Proc (Control Theory Appl) 130(10):17–27

Peng C, Li J, Fei M (2017) Resilient event-triggering $h_\infty$ load frequency control for multi-area power systems with energy-limited dos attacks. IEEE Trans Power Syst 32(5):4110–4118

Sarangan S, Singh VK, Govindarasu M (2018) Cyber attack-defense analysis for automatic generation control with renewable energy sources. In: 2018 North American power symposium (NAPS), Fargo, ND, USA, pp 1–6

Shen Y, Fei M, Du D, Zhang W, Stanković S, Rakić A (2017) Cyber security against denial of service of attacks on load frequency control of multi-area power systems. In: Li K, Xue Y, Cui S, Niu Q, Yang Z, Luk P (eds) Advanced computational methods in energy, power, electric vehicles, and their integration. Singapore, Springer, pp 439–449

Singh AK, Pal BC (2019) Decentralized dynamic estimation using PMUs. In: Dynamic estimation and control of power systems. Academic Press, pp 61–91

Sridhar S, Manimaran G (2010) Data integrity attacks and their impacts on SCADA control system. In: IEEE PES general meeting. Minneapolis, MN, USA, pp 1–6

Tan R, Nguyen HH, Foo EYS, Yau DKY, Kalbarczyk Z, Iyer RK, Gooi HB (2017) Modeling and mitigating impact of false data injection attacks on automatic generation control. IEEE Trans Inf Forensics Secur 12(7):1609–1624

Wang Q, Tai W, Tang Y, Zhu H, Zhang M, Zhou D (2019) Coordinated defense of distributed denial of service attacks against the multi-area load frequency control services. Energies 12(13)

Wood AJ, Wollenberg BF, Sheble GB (2013) Power generation automation and control. Wiley, New York

Wu G, Sun J, Chen J (2018) Optimal data injection attacks in cyber-physical systems. IEEE Trans Cybern 48(12):3302–3312

Wu Y, Wei Z, Weng J, Li X, Deng RH (2018) Resonance attacks on load frequency control of smart grids. IEEE Trans Smart Grid 9(5):4490–4502

Wu Y, Weng J, Qiu B, Wei Z, Qian F, Deng RH (2019) Random delay attack and its applications on load frequency control of power systems. In: 2019 IEEE conference on dependable and secure computing (DSC). Hangzhou, China, pp 1–8

Zhao J, Mili L, Wang M (2018) A generalized false data injection attacks against power system nonlinear state estimator and countermeasures. IEEE Trans Power Syst 33(5):4868–4877

# Chapter 4
# Vulnerability Assessment for Multi-area Load Frequency Control


**Abstract** This chapter offers a comprehensive exploration of essential aspects related to assessing the vulnerability of modern smart grids. Beginning with an introduction, the chapter proceeds to delve into key methodologies and models designed to evaluate and enhance grid resilience. The chapter introduces the concept of Data Penetration Testing, shedding light on its importance in gauging the robustness of grid systems. It then addresses the Cascading Outage Model, a critical component in understanding the potential ripple effects of vulnerabilities in load frequency control across interconnected areas. Central to the chapter is the detailed Vulnerability Assessment section, which encompasses the identification of threats and vulnerabilities, quantifying risk, and prioritizing these risks. This methodology serves as the cornerstone for strengthening grid security in the face of evolving threats. Furthermore, the chapter explores a comprehensive Detailed Risk Quantification Methodology, offering an in-depth approach to assessing and quantifying vulnerabilities within multi-area load frequency control systems. The practical application of these methodologies is exemplified through two case studies: Vulnerability Assessment for the 9-bus System and the 39-bus New England System.

**Keywords** Penetration testing · Cascading outage · Vulnerability assessment · VAPT · Grid resilience

## 4.1 Introduction

As we have seen in Chap. 3, even a minor attack can penetrate the system and cause widespread damage. Thus, to implement efficient detection and mitigation strategies, it is first necessary to identify the impacts of these attacks. Thus, Vulnerability Assessment of such attacks is important to protect these critical systems.

The system operators use Vulnerability Assessment (VA) and Penetration testing to identify the system's vulnerabilities and take necessary steps to mitigate them (Liu et al. 2017). IT and OT systems differ from each other from the cyber-security point of view. Traditional Vulnerability Assessment provides a detailed and comprehensive assessment of hardware and software assets, identifies their vulnerabilities, and



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_4

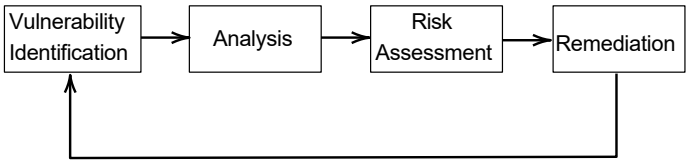



**Fig. 4.1** Vulnerability analysis steps

provides a suitable risk score. Nonetheless, it is vital to analyze the effects of these attacks on the system level characteristics and offer a risk score that can be clearly understood and managed by power system operators. Once the vulnerabilities are identified and ranked, the system operator can determine the steps for attack mitigation or system protection. The basic Vulnerability Assessment steps are given in Fig. 4.1.

This chapter presents a Vulnerability Assessment framework for a Multi-area Load Frequency Control (LFC) system from a power systems engineer's perspective. The attack can be imposed on the sensor measurements, the tie-line MegaWatt (MW) and frequency measurements, or actuator signals. The first step is Penetration testing, which assesses the vulnerabilities that might not be detectable with network or system scans using a gray-box approach. We then develop a detailed Vulnerability Assessment framework for the system.

## 4.2 Data Penetration Testing

Penetration testing simulates a hacker and is used for assessing the vulnerabilities in a system that might not be detectable with network or system scans. However, since penetration or injection of attack is not safe in a live system, we introduce a mathematical approach for the same.

There are various approaches to penetration testing wherein the test assumes different levels of data available to the attacker:

1. White box: All the system and network details are available.
2. Black box: No system or network details are available.
3. Gray box: Partial information is available.

In the presented attack problem, the assumption that an attacker can have all the system details is not practical. However, since specific power system data are easily accessible or available in the public domain, it is best to adopt the gray-box testing approach.

Figure 4.2 gives the detailed representation of the power system with MA-LFC. A 3 area system connected by tie-lines is shown at the physical layer. The frequency and tie-line data are sent to the SCADA function layer through the measurement layer that encapsulates both the sensors and the communication channel. At the function layer, the data are first preprocessed and then fed into the control system. The output

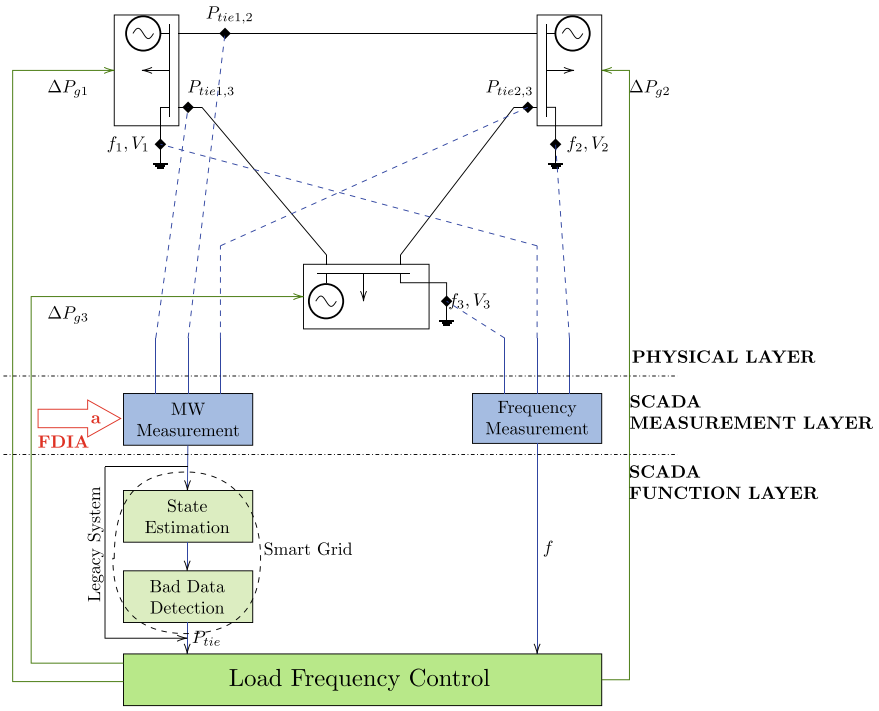



**Fig. 4.2** Attack on MW measurements on a grid system with bad data detection

of the function layer is then sent back to the physical layer through actuators. The following assumptions are made in the assessment:

1. Frequency Measurements: Directly used in the control system.
2. Tie-line Measurements: Undergo initial processing to detect Bad Data.
3. Attacker can spoof into the sensor measurements to get the data values.
4. Attacker does not have knowledge about the topology of the system.

Thus, penetration testing aims to show that an external entity can hack into the system, inject data into the Tie-line sensor measurements and bypass the Bad Data Detection (BDD).

If $z_i$ are the sensor measurements, the tie-line flows can be estimated using state estimation as

$$\hat{\mathbf{z}} = H\hat{\mathbf{x}} \tag{4.1}$$

BDD using the method of residues identifies an erroneous measurement by comparing the measurement residue with a threshold and raises an alarm if the residue is greater than the threshold.

$$r = \|\mathbf{z} - \hat{\mathbf{z}}\|_2 > \text{Threshold} \tag{4.2}$$

If the attack is modeled as $\mathbf{a} = \mathbf{Hc}$, the residue will be as shown

$$r_a = \|\mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{x}} + \mathbf{c})\|_2 = \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\|_2 = r \tag{4.3}$$

From (4.3), it is evident that the residue in the event of an attack is equal to the actual residue. Thus, to model an attack such that bad data detection tests fail, the adversary should have complete information about the measurement Jacobian H (Liu et al. 2009). This information is virtually unavailable to an external entity.

However, since the attack here is only on the tie-line measurements, we can determine **H** using the available measurements alone.

Let $\tilde{\mathbf{H}}$ be a matrix formed by collecting the rows of **H** that correspond to tie-lines. Since $\mathbf{a} = \tilde{\mathbf{H}}c$, it is evident that '**a**' is in the column space of $\tilde{\mathbf{H}}$. From the properties of matrices, '**a**' can be obtained as a linear combination of the basis vectors of $\tilde{\mathbf{H}}$. Let $z_i,\ i = 1, 2, ..., n_o$ be the tie-line measurements obtained at various time instants. Thus, **z** can be written as

$$\mathbf{z} = [z_1\ z_2\ \ldots z_{n_o}] \tag{4.4}$$

Since $\mathbf{z} = \tilde{\mathbf{H}}\mathbf{x}$, each of these observed measurements will be in the column space of $\tilde{H}$. Choose $n_t$ measurements (where $n_t$ is the number of tie-lines) from the total available measurements, $n_o$ such that these measurements are independent of each other, i.e.,

$$\begin{aligned} \text{Rank}(\mathbf{z}) &= n_t \\ \text{where, } \mathbf{z} &= [z_1\ z_2\ \ldots z_{n_t}] \end{aligned} \tag{4.5}$$

The above vectors thus form a basis for the matrix $\tilde{\mathbf{H}}$. This condition can be used to model the attack vector as

$$a = b_1\, z_1 + b_2\, z_2 + \ldots + b_{n_t}\, z_{n_t} \tag{4.6}$$

where $b_1, b_2, \ldots b_{n_t}$ are arbitrary values that depend on attack objective.

It is evident from the above analysis that a successful attack can be implemented on the system with limited system knowledge.

## 4.3 Cascading Outage Model

Before we move on to the Vulnerability Assessment framework, we first discuss the cascading analysis model in this section.

This study uses the Cascading Outage Simulator with Multiprocess Integration Capabilities (COSMIC) model (Song et al. 2016). This cascading outage model is represented using the following set of equations:

1. **Differential equations:** Used to simulate dynamic components in COSMIC, including rotating machines, exciters, and governors:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}(t), \mathbf{y}(t), \mathbf{z}(t)). \tag{4.7}$$

The vector $\mathbf{x}$ is used to represent continuous state variables such as voltage, and rotor angle that change with time according to a set of differential equations. These equations include swing equations, exciter equations and rotor angle equations.

2. **Nonlinear power flow equations:** Used to characterize the power flows:

$$\mathbf{g}(t, \mathbf{x}(t), \mathbf{y}(t), \mathbf{z}(t)) = \mathbf{0}. \tag{4.8}$$

The vector $\mathbf{y}$ represents the set of variables that vary according to algebraic relations such as real and reactive powers.

3. A series of equations that reflect the **distance to thresholds** that cause discrete variations are used to characterize discrete changes (such as component failures or load shedding), and load voltage responses are explicitly represented:

$$\mathbf{h}(t, \mathbf{x}(t), \mathbf{y}(t), \mathbf{z}(t)) < 0 \tag{4.9}$$

For example, if the 'line flow > threshold', then the corresponding entry in $\mathbf{z}$ changes state.

4. During an event, COSMIC employs a recursive approach to solve the differential algebraic equations (DAEs) while keeping an eye out for discrete events, such as those that divide the network into islands.

In Eqs. 4.7–4.9, $\mathbf{x}$ and $\mathbf{y}$ are the vectors of continuous state variables, and they change according to differential equation and algebraic equation, respectively. The variable $\mathbf{z}$ takes discrete values and changes state when the constraint $h_i(\cdots) < 0$ fails. The detailed equations are given in Appendix B.

Four different types of protective relays are modeled: (1) under-voltage load shedding (UVLS), (2) under-frequency load shedding (UFLS) relays for stress reduction, and (3) over-current (OC) and (4) distance (DIST) relays for transmission line protection. OC relays monitor the instantaneous current flow along each branch. DIST relays serve as a Zone 1 relay that keeps track of the transmission line's apparent admittance. The UVLS and UFLS relays operate at 0.9 pu and 0.95 pu, respectively, and shed 25% of load. The OC and TEMP relay settings are obtained from the line MVA ratings.

## 4.4 Vulnerability Assessment

To perform a Vulnerability Assessment, we summarize the attack and its impact on the grid as explained below.

"False data injected into the sensors, RTUs, or actuators leads to false generation dispatches. These dispatches are adjusted to create overloads over the system

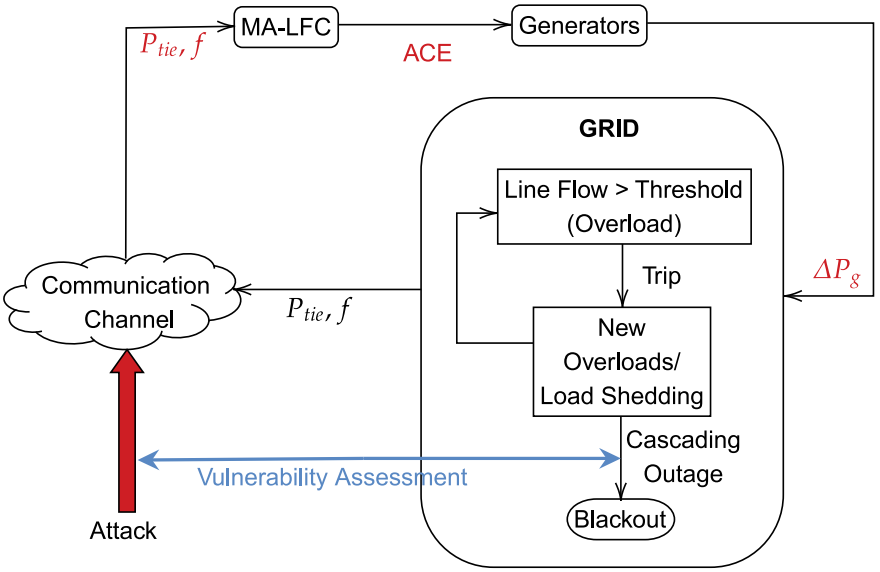



**Fig. 4.3** Attack propagation flow

transmission lines. Protective relays disconnect the overloaded lines. These disconnections lead to further overloads and outages, thus leaving the system under a state of blackout."

The above flow is represented in Fig. 4.3.

The Vulnerability Assessment involves three major steps (Vaiman et al. 2012):

1. Identifying the threats and vulnerabilities in the system.
2. Quantifying the risk due to the threats.
3. Prioritizing the risks.

### *4.4.1 Identification of Threats and Vulnerabilities*

Assuming that the threat is external and not internal to the system, the attack surface consists of frequency and tie-line sensors and the actuation signal, i.e., the Area Control Error.

### *4.4.2 Quantifying Risk*

VA defines risk as the product of an attack's likelihood and severity. This analysis assumes that frequency, tie-line, and actuator data are transmitted over a tainted

channel with an equal probability of being attacked. Since we are concerned with the cascading effect of attack on the grid in the shortest possible period, we propose two outage indices:

1. Load Shedding Index ($r_{LS}$): Defined as the ratio of Net Load Shedding to Blackout Time.
2. Line Outage Index ($r_{LO}$): Defined as the ratio of total line outages to Blackout Time.

The measurements have a different range of values; frequency change is around $10^{-2}$, and tie-line power change is around 10. We use mean normalization so that different attack vectors are updated to the same level for proper comparison. A maximum outage index will indicate a higher outage in a lower time. The contingency leading to the maximum outage index will be used to determine the risk.

The worst-case attack scenario would be to cause maximum damage with minimum attack effort or attack input. Thus, the final proposed net risk index is defined as a ratio of the outage factor to the normalized attack ($\bar{\mathbf{a}}$):

$$\text{Risk} = (r_{LS} + r_{LO})/\bar{\mathbf{a}} \tag{4.10}$$

### *4.4.3 Prioritizing the Risk*

After obtaining all risk indices, these vulnerabilities are prioritized based on their risk indices. This ranking can also be used to determine the design processes for mitigation solutions.

## 4.5 Detailed Risk Quantification Methodology

We evaluate the risk using cascading failure analysis (Baldick et al. 2008; Che et al. 2019). There are three stages involved in the quantification of risk.

### *4.5.1 Stage 1: Initiating Event Identification*

To determine the triggering event, we remove the network's connections sequentially and conduct a dynamic analysis using a differential equation analysis (DEA) algorithm. Each instant an element exceeds its threshold, the dynamic process is paused, the associated element is removed from network, and operations are resumed. The detailed procedure is depicted in Fig. 4.4.

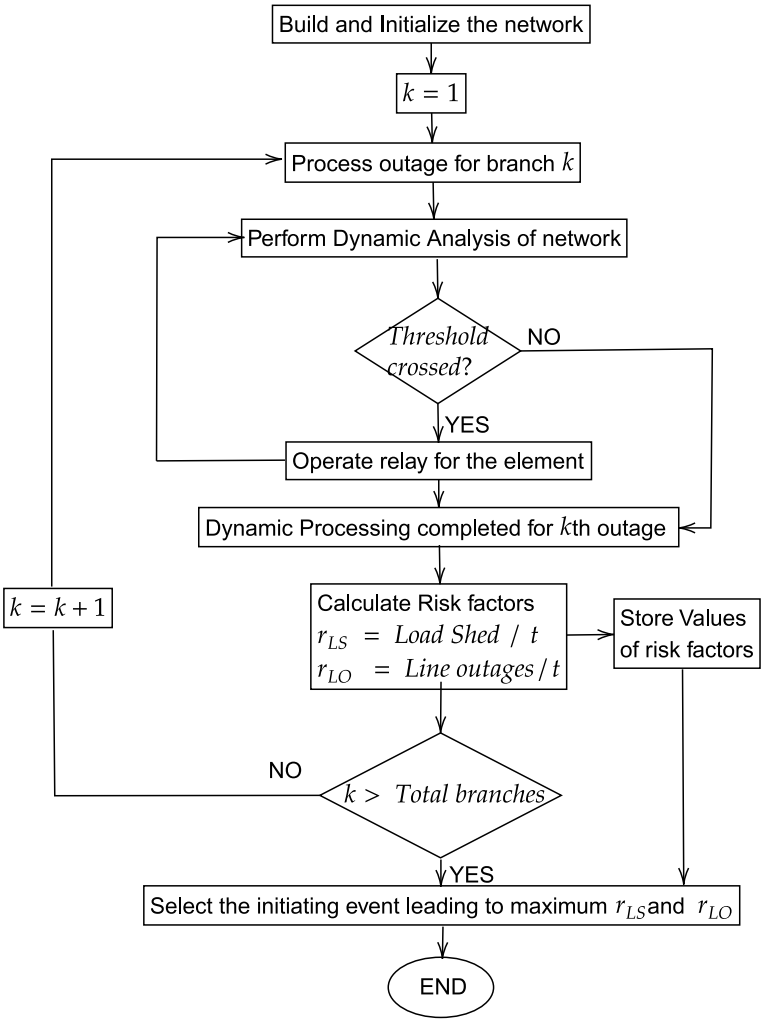



**Fig. 4.4** Flowchart for initiating event identification

### *4.5.2 Stage 2: Determination of Required Change in Generation*

The second phase involves determining the change in generation that can lead to the overloading of the line obtained in Stage 1. To achieve this, an optimization problem is formulated to find the minimum shift in generation from the current value that can lead to overloading of the lines. We utilize a Power Transfer Distribution

Factor (PTDF) matrix that gives the relationship between changes in generation and changes in line flows. Assuming that a line '$ij$' is obtained as the output of Stage 1, we formulate a corresponding optimization problem as given in (4.11):

$$\text{Minimize} \sum dP_{g_i}^2, i = 1, ..n_g \tag{4.11a}$$

$$\text{Subject to,} \begin{bmatrix} p_{ij_1} & p_{ij_2} & \dots & p_{ij_n} \end{bmatrix} \begin{bmatrix} dP_{g_1} \\ dP_{g_2} \\ . \\ . \\ dP_{g_n} \end{bmatrix} + P_{ij}^0 \geq P_{ij,max} \tag{4.11b}$$

$$P_{gmin} \leq dP_{g_i} + P_g^0 \leq P_{gmax} \tag{4.11c}$$

where $dPg_i$ is the change in generation for unit $i$. $P_{gmin}$ and $P_{gmax}$ are the generation limits which are fixed for a generating unit. $P_{ij,max}$ is the line limit for the line $ij$ of interest and $\begin{bmatrix} p_{ij1} & p_{ij2} & \dots & p_{ijn} \end{bmatrix}$ is the row of Power Transfer Distribution Factor (PTDF) matrix corresponding to line $ij$. $P_{ij,max}$ is based on the line thermal ratings and is constant for a given line.

The constraint (4.11b) is that the line flow should be greater than the line limit. The constraint (4.11c) sets the generator limits.

### *4.5.3 Stage 3: Optimal Attack Vector and Risk Calculation*

Using (7.1), the steady-state values of the generation shifts are calculated for a unit step change in the attack vectors. The magnitude of the step attack is calculated by averaging the maximum and minimum settling values. After determining the values for unit step attacks, the best attack vector is determined by solving the algebraic equations.

Finally, the value of risk is calculated using (4.10).

## 4.6 Case Study: Vulnerability Assessment for 9-Bus and 39-Bus New England Systems

### *4.6.1 Example 4.1: VA on 9-Bus System*

We first perform the VA on a 9-bus system with three generators. Each generator is assumed to be in one area. The 9-bus system is as shown in Fig. 4.5.

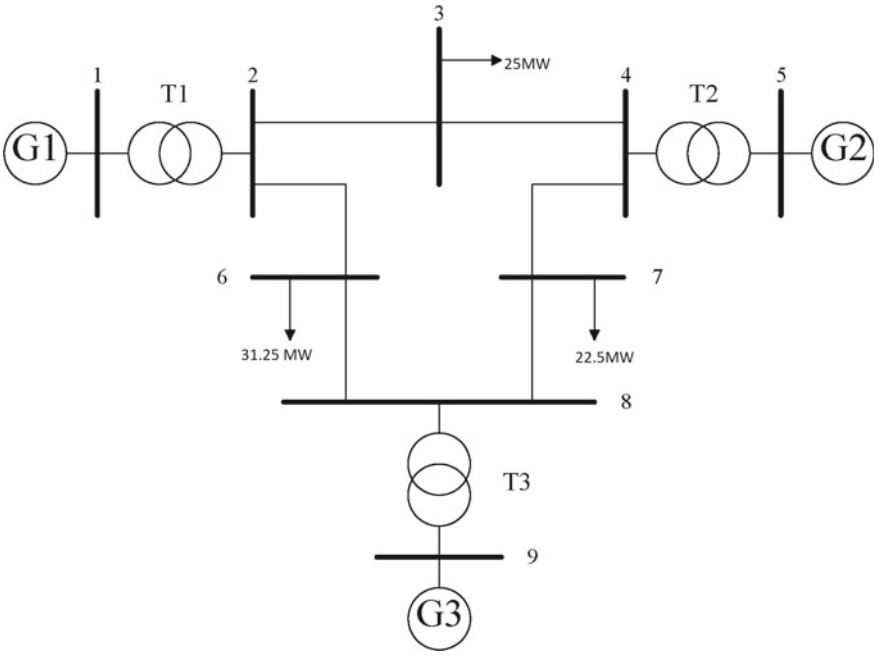



**Fig. 4.5** IEEE 9-bus system (Anderson and Fouad 2003)

**Table 4.1** Stage 1 Output: $r_{LS}$ for 9-bus system

| Branch outage in stage 1 | Time (s) | Total load shed (MW) | $r_{LS}$ |
|---|---|---|---|
| 4 (7–8) | 7.50 | 22.50 | 3.000 |
| 6 (2–6) | 7.50 | 31.25 | 4.167 |
| 8 (2–3) | 7.54 | 25.00 | 3.315 |

#### 4.6.1.1 Stage 1: Initiating Event Identification

Table 4.1 gives the Stage 1 output. It can be seen from the table that the outage of branch 6(2–6) leads to maximum load shedding and $r_{LS}$.

Thus, the line 2–6 outage is taken as the final output of Stage 1 as indicated in Fig. 4.6.

#### 4.6.1.2 Stage 2: Change in Generation

In the second stage, we find the value of $dP_{gi}$ for all generators such that line 2–6 (obtained from Stage 1) exceeds limits.

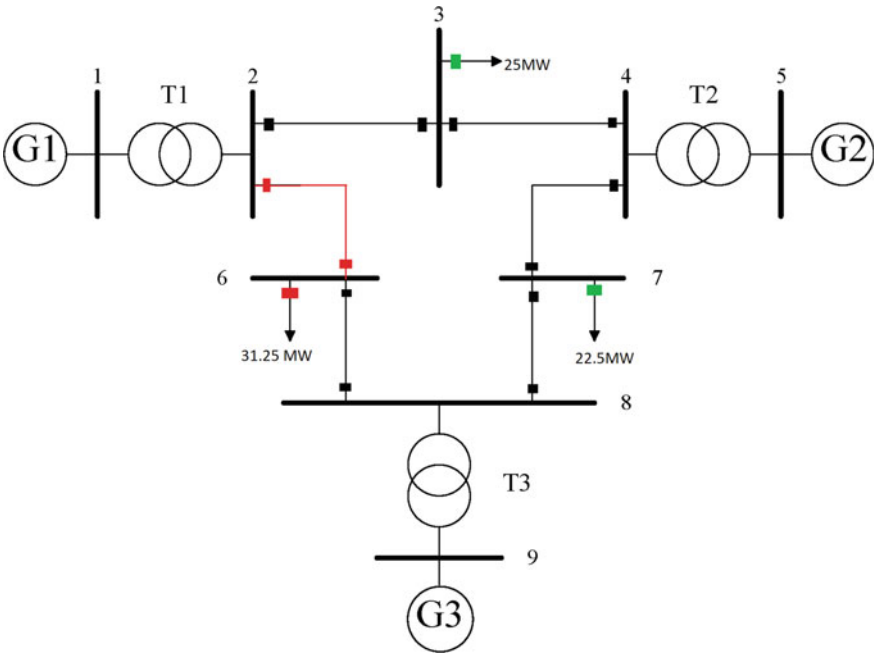



**Fig. 4.6** Stage 1 output: outage of line 2–6

**Table 4.2** Risk values for 9-bus system

| Attacked measurement | Actual attack values | Nominal attack values (pu) | Risk |
|---|---|---|---|
| Tie-line sensor | 96.8931 MW | 0.5014 | 3.9128 |
| | 13.0005 MW | 0.0673 | |
| | −96.3663 MW | −0.4986 | |
| Frequency sensor | 1.5254 Hz | 0.3374 | 1.3374 |
| | −0.3824 Hz | −0.0846 | |
| | 4.1381 Hz | 0.9154 | |
| ACE Actuator | 0.0085 | 0.0244 | 4.0765 |
| | −0.1520 | −0.4369 | |
| | 0.1958 | 0.5631 | |

#### 4.6.1.3 Stage 3: Optimal Attack and Risk Calculation

Finally, the attack vectors to be injected into different sensors to obtain this change in the generation are given in Table 4.2.

From Table 4.2 it is evident that the maximum risk is to the ACE Actuator, i.e., a very small attack value is sufficient to create the outage in the system. Since the

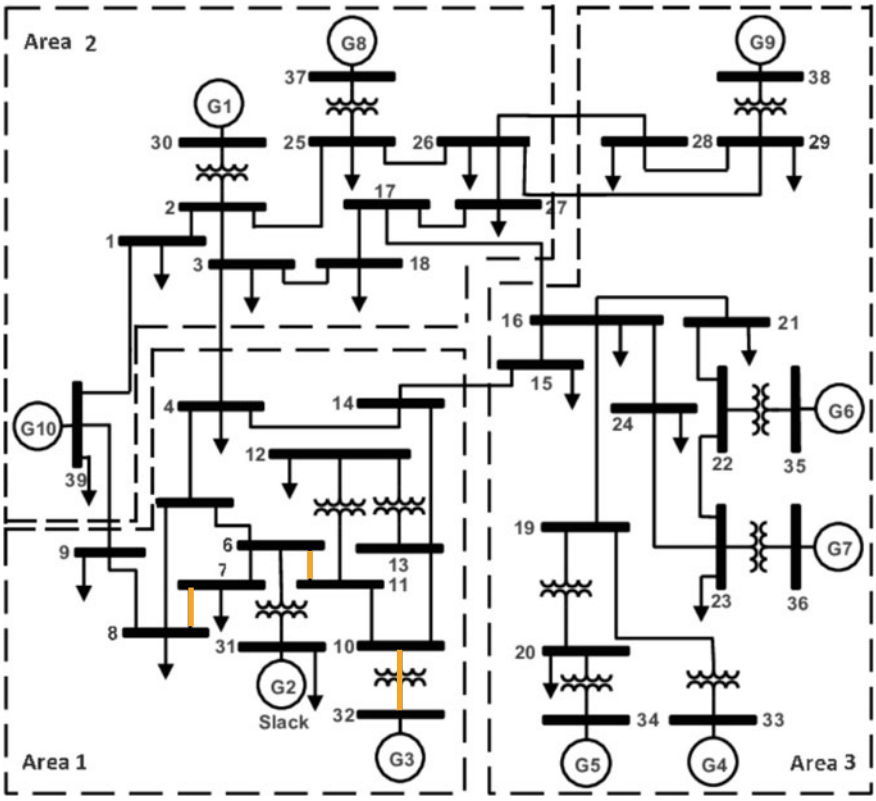



**Fig. 4.7** 39-bus 3 area New England test system (Bevrani 2014)

above system is small, the effects of line outages are not very evident. We, therefore, perform the VA on a more extensive system in the next section.

### *4.6.2 Example 4.2: Vulnerability Assessment for 39-Bus New England System*

In this section, we carry out the Vulnerability Assessment described above using the 39-bus New-England test system with 3 areas. It is the most common test system used in LFC analysis and studies (Rerkpreedapong et al. 2003). The load data were obtained from the New-England ISO website (England 2024). The bus, branch, and line ratings are obtained from the MATPOWER 39-bus data.

We carry out the Vulnerability Assessment on the 39-bus New England System shown in Fig. 4.7.

**Table 4.3** 39-bus system outage indices

| Branch outage in stage 1 | Time (s) | Total load shed (MW) | $r_{LS}$ | $r_{LO}$ |
|---|---|---|---|---|
| 12 (6–11) | 217.969 | 845.57 | 3.8793 | 0.0046 |
| 13 (7–8) | 217.969 | 845.57 | 3.8793 | 0.0046 |
| 16 (10–11) | 325.985 | 648.42 | 1.9891 | 0.0031 |
| 17 (10–13) | 89.510 | 0 | – | 0.0112 |
| 18 (13–14) | 89.839 | 714.8 | 7.9565 | 0.0111 |
| 23 (16–21) | 245.749 | 265.57 | 1.0806 | 0.0040 |
| 27 (21–22) | 19.157 | 150.62 | 7.8625 | 0.0522 |
| 29 (23–24) | 224.603 | 396.07 | 1.7634 | 0.0045 |
| 37 (6–31) | 60.727 | 316.07 | 5.2047 | – |
| 38 (10–32) | 53.852 | 566.07 | 10.5115 | – |
| 43 (25–37) | 81.996 | 316.07 | 3.8550 | – |

#### 4.6.2.1 Stage 1: Initiating Event Identification

Table 4.3 provides the quantity of load shedding, the duration of its occurrence, and the respective outage indices. Due to the failure of line 12 (linking buses 6–11) or 13 (linking buses 7–8) in the initial stage, the highest load shedding is 848.57MW; however, it takes around four minutes to impact. The maximum $r_{LS}$ is for the line 38 interruption (linking buses 10–32). In this instance, the net load shed is 566.07MW and transpires within less than one minute. Thus, this case advances to stage two.

#### 4.6.2.2 Stage 2: Change in Generation

In the second stage, we find the value of $dP_{g_i}$ for all generators such that line 10–32 (obtained from Stage 1) exceeds limits using (4.11).

#### 4.6.2.3 Stage 3: Optimal Attack and Risk Calculation

Based on the generation changes, the optimal attack vector and risk corresponding to each attack are calculated. The risk values corresponding to an attack on the sensor and actuators are presented in Table 4.4.

It can be observed from the risk levels that the most significant risk occurs when frequency sensors are attacked. It should be noticed that the attacked frequency values exceed the acceptable range. This high-frequency deviation may furthermore activate frequency relays, resulting in substantial damage. However, it is possible to identify such systemic abnormalities accurately.

**Table 4.4** Risk values for 39-bus system

| Attacked measurement | Actual attack values | Normalized attack values | Risk |
|---|---|---|---|
| Tie-line sensor | 172.1333 | 0.5591 | 4.0192 |
| | −12.0141 | −0.0390 | |
| | −135.7383 | −0.4409 | |
| Frequency sensor | 3.1772 | −0.5000 | 21.7197 |
| | 3.6680 | 0.4840 | |
| | 3.6760 | 0.5000 | |
| Actuator | 0.2946 | −0.5799 | 9.0632 |
| | 0.5029 | 0.1598 | |
| | 0.5761 | 0.4201 | |

## 4.7 Summary

This chapter describes the proposed cyber-attack Vulnerability Assessment on the Multi-Area Load Frequency Control of a power system. The first step is mathematical Data Penetration Testing, which shows that it is possible to model an attack that passes through the system detection tests unidentified. Thus the system can be affected by stealth attacks.

In the Vulnerability Assessment framework, risk indices are proposed that effectively capture the relation between attacks on the LFC and the cascading outages in the grid system. The vulnerability assessment is done on the modified IEEE 9-bus system and 39-bus New-England system, and an attack has been identified, leading the system into a blackout state. Since the frequency deviation is within limits, frequency-based actions are not taken.

The data obtained from the Vulnerability Assessment can further be used to design efficient attack identification and mitigation strategies.

## References

Anderson PM, Fouad AA (2003) Power system control and stability. IEEE Press, New York

Baldick R, Chowdhury B, Dobson I, Zhaoyang Dong, Bei Gou, Hawkins D, Huang H, Joung M, Kirschen D, Fangxing Li, Juan Li, Zuyi Li, Chen-Ching Liu, Mili L, Miller S, Podmore R, Schneider K, Kai Sun, Wang D, Zhigang Wu, Pei Zhang, Wenjie Zhang, Xiaoping Zhang (2008) Initial review of methods for cascading failure analysis in electric power transmission systems ieee pes cams task force on understanding, prediction, mitigation and restoration of cascading failures. 2008 IEEE power and energy society general meeting–conversion and delivery of electrical energy in the 21st Century. Pittsburgh, PA, USA, pp 1–8

Bevrani H (2014) Robust power system frequency control. Springer, Berlin

Che L, Liu X, Li Z, Wen Y (2019) False data injection attacks induced sequential outages in power systems. IEEE Trans Power Syst 34(2):1513–1523

England IN Reliable electricity. Competitive prices. Clean-energy transition
Liu X, Shahidehpour M, Li Z, Liu X, Cao Y, Li Z (2017) Power system risk assessment in cyber attacks considering the role of protection systems. IEEE Trans Smart Grid 8(2):572–580
Liu Y, Ning P, Reiter MK (2009) False data injection attacks against state estimation in electric power grids. In: Proceedings of the 16th ACM conference on computer and communications security, CCS '09, pp 21–32
Rerkpreedapong D, Hasanovic A, Feliachi A (2003) Robust load frequency control using genetic algorithms and linear matrix inequalities. IEEE Trans Power Syst 18(2):855–861
Song J, Cotilla-Sanchez E, Ghanavati G, Hines PDH (2016) Dynamic modeling of cascading failure in power systems. IEEE Trans Power Syst 31(3):2085–2095
Vaiman B, Chen C, Dobson H, Papic M, Zhang (2012) Risk assessment of cascading outages: Methodologies and challenges. IEEE Trans Power Syst 27(2):631–641

# Chapter 5
# MITRE ATT&CK for Smart Grid Cyber-Security


**Abstract** The MITRE ATT&CK framework is a recent web-based tool that has been widely used in the field of cyber-security. It is popular due to its open-source and crowd-sourced nature and comprehensive cataloging of adversary tactics and techniques that are used to launch successful attacks. This chapter delves into the integration of the MITRE ATT&CK framework into the realm of power systems, with a particular focus on smart grids. In this chapter, we discuss in detail the usage of MITRE ATT&CK framework for threat analysis in power systems.We begin by mapping the Threats in MITRE to Smart Grids which can help in detection and mitigation planning. We combine the MITRE framework with a probabilistic approach to rank the attack points. A practical illustration for VAPT using MITRE framework is provided using the substation automation system as a case study. The approach used in the book can be further extended to other power system and industrial control system applications for vulnerability assessment and penetration testing. Overall, this chapter serves as a comprehensive guide for security practitioners, researchers, and stakeholders seeking to fortify power systems against cyber threats, harnessing the analytical power of the MITRE ATT&CK framework to safeguard critical infrastructure in an increasingly digital landscape.

**Keywords** MITRE ATT&CK · Attack tactics · VAPT · Substation Automation Systems (SAS) · Industrial control security

## 5.1 Introduction

In this age of digital transformation, the concept of cyber-security has become paramount, and innovative tools and strategies are imperative to protect critical infrastructure. One of such tools that has gained prominence in the cyber-security landscape is the MITRE ATT&CK framework. Originally developed as a knowledge base to understand the tactics, techniques, and procedures (TTPs) used by adversaries in the realm of IT security, MITRE ATT&CK has transcended its roots to find valuable applications in diverse domains, including the protection of smart grids.



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_5

MITRE ATT&CK, an acronym for Adversarial Tactics, Techniques, and Common Knowledge, is a comprehensive framework that meticulously documents the modus operandi of cyber adversaries. Its strength lies in its ability to categorize and describe the myriad tactics and techniques used by malicious actors during various stages of a cyber attack, from initial reconnaissance to data exfiltration. By providing a standardized and structured taxonomy of adversarial behaviors, MITRE ATT&CK equips cyber-security professionals with a powerful tool to understand, detect, and mitigate cyber threats.

In the context of smart grid cyber-security, the MITRE ATT&CK framework takes on a new and indispensable role. Smart grids, which encompass a multitude of interconnected devices, communication networks, and software systems, are fertile grounds for cyber threats. As the backbone of a nation's critical infrastructure, the smart grid plays a pivotal role in delivering reliable and efficient electricity to consumers. However, the interconnectivity and reliance on digital technologies in these systems also expose them to a broad spectrum of cyber-security challenges.

The vulnerabilities within the smart grid ecosystem underscore the urgent need for robust cyber-security measures. Traditional security models are no longer sufficient in this dynamic and ever-evolving landscape. Here, MITRE ATT&CK emerges as a beacon of insight and resilience. In the following sections of this chapter, we will explore how MITRE ATT&CK can be tailored and applied to the unique challenges of smart grid cyber-security, offering a structured approach to understanding and countering cyber threats in this critical domain.

## 5.2 Understanding MITRE ATT&CK

In the realm of cyber-security, where the landscape is ever-evolving and adversaries continuously adapt their tactics, techniques, and procedures (TTPs), understanding and countering these threats is an ongoing challenge.

MITRE ATT&CK is a comprehensive knowledge base developed by the MITRE Corporation, a not-for-profit organization with a long-standing history of conducting research and development in various technology domains. The framework's primary objective is to codify and categorize the behaviors and methodologies employed by cyber adversaries during the various stages of a cyber attack.

MITRE ATT&CK is structured around two primary components:

1. **Tactics**: These are the high-level objectives that adversaries aim to achieve during an attack. Examples of tactics include initial access, execution, persistence, and exfiltration. Each tactic represents a key phase in the attack lifecycle.
2. **Techniques**: Techniques are the specific methods and procedures employed by adversaries to accomplish their tactical objectives. Each technique falls under one of the tactics and is accompanied by detailed descriptions, examples, and potential detection strategies.

The MITRE ATT&CK framework is not limited to a specific industry or technology stack. Instead, it provides a common language and taxonomy that enables cyber-security professionals to

- **Understand Threat Behaviors**: By detailing the tactics and techniques used by adversaries, MITRE ATT&CK offers a deep understanding of how cyber attacks are orchestrated.
- **Enhance Threat Detection**: Organizations can leverage the framework to enhance their threat detection capabilities, as it provides guidance on recognizing and countering specific techniques.
- **Improve Incident Response**: When a cyber incident occurs, MITRE ATT&CK assists in incident response efforts by helping teams understand the nature of the attack and its potential impact.

### *5.2.1 Evolution of MITRE ATT&CK*

MITRE ATT&CK was initially developed by MITRE Corporation in 2013 as a research project focused on documenting cyber adversary behavior. Over time, it has evolved into a globally recognized and widely adopted framework. It is important to note that MITRE ATT&CK is continually updated and refined to reflect emerging threats and changing adversary tactics.

The versatility of MITRE ATT&CK is a key reason for its widespread adoption. It has found utility across various cyber-security domains, including, but not limited to,

- **Threat Intelligence**: Security analysts and threat intelligence teams use MITRE ATT&CK to map observed adversary behaviors to known tactics and techniques, aiding in attribution and identifying patterns.
- **Security Assessment and Red Teaming**: Organizations employ MITRE ATT&CK to assess their security posture and test defenses by simulating real-world attacks, known as red teaming exercises.
- **Security Operations**: Security operation centers (SOCs) leverage MITRE ATT&CK to enhance monitoring, alerting, and incident response capabilities.
- **Compliance and Frameworks**: MITRE ATT&CK is increasingly referenced in industry-specific cyber-security regulations and frameworks, making it a valuable resource for compliance efforts.

### *5.2.2 Relevance of MITRE ATT&CK in Smart Grid Cyber-Security*

The smart grid, with its intricate blend of hardware, software, and critical infrastructure, introduces a unique set of cyber-security challenges. The application of MITRE ATT&CK to smart grid cyber-security allows for a structured and standardized approach to understanding, detecting, and mitigating threats specific to this domain.

Given the critical nature of smart grid operations and the potential consequences of cyber attacks on power generation, distribution, and control systems, the relevance of MITRE ATT&CK cannot be overstated. In the following sections, we will explore how MITRE ATT&CK can be adapted and customized to address the distinctive challenges faced by smart grid cyber-security professionals, providing a robust framework for defending against evolving threats in this critical sector.

The integration of the MITRE framework in the Smart Grid VAPT design is relevent due to several reasons:

1. MITRE ICS framework is the most comprehensive framework available for threat modeling.
2. The MITRE matrices give a good correlation between tactics, techniques, and mitigations.
3. It is comprehensive and analyzes security properties against each system component
4. It is a widely accepted framework for ICS threat modeling and vulnerability assessment.

## 5.3 Mapping Threats to Smart Grids

The different attacker goals used in the MITRE framework are called tactics. The various tactics that are used in the Smart Grid attack model are as follows.

1. Initial Access: The Initial access can be gained into the substation or control center either from outside or through an internal device. This step is used to gain an initial foothold.
2. Execution: Once the attacker is inside the substation, he can start executing commands to disrupt the actual behavior of the system.
3. Persistence: The attacker then modifies programs and configurations to continue to be in the system and maintain access.
4. Privilege Escalation: An attacker may not have all the permissions required during initial access. Once inside, he can enter more secure and critical data and controls of the HMI and IEDs.
5. Evasion: The attackers can evade various detection and protection methods, such as firewalls, by spoofing communication and exploiting software vulnerabilities.
6. Discovery: Remote discovery can be used to understand the control center topology. This discovery can help the proper subsequent movement to reach the desired target.
7. Lateral Movement: Once the attacker has the complete control center data and topology, he can devise methods to move from the initial access points to the targets.
8. Collection: At each intrusion point, the attacker collects data that can be used to exploit the controls of the control center.
9. Command and Control: Using the data at hand, the attacker finally implements control commands to alter the working of the controls.

| Initial Access | Execution | Persistence | Privilege Escalation | Evasion | Discovery | Lateral Movement | Collection | Command and Control | Inhibit Response Function | Impair Process Control |
|---|---|---|---|---|---|---|---|---|---|---|
| Exploit Public-Facing Application | Change Operating Mode | Modify Program | Exploitation for Privilege Escalation | Change Operating Mode | Network Connection Enumeration | Default Credentials | Automated Collection | Commonly Used Port | Activate Firmware Update Mode | Brute Force I/O |
| Exploitation of Remote Services | Command-Line Interface | Module Firmware | | Exploitation for Evasion | Network Sniffing | Exploitation of Remote Services | Data from Information Repositories | Connection Proxy | Alarm Suppression | Modify Parameter |
| External Remote Services | Execution through API | Project File Infection | | Indicator Removal on Host | Remote System Discovery | Lateral Tool Transfer | Detect Operating Mode | Standard Application Layer Protocol | Block Command Message | Module Firmware |
| Internet Accessible Device | Graphical User Interface | System Firmware | | Masquerading | Remote System Information Discovery | Program Download | I/O Image | | Block Reporting Message | Spoof Reporting Message |
| Remote Services | Modify Controller Tasking | Valid Accounts | | Rootkit | | Remote Services | Man in the Middle | | Block Serial COM | Unauthorized Command Message |
| Replication Through Removable Media | Native API | | | Spoof Reporting Message | | Valid Accounts | Monitor Process State | | Data Destruction | |
| Rogue Master | Scripting | | | | | | Point & Tag Identification | | Denial of Service | |
| Spearphishing Attachment | User Execution | | | | | | Program Upload | | Device Restart/Shutdown | |
| Supply Chain Compromise | Modify Controller Tasking | | | | | | | | Manipulate I/O Image | |
| Transient Cyber Asset | | | | | | | | | Modify Alarm Settings | |
| | | | | | | | | | Rootkit | |
| | | | | | | | | | Service Stop | |
| | | | | | | | | | System Firmware | |

**Fig. 5.1** MITRE ATT&CK matrix for substation automation system

10. Inhibit Response Function: Commands can also be introduced to inhibit safety controls and functions that respond during an emergency situation.
11. Impair Process Control: Finally, the attacker can disable or even damage the complete physical process of the control center, leading to an outage of the entire system.

Under each tactic, some subcategories are identified, and the final matrix for the Smart Grid will be similar to Fig. 5.1.

### *5.3.1 Mapping Attacks to the MITRE Framework*

Let's examine how real-world cyber attacks on smart grids can be mapped to the MITRE ATT&CK framework:

**Stuxnet Worm**: Stuxnet is a notorious example of an attack on industrial control systems (ICS) similar to those used in smart grids. It employed various techniques, such as spear-phishing (a technique categorized under 'Initial Access' in MITRE ATT&CK) to infect systems. Stuxnet's payload included zero-day exploits ('Exploitation of Vulnerability' tactic) to manipulate programmable logic controllers (PLCs) and disrupt uranium enrichment facilities.

**Ukraine Power Grid Attack**: In December 2015 and 2016, Ukraine experienced multiple power outages due to cyber attacks. These attacks, attributed to APT groups, involved techniques like remote access ('External Remote Services' tactic) and disabling protective relays ('Impair Process Control' tactic). These tactics are well-documented within MITRE ATT&CK.

**BlackEnergy Malware**: The BlackEnergy malware was responsible for a cyber attack on Ukraine's power grid. It utilized spear-phishing ('Initial Access' tactic) and exploited vulnerabilities ('Exploitation of Vulnerability' tactic) to gain access to

critical systems. It also used a 'KillDisk' component to destroy data, falling under the 'Impact' tactic.

**NotPetya Ransomware**: Although NotPetya was initially a ransomware attack, it quickly propagated across networks, disrupting systems worldwide. It leveraged credential theft ('Credential Access' tactic) and lateral movement ('Lateral Movement' tactic) to spread within networks.

These examples illustrate how real-world cyber attacks on smart grids align with the MITRE ATT&CK framework's tactics and techniques. By mapping such attacks to MITRE ATT&CK, smart grid defenders can better understand the adversary's behavior, enhance threat detection and response, and fortify their cyber-security measures against evolving threats. This mapping enables a structured approach to safeguarding the critical infrastructure of the smart grid and maintaining reliable electricity delivery to consumers.

## 5.4 Using MITRE ATT&CK for Smart Grid Defense

After completing the threat identification phase using the MITRE ATT&CK framework, the next crucial step is to use the insights gained to develop effective mitigation strategies for the identified vulnerabilities in your smart grid environment. Here's how MITRE ATT&CK can be leveraged for this purpose.

### *5.4.1 Tactic/Technique-Based Mitigation*

Tactic/technique-based mitigation can be achieved by tactic mapping in combination with the mitigation library:

**Tactic Mapping**: MITRE ATT&CK categorizes adversary behaviors into tactics and techniques. Review the tactics associated with the identified threats. For instance, if you've identified a threat under the 'Execution' tactic, focus on mitigations related to that specific tactic.

**Mitigation Library**: MITRE ATT&CK offers a mitigation section that provides recommendations for countering each technique. Explore this library to identify relevant mitigations that align with the tactics used by potential adversaries.

### *5.4.2 Customized Mitigation Strategies*

Once the threats are identified, tailor-made strategies can be implemented to overcome the identified vulnerabilities:

**Tailor Mitigations**: Recognize that not all MITRE ATT&CK-recommended mitigations may be applicable or feasible in your specific smart grid context. Customize the mitigation strategies to align with the unique characteristics of your grid.

**Prioritization**: Assess the criticality and potential impact of each vulnerability and prioritize mitigation efforts accordingly. Some vulnerabilities may require immediate attention, while others can be addressed over time.

### 5.4.3 Incident Response Playbooks

Develop incident response playbooks based on MITRE ATT&CK insights. These playbooks should outline step-by-step procedures to respond to specific tactics and techniques. Include details on how to identify, mitigate, and recover from each threat.

### 5.4.4 Continuous Monitoring and Testing

Even after the application of defense strategies, the MITRE framework can be leveraged for continuous monitoring.

**Ongoing Assessment**: Continuously monitor your smart grid environment for emerging threats and vulnerabilities. Regularly revisit the MITRE ATT&CK framework to update your mitigation strategies in response to evolving threats.

**Red Teaming and Testing**: Conduct red-teaming exercises that simulate real-world attacks based on MITRE ATT&CK tactics and techniques. Use the results to validate the effectiveness of your mitigation strategies and identify areas for improvement.

### 5.4.5 Vendor and Technology Collaboration

Vendors and product specialists play an important role in the cyber-security of grid systems.

**Engage Vendors**: Collaborate with technology vendors and solution providers to implement security features and updates that align with MITRE ATT&CK-based mitigations.

**Security Training**: Ensure that staff members are well trained in cyber-security best practices and are aware of the MITRE ATT&CK framework to effectively implement and manage mitigation strategies.

### 5.4.6 Documentation and Compliance

Maintain detailed records of the mitigation strategies implemented in your smart grid environment. This documentation is essential for audit purposes and to demonstrate compliance with regulatory requirements.

By systematically applying MITRE ATT&CK-based mitigation strategies, you can significantly improve the security posture of your smart grid infrastructure and reduce the risk of cyber attacks. Keep in mind that cyber-security is an ongoing process, and staying vigilant, adapting to emerging threats, and continuously improving your mitigation measures are essential for protecting critical infrastructure.

## 5.5 MITRE ATT&CK for Vulnerability Assessment and Penetration Testing (VAPT)

In this section, we present a case study of employing MITRE ATT&CK for enhancing the cyber-security of a smart grid infrastructure. This case study has been adopted with permission from a cyber-security organization, Gridsentry (gs).

The case study combines multiple standards in combination with MITRE threat modeling for VAPT assessment of a Substation Automation System which is similar to the control center. A similar procedure can be adopted by various utilities for the VAPT of their substations or control centers.

The proposed VAPT process for a substation has the following steps:

1. **Map**: Map exploits to **MITRE** threat model for vulnerability assessment.
2. **Analyze**: Determine the likelihood, impact, and risk scores using **NIST-800-30** and attack trees.

### *5.5.1 Mapping of Exploits to MITRE ATT&CK*

Figure 5.2 gives a high-level flow of how the exploit is executed.

Each node in the high-level tree can be accomplished using a chain of events derived using the ATT&CK matrix. For example, to implement a Denial of Service (DoS) exploit at the station bus, the first step is to gain access to the station switch or LAN. This step can be mapped to the initial access steps as below.

$$\begin{aligned}(\textit{Exploit Remote Services})\, OR\, (\textit{Internet Accessible Device})\\ OR\, (\textit{Replication through removable media})\\ OR\, (\textit{Spearphishing Attachment})\end{aligned}$$

Similarly, the next step can be further mapped to other cells of the matrix to obtain the detailed or lower level attack flow.

## 5.6 Analyze the Likelihood, Impact, and Risk Scores

Once attacks are mapped to the threat model to create a cyber kill chain, the next step is to analyze the likelihood and impact, leading to the final risk scores for each exploit. This process of risk assessment is carried out using the steps below

1. Create Attack trees using the MITRE mapping (lower levels) and power system knowledge (upper levels).
2. Assign likelihood scores for each node using CVSS and NESCOR-based scoring.
3. Calculate net likelihood of tree using probability theory.

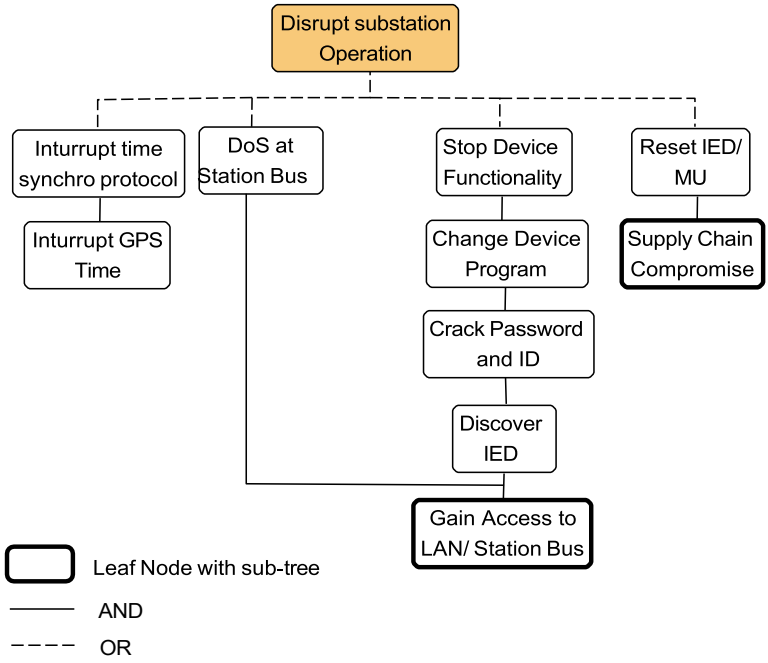



**Fig. 5.2** High-level attack tree

4. Assign impact scores for the final nodes in each tree.
5. Combine the impact and likelihood scores to arrive at the final risk value using NIST-800 guidelines.

The most important step in the risk assessment is the designing or construction of attack trees.

### *5.6.1 Substation Attack Trees*

Various attack trees are built to achieve attacker goals based on the tactics and techniques discussed in previous sections. Attack trees are an efficient way to represent the movement of attacks from their initial onset until the final attack impact. For example, Fig. 5.2 shows a part of the attack tree that causes a substation disruption.

Attack tree in Fig. 5.2 shows various paths through which the primary goal of disrupting substation can be achieved. Every node in the attack tree can be a part of multiple attack trees. To reach each node, there can be further sub-trees. Leaf nodes with further sub-branches are indicated with a bold outline as shown for the 'Gain Access to station bus'.

**Table 5.1** Impact scores

| Final node | Impact score | Color |
|---|---|---|
| Tripping of critical assets like feeder, transformer | Very high | Red |
| Tripping of non-critical assets | High | Orange |
| Tripping of bay-level devices | Moderate | Yellow |
| Tripping of ancillary services/Data collection | Low | Blue |
| No impact on physical layer | Very low | Green |

Each node can now be assigned likelihood values based on NESCOR guidelines. These individual scores can be used to determine the net likelihood. The determination of the net likelihood is discussed in detail in the following subsection.

### *5.6.2 Impact Scores*

The impact score is assigned based on the final node of the attack tree that is reachable by a penetration testing module. The impact scores are assigned, and the nodes in attack trees are colored based on the values in Table 5.1.

The system impact will be highest if the attack leads to the failure of critical devices in the substations, such as a breaker or transformer failure of critical feeders, due to which we assign the maximum impact score for these failures. Tripping of bay-level devices has a medium impact, and, finally, tripping of ancillary services is given a low score. There can also be attacks on the system that do not impact the physical system but may impact the IT system to get data. Such attacks fall into the very low-impact category.

It is possible that due to implementation of security features in a substation, the penetration testing penetrates only until an intermediate leaf of the attack tree. Thus, the impact score of this leaf node will be considered for final risk calculation.

### *5.6.3 Likelihood Scores*

The likelihood scores are assigned in different steps. Firstly, likelihood scores are given for each leaf node. Then, probability theory calculates a net likelihood score for the entire attack tree. The net likelihood score is then designated as High, Low, and Medium based on the range in which the values fall.

The likelihood scores for each leaf node are decided using NESCOR Electric Sector Failure Scenarios and Impact Analyses (NESCOR 2015). CVSS scores are another means for assigning the scores. Comparing CVSS and NESCOR 2015 shows

**Table 5.2** Likelihood scores

| Criterion | Sub-criterion | Scores |
|---|---|---|
| Skill required | Deep knowledge on domain and cyber attacks | 0.3 |
| | Insider knowledge required | 0.5 |
| | Basic domain and cyber skills | 0.9 |
| Accessibility | High expertise required to gain access | 0.3 |
| | Publicly accessible but not known commonly | 0.5 |
| | Common knowledge | 0.9 |
| Attack vector | Attack knowledge available theoretically | 0.3 |
| | Past history of attack but with no attack scripts available | 0.5 |
| | Attack scripts/tools directly available in public domain | 0.9 |
| Common vulnerability | Isolated occurrence | 0.3 |
| | More than one utility | 0.5 |
| | More than half of the utilities | 0.9 |

that all the parameters used are similar, with only name changes. The individual leaf likelihood scores are decided based on Table 5.2, which is derived from the NESCOR.

The NESCOR document scores the likelihoods with discrete values (0, 1, 3, and 9). We combine specific categories and further divide the values by a factor of 10 to arrive at the values in Table 5.2. The division by ten is used to obtain probability-like values. This division helps calculate the net likelihood using the principles of probability theory.

#### 5.6.3.1 Net Likelihood Calculation

The overall likelihood scores are then generated based on probability theory. It is considered that each event or leaf node leading to the attack is independent of the other. Thus, the likelihoods of nodes combined using an OR are added together, and the likelihoods combined using AND are multiplied to arrive at the net probability for each high-level node. The net likelihood is then divided into different categories using Table 5.3.

**Table 5.3** Net likelihood scores

| Final net likelihood | Likelihood score |
|---|---|
| ≥0.7 | Very high |
| 0.5–0.7 | High |
| 0.3–0.5 | Moderate |
| 0.1–0.3 | Low |
| ≤0.1 | Very low |

**Table 5.4** Risk: combination of likelihood and impact

| Likelihood scores/Impact scores | Very high | High | Moderate | Low | Very low |
|---|---|---|---|---|---|
| Very high | Very high | High | Moderate | Low | Very low |
| High | Very high | High | Moderate | Low | Very low |
| Moderate | High | Moderate | Moderate | Low | Very low |
| Low | Moderate | Low | Low | Low | Very low |
| Very low | Low | Low | Very low | Very low | Very low |

### *5.6.4 Risk Scores*

The risk scores are a combination of the likelihood and the impacts. This is obtained using the NIST Guide for Conducting Risk Assessments (NIST-SP-800-30) (Division, 2012). The likelihood and impact scores can be combined as shown in Table.5.4 to obtain the risk scores.

The final risk scores determine the security levels of the system. These scores will change as new security systems are introduced into the system. Thus the VAPT is a continuous process.

## 5.7 Case Study: MITRE ATT&CK for Substation VA

In the example case considered here, the aim of the penetration testing is to Open/Close the circuit breaker. This is achieved using a data manipulation attack-based penetration testing. Figure 5.3 shows all the steps involved in the VAPT process.

We perform a network penetration testing at the station bus level. The exploit captures the packets, modifies them, and re-injects it into the network to give wrong commands for operation. The exploit is mapped to the MITRE ATT&CK to get the detailed attack tree as shown in Fig. 5.3. The bold numbers in Fig. 5.3 are the likelihood scores assigned to each leaf node.

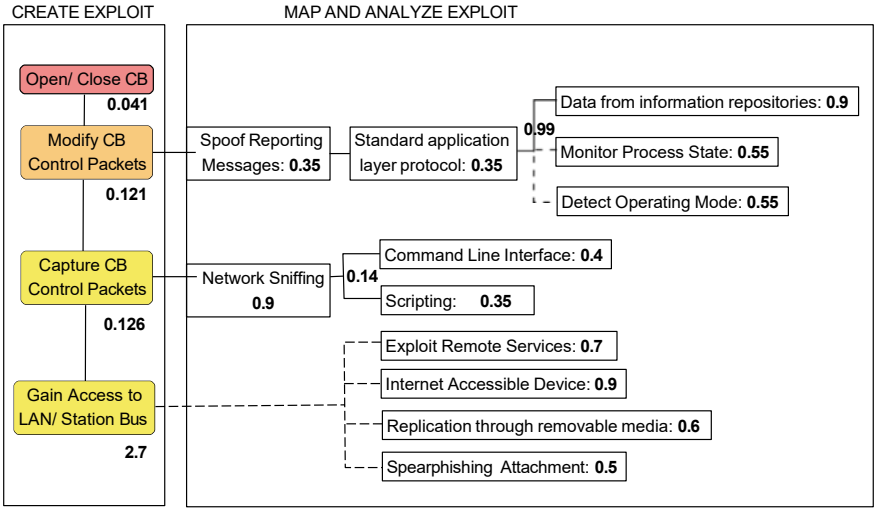



**Fig. 5.3** VAPT for case 1

### *5.7.1 Attack Penetrates to Final Node*

Assuming that the circuit breaker is connected to a critical feeder line, the impact score assigned to the final event is 'Very High'.

The final likelihood score of 0.041 is obtained using probability theory as shown in (5.1)

$$Net\ Likelihood = 2.7 * 0.128 * 0.121 = 0.041 \tag{5.1}$$

The final likelihood is thus 'Very Low' and Impact is 'Very High'. Thus, the final risk is obtained from Table 5.4 as 'Low'. The risk is low since it is a complicated attack.

### *5.7.2 Attack Stops at Capture Packets*

Let us assume that there are security features in the substation that prevent the spoofing of message packets. This could be achieved using Intrusion Detection Systems and Intrusion Protection Systems. In this case, the attack is only able to reach until the second stage giving us a net likelihood value of 0.3456 as given in (5.2)

$$Net\ Likelihood = 2.7 * 0.128 = 0.3456 \tag{5.2}$$

The final likelihood is thus 'Moderate' and Impact is 'Low'. Thus, the final risk is obtained from Table 5.4 as 'Low'. The risk is low since the attack does not impact any critical devices.

## 5.8 Summary

This chapter underscores the critical role of MITRE ATT&CK in fortifying smart grid cyber-security, with several key takeaways:

1. **Enhancing Cyber-Security**: MITRE ATT&CK provides a comprehensive framework to identify, categorize, and counteract cyber threats in smart grid systems. Its structured approach aids in understanding and addressing vulnerabilities effectively.
2. **Threat Mapping**: MITRE ATT&CK enables the mapping of specific threats and attack vectors to smart grid systems, making it easier to recognize potential risks and vulnerabilities unique to the energy sector.
3. **Proactive Defense**: By using MITRE ATT&CK, organizations can implement proactive threat detection and defense strategies. This approach shifts the focus from reactive measures to proactive threat hunting and mitigation.
4. **Adaptive Defense**: Smart grids evolve, and so do cyber threats. MITRE ATT&CK's adaptability allows for continuous monitoring and adaptation to emerging threats, enhancing the long-term resilience of smart grid systems.
5. **Collaboration and Knowledge Sharing**: The framework encourages collaboration among organizations and industries, fostering knowledge sharing and collective defense against cyber threats.

For practitioners and policymakers in the smart grid industry, the following insights and recommendations are offered:

1. **Implement MITRE ATT&CK**: Organizations should consider integrating MITRE ATT&CK into their cyber-security strategies. It serves as a powerful tool for assessing, planning, and implementing defense mechanisms.
2. **Training and Awareness**: Adequate training and awareness programs should be conducted to educate personnel about MITRE ATT&CK and its application in smart grid security.
3. **Incident Response**: Develop and refine incident response plans that incorporate MITRE ATT&CK to ensure efficient and effective responses to cyber incidents.
4. **Regulations and Standards**: Policymakers should consider incorporating MITRE ATT&CK into regulatory frameworks and industry standards to promote its widespread adoption and ensure a consistent approach to cyber-security in the smart grid sector.
5. **Collaboration**: Encourage collaboration among utilities, vendors, and government agencies to share threat intelligence and best practices, leveraging MITRE ATT&CK as a common language for communication.

6. **Continuous Improvement**: Recognize that cyber-security is an evolving field. Regularly update and adapt strategies based on the evolving threat landscape and insights gained from MITRE ATT&CK assessments.

In conclusion, MITRE ATT&CK offers a powerful framework for strengthening smart grid cyber-security by enabling threat identification, proactive defense, and collaborative efforts. Its integration and adoption can significantly enhance the resilience of smart grid systems in the face of evolving cyber threats.

## References

Division CS (2012) Guide for conducting risk assessments. In: National institute of standards and technology (NIST), p I1

NESCOR (2015) Electric sector failure scenarios and impact analyses-version 3.0. In: National electric sector cybersecurity organization resource, p 8

# Part III
# Attack Detection and Mitigation

# Chapter 6
# Signal Processing-Based Attack Detection


**Abstract** The chapter delves into an innovative approach to improve the security of smart grid systems through signal processing techniques. The chapter begins with an insightful introduction, highlighting the pressing need for robust attack detection mechanisms in the evolving landscape of smart grids. It then unfolds a multi-level attack detection strategy, emphasizing the importance of a comprehensive defense framework. Singular Spectral Analysis (SSA) emerges as a key player, and its application in attack detection is thoroughly explored. Further, the focus extends to multivariate SSA for control center-level detection, showcasing extensions in both training and detection phases. The chapter meticulously evaluates the performance of the detection algorithm, with a dedicated section on performance enhancement strategies. The heart of this chapter lies in presenting real-world results of multi-level attack detection, including at the RTU/IED and control center levels. Hypothesis testing-based attack detection, particularly SSA Hoeffding Test, takes the stage, accompanied by adaptive threshold selection techniques. The results of adaptive attack detection are dissected, including performance under load variations, comparisons with existing strategies, and scalability evaluations.

**Keywords** Singular spectrum analysis · Adaptive attack detection · Multivariate time series analysis · Hypothesis testing

## 6.1 Introduction

Over the past few years, the increased number of attacks on the power system has shown that the attackers are highly sophisticated and technologically advanced. The grid control systems use the Distributed Network Protocol (DNP3) for their communication which is also highly vulnerable (East et al. 2009; Darwish et al. 2015). Thus, the operators and the grid system should be capable enough to handle such attacks



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_6

and respond and recover faster. Fast detection gives sufficient time to mitigate the attacks' effects by isolating the system or implementing emergency control actions.

Different types of methods have been proposed in the literature for attack detection at various levels of the power system. Kalman filter (Akbarian et al. 2020) and Stochastic Unknown Input Estimators (Ameli et al. 2018) can be used to estimate LFC states using outputs and initial states. These estimates are further compared with measurements to detect attacks. Such model-based detection strategies depend on the accuracy of system models used for estimation. Bi et al. (2019a) suggests a game-theoretic approach to model the detection based on attack patterns. Attack-specific detection strategies suggest that analysis of specific attack strategies is necessary for detection, and Chen et al. (2018) presents a unified model. Bi et al. (2019b) discuss Fixed and Variable attacks and describe the variations in their impacts and detection. In (Wang and Govindarasu 2018), the authors derive conformity metrics that are used to detect abnormal generation controls induced by cyberattacks using a semi-supervised clustering. A set of attack templates are used to train the model. It utilizes raw data, and therefore, it is a data-driven algorithm.

Existing model-based algorithms depend on the model's accuracy, and changes in the system can affect these methods. Detection strategies built using specific attack patterns can fail to detect new or zero-day attacks. Thus, it is essential to devise new methods that are fast, adaptive, and independent of attack templates.

This chapter proposes a Spectral Analysis-based approach that utilizes the dynamic variations of signals during normal conditions to detect attacks effectively. An important attribute that is taken advantage of in the proposed work is that the grid control systems have somewhat regular dynamics that can be obtained using the massive amount of data that is available through Phasor Measurement Units (PMUs). Methods based on spectral analysis assume stationarity. The normal data are obtained from an underlying model which brings in the stationarity property to the normal data. By definition, attacks on the system are non-stationary, and thus any attacks into the system will lead to deviation from the normal characterization obtained. This deviation can be utilized to detect attacks.

The algorithm's significant advantage is its speed and low computational burden, making its practical implementation highly feasible in the Smart Grid environment.

Different types of methods have been proposed in the literature for attack detection at various levels of the power system. These can be broadly classified into model-based and data-based methods.

Table 6.1 summarizes the various features of existing attack detection algorithms.

The computational complexity is evaluated based on the most complex step in both the training and detection stages. Kalman filter and SUIE are model-based methods and depend on the total number of states and measurements. The data-based techniques depend on the number of measurements and training samples. To compare these different methods, we thus consider '$n$' as the number of states and measurements combined.

As seen from Table 6.1, both the data-based and model-based techniques have their own advantages and disadvantages. Thus, it is necessary to devise a detection method that is fast, accurate, and adaptive.

**Table 6.1** Attack detection techniques ($n$: combined number of states & measurements, $n_t$: Number of training data)

| No. | Algorithm | Complexity | | Advantages | Disadvantages |
|---|---|---|---|---|---|
| | | Estimation/learn | Detection | | |
| Model-Based | | | | | |
| 1 | Kalman Filter (Khalaf et al. 2019) | $O\left(n^3\right)$ | $O\left(n^2\right)$ | Does not use attack data. Detects zero-day attacks | Fails to detect noise level attacks. Not adaptive to system changes. Detection at the system level. Cannot be implemented at RTU/IED level |
| 2 | Stochastic Unknown Input Estimator (Ameli et al. 2018) | $O(n^3)$ | $O(n^2)$ | Fast detection with good accuracy for known attacks | Model is built based on attack data and hence cannot detect zero-day attacks. Not adaptive to system changes. Linear model for LFC and cannot be used at RTU/IED level |
| Data-Based | | | | | |
| 3 | Recurrent Neural Network (Ayad et al. 2018) | $O(n_t)$ | $O(n)$ | The model adapts to system changes and can detect attacks fast. | Uses labeled attack data during training. Therefore, fails to detect zero-day attacks. |
| 4 | GAN-Based (Li et al. 2021) | $O(n^2 n_t)$ | $O(n^2)$ | Fast and adapts to system changes. Semi-supervised can detect zero-day attacks | Fails to detect stealth attacks. Cannot be implemented at the RTU/IED level |
| 5 | Relation-Based dynamic analysis (Bi et al. 2019b) | $O(n^2 n_t)$ | $O(n^2)$ | Does not require attack model for training, and hence can detect zero-day attacks. Nonlinearities do not affect detection | High probability of false detection due to omission of modes in normal system dynamics |

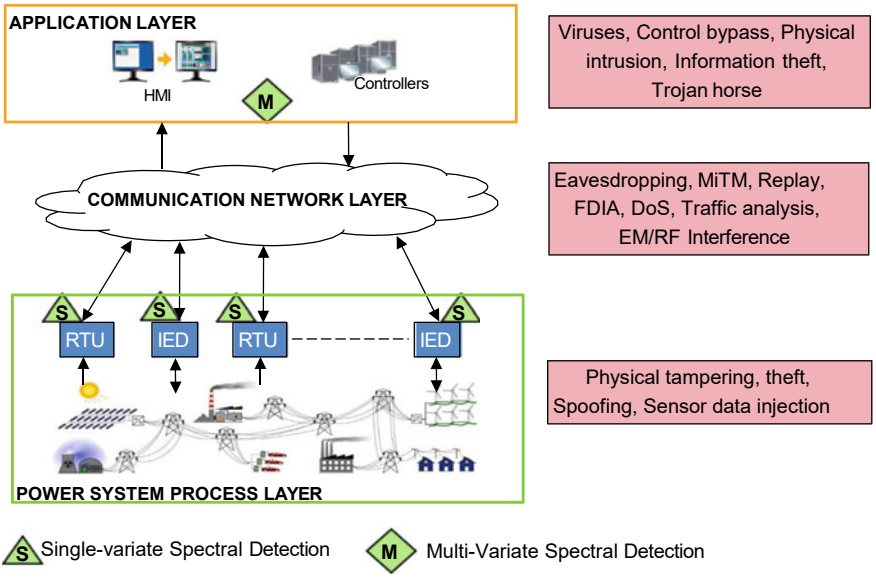



**Fig. 6.1** Multi-level attack detection

## 6.2 Multi-level Attack Detection

In the grid system, the lowest level of data acquisition is at the Remote Terminal Units (RTUs) or Intelligent Electronic Devices (IEDs). Thus, an attack detection at the lowest level of data acquisition can support local detection at a very high speed. If one device is tampered with, the changes will be immediately reflected within the other local RTUs. Detection accuracy can be further improved by utilizing various signals' relationships. Multiple signals would be available at the control center and certain RTUs and IEDs. In this multi-level detection, we propose a single variable detection at the RTU/IED level. A multivariate detection is proposed at the control center, where multiple signals are available. The overall architecture is as shown in Fig. 6.1

The measurements of the normal working conditions of a power system are easy to obtain. However, obtaining an exhaustive set of the attacked measurements is not feasible since the attacking pattern of attackers keeps changing and cannot be pre-determined. The proposed method's significant advantage is that only routine condition measurements are necessary for the training phase. We first discuss Singular Spectrum Analysis (SSA) preliminaries and then discuss the proposed detection algorithm for a single variable case. We finally extend the proposed algorithm to a multivariate detection algorithm.

## 6.3 Singular Spectral Analysis (SSA)-Based Attack Detection

This section proposes a cyber-attack detection framework assuming measurement and data to be noisy. The proposed method utilizes the representative behaviors in the system's dynamical changes.

The Singular Spectrum Analysis (SSA) is a data-based time-series method primarily used for spectral estimation. It decomposes the time series into a sum of components, each having a meaningful interpretation. Its advantage is that it does not require a defined known model and takes no a priori statistical assumptions on the signal (Hossein Hassani 2018). Due to these benefits, SSA is used in a wide range of applications. SSA primarily involves two stages to create noise-free data: decomposition and reconstruction. We use the reconstructed data in the proposed method to represent the system's normal behavior (Malioutov et al. 2005). Singular value decomposition (SVD) obtains the dominant eigenvectors corresponding to the signal subspace. SVD thus aids in separating noise from data and performs dimensionality reduction to improve the real-time computation speed.

The algorithm can be broken down into the following steps:

1. **Singular Spectrum Analysis (SSA)**: Used to break down the measurement data into components that are representative of the normal behaviors using SVD. Based on these components, a projection matrix is obtained.
2. **Normal Data Cluster Analysis**: Training data are projected onto signal subspace based on the projection matrix. This can be considered as a low-rank approximation of the signal and helps in eliminating components of the noise that are in the noise subspace. These projected data form a cluster in the signal subspace characterized by a center.
3. **Detection**: New measurements are projected to signal subspace, and their distance from the cluster's center is determined. If the data is far from the cluster, it indicates an attack.

## 6.4 Process Level Single Variate Attack Detection

Figure 6.2 gives the overall steps involved in the attack detection process for both single and multivariate approaches. The different steps involved in attack detection are discussed below.

### *6.4.1 Signal Subspace Determination*

Since the measurements are recorded and transmitted over Power line carrier communication channels, they usually contain noise. Thus, the proposed method's first

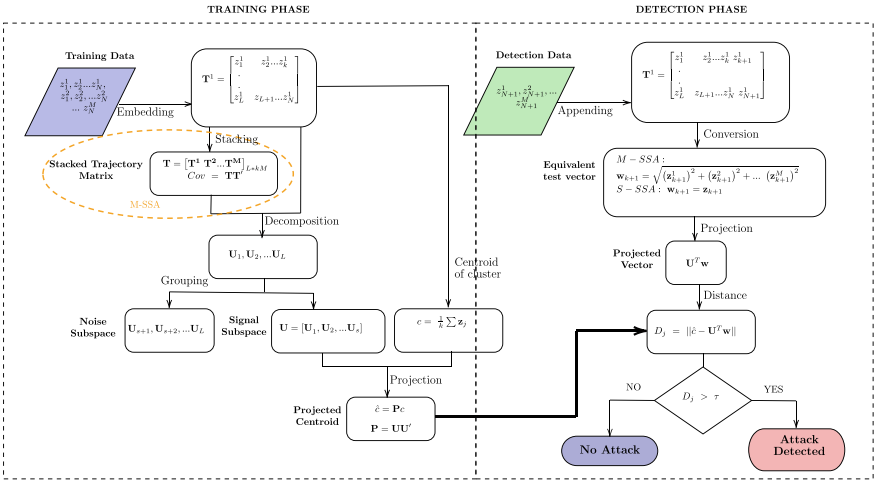



**Fig. 6.2** Spectrum analysis-based detection flowchart

step is determining the Signal subspace from the incoming measurements to separate noise components.

We collect $N$ samples of each sensor data and embed it in an $L$-dimensional Euclidean space $\mathbb{R}^L$ to form the trajectory matrix $\mathbf{T} \in \mathbb{R}_{L\times k}$ ($L \leq N/2, k = N - L + 1$). The trajectory matrix of each sensor $i$ is given by

$$\mathbf{T}^{(i)} = \mathbf{T} = \begin{bmatrix} z_1^{(i)} & z_2^{(i)} & \cdots & z_k^{(i)} \\ z_2^{(i)} & z_3^{(i)} & \cdots & z_{k+1}^{(i)} \\ . & . & . & . \\ . & . & . & . \\ z_L^{(i)} & z_{L+1}^{(i)} & \cdots & z_N^{(i)} \end{bmatrix} \tag{6.1}$$

This matrix inherits a Hankel structure.

We then use Singular Value Decomposition(SVD) to obtain the eigenvectors $U_1, U_2, ...U_L$ of the covariance matrix $\text{Cov} = \mathbf{TT}'$.

Since the dominant eigenvalues correspond to the system dynamics, the next step is to decompose the column space into dominant $(U_1, U_2, \ldots, U_s)$ and non-dominant $(U_{s+1}, \ldots, U_L)$ subspaces. The dominant subspace is referred to as the signal subspace, while the non-dominant subspace corresponds to noise. We choose SVD for this owing to its computational robustness, and high-resolution discrimination against noise contamination (Klema and Laub 1980). Thus, any vector in the signal subspace will be a linear combination of $U_1, U_2, \ldots, U_s$. If $S^s$ is the signal subspace,

$$\begin{aligned} S^s &= \text{Span}(U_1, U_2, \ldots U_s) \\ \mathbf{U} &= [U_1 U_2 ... U s] \end{aligned} \tag{6.2}$$

### 6.4.2 Signal Subspace Projection

The next step is to project each of the vectors $\mathbf{z}_j = \left[z_1^{(i)}\ z_2^{(i)}\ \ldots\ z_L^{(i)}\right]'$ (i.e., each column of matrix $\mathbf{T}$) of the trajectory matrix onto the signal subspace to get a mathematical representation of the normal process behavior. For this, we find a $\mathbf{p} \in S^s$ that is nearest to $\mathbf{z}$,i.e., $||\mathbf{p} - \mathbf{z}||$ is minimum. This is equivalent to finding an orthogonal projection matrix $\mathbf{P}$ which projects $\mathbf{z}$ onto $S^s$

$$\mathbf{P} = \mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}' \tag{6.3}$$

Since columns of U are orthonormal, $\mathbf{U}'\mathbf{U} = \mathbf{I}$.

$$\mathbf{P} = \mathbf{U}\mathbf{U}' \tag{6.4}$$

Since the signal subspace $S^s$ represents the system's normal behavior, all the training vectors form a cluster. The centroid of the cluster is given by

$$c = \frac{1}{k}\sum_{j=1}^{k}\mathbf{z}_j \tag{6.5}$$

where $\mathbf{z}_j$ is each of the column vectors of $\mathbf{T}$ This centroid is represented in the signal subspace as

$$\tilde{c} = \mathbf{P}c \tag{6.6}$$

If an attack is injected into the system measurements, the system dynamics changes and the attack pushes the incoming measurements away from the cluster of normal behavior vectors.

### 6.4.3 Detection Phase

Every incoming measurement sample is added to the trajectory matrix to form a new lagged vector of measurements in the detection phase. If $z_{N+1}$ is the most recent incoming sample, the lagged vector is given by

$$\mathbf{z}_j = [z_{k+1}^{(i)} \ldots z_N^{(i)}\ z_{N+1}^{(i)}]' \tag{6.7}$$

The final step is to calculate the squared Euclidean distance of each new lagged vector $\mathbf{z}_j$ from the centroid in $S^s$

$$D_j = ||\tilde{c} - \mathbf{P}\mathbf{z}_j|| \tag{6.8}$$

An attack is detected if this distance exceeds a certain threshold ($D_n > \tau$). The calculation of this threshold value will be discussed in detail in this chapter.

### 6.4.4 Selection of Parameters

The two major parameters that need to be selected in the algorithm are the window length or the number of rows in the training phase, $L$, and the value of the number of eigenvalues to determine signal subspace $s$. The value of the window length $L$ depends on the total number of training samples and is usually taken as $L \leq N/2$. There are several methods for choosing the value of $s$ to separate signal and noise effectively. In the given simulation, we use the concept that pure noise series typically produces a slowly decreasing sequence of singular values. This can be estimated well from a screen plot which is a plot of the logarithm of the singular values.

## 6.5 Multivariate SSA for Control Center Level Detection

SSA is primarily designed for single time-series data. However, the single variable SSA can be effectively extended to a multivariate method. All the sensor measurements and actuator signals are available at the control center. These measurements have certain relations which define the dynamical system better. Thus, the single variable SSA can be extended to include multiple measurements and get a better characterization for the LFC process (von Bünau et al. 2009; Hossein Hassani 2018).

The extension of the proposed method to multivariate cases mainly involves changes in two steps, as shown by the orange dashed area in Fig. 6.2, i.e., in the formation of the matrix $\mathbf{T}$ and in forming the equivalent test vector during detection.

### 6.5.1 Extension in Training Phase

To include multiple measurement values, we form a stacked matrix using the various $\mathbf{T}^{(i)}$ matrices of step 1. The stacking can be horizontal or vertical. It will be shown in the subsequent sections that the computational burden depends on the number of rows. Thus, Horizontal stacking has the advantage that the number of rows (and the computation burden) remains constant irrespective of the number or measurements considered. The stacked trajectory matrix for $M$ sensors is

$$\mathbf{T}_{L\times(KM)} = [\mathbf{T}^{(1)}_{L\times K}\ \mathbf{T}^{(2)}_{L\times K} \ldots \mathbf{T}^{(M)}_{L\times K}] \tag{6.9}$$

### 6.5.2 Extension in Detection Phase

The major challenge in the multivariate algorithm during the detection phase is to find a new vector that should be the equivalent of all the lagged measurement vectors. In order to obtain the equivalent vector with the new sample, we analyze the covariance matrix Cov. When a new sample is included in the trajectory matrix, the matrix $\mathbf{T}$ is appended as

$$\mathbf{T}_{new}^{(i)} = \begin{bmatrix} z_1^{(i)} & z_2^{(i)} & \cdots & z_k^{(i)} & z_{k+1}^{(i)} \\ z_2^{(i)} & z_3^{(i)} & \cdots & z_{k+1}^{(i)} & z_{k+2} \\ . & . & . & . & \\ . & . & . & . & \\ z_L^{(i)} & z_{L+1}^{(i)} & \cdots & z_N^{(i)} & z_{N+1}^{(i)} \end{bmatrix} = [\mathbf{T}^{(i)} : \mathbf{z}_{k+1}^{(i)}] \tag{6.10}$$

The stacked trajectory matrix can now be represented as

$$\mathbf{T} = [\mathbf{T}^{(1)} : \mathbf{z}_{k+1}^{(1)} : \mathbf{T}^{(2)} : \mathbf{z}_{k+1}^{(2)} : \ \ldots \mathbf{T}^{(M)} : \mathbf{z}_{k+1}^{(M)}] \tag{6.11}$$

We then construct the covariance matrix as

$$\begin{aligned} Cov &= \mathbf{T}\mathbf{T}' \\ &= \mathbf{T}^{(1)}\mathbf{T}^{(1)'} + \mathbf{z}_{k+1}^{(1)}\mathbf{z}_{k+1}^{(1)'} \\ &\quad + \mathbf{T}^{(2)}\mathbf{T}^{(2)'} + \mathbf{z}_{k+1}^{(2)}\mathbf{z}_{k+1}^{(2)'} \\ &\quad + \ldots + \mathbf{T}^{(M)}\mathbf{T}^{(M)'} + \mathbf{z}_{k+1}^{(M)}\mathbf{z}_{k+1}^{(M)'} \end{aligned} \tag{6.12}$$

Thus, the vector $\mathbf{w}$ equivalent to a combination of all measurements is given by

$$\mathbf{w}_{k+1}\mathbf{w}'_{k+1} = \mathbf{z}_{k+1}^{(1)}\mathbf{z}_{k+1}^{(1)'} + \mathbf{z}_{k+1}^{(2)}\mathbf{z}_{k+1}^{(2)'} + \ldots + \mathbf{z}_{k+1}^{(M)}\mathbf{z}_{k+1}^{(M)'} \tag{6.13}$$

$$\mathbf{w} = \sqrt{(\mathbf{z}^{(1)})^2 + (\mathbf{z}^{(2)})^2 \ldots (\mathbf{z}^{(M)})^2} \tag{6.14}$$

Thus, every incoming sample of data is used to create a new time-lagged vector $\mathbf{z}_{k+1}$ and $\mathbf{w}$, as obtained in (6.14) is used to create the test vector.

## 6.6 Performance Analysis of Detection Algorithm

Since the LFC is a critical control system operating in real time, it is essential to consider the time and computational complexity of the attack detection algorithm. The computation burden should be low so that the algorithm can be implemented using limited hardware and it does not significantly impact the overall performance

**Table 6.2** Computational complexities for single variate and multivariate detection

| Computation step | | Single variate | Multivariate |
|---|---|---|---|
| Training phase | Covariance matrix formation | $O(L^2k)$ | $O(L^2kM)$ |
| | SVD | $O(L^3)$ | $O(L^3)$ |
| | Total | $O(L^2k + L^3)$ | $O(L^2kM + L^3)$ |
| Detection phase | Equivalent vector computation | – | $O(M)$ |
| | Distance $(\tilde{c} - \mathbf{P}\mathbf{w}_j)$ | $O(L^2s)$ | $O(L^2s)$ |
| | Norm | $O(s)$ | $O(s)$ |
| | Total | $O(L^2)$ | $O(L^2 + M)$ |
| | Enhanced | $O(L)$ | $O(L)$ |

of the LFC system. The computational complexity of each step for both the single variate and multivariate algorithms is given in Table 6.2.

During the training phase, singular value decomposition is the heaviest computation step. The covariance calculation of the trajectory matrix has a computational time proportional to $O(L^2k)$ for the single variate algorithm. It depends further on the number of sensors in multivariate case $O(L^2kM)$. For SVD, it is $O(L^3)$ for both the cases since the Covariance matrix has the same size $(L \times L)$. Thus, the time complexity of the training phase is approximate $O(L^2k + L^3)$. However, this value does not significantly affect the performance as it is not a real-time operation.

The detection phase computational burden is the time required to calculate the distance, $D_j = ||\tilde{c} - \mathbf{P}z_j||$. Since $\mathbf{P} = \mathbf{UU'}$, the computational complexity of $(\tilde{c} - \mathbf{P}z_j)$ is $O(L^2s)$ and that for norm calculation is $O(s)$ which leads to an overall complexity of $O(L^2s + s)$. Since $s \ll L$, the computation complexity for the detection phase is approximately quadratic in $L$ or $O(L^2)$.

In the multivariate case, the computational burden further involves the time required to compute the aggregate vector $\mathbf{w}_j$. From (6.14), each element of $\mathbf{w}$ is

$$w_i = \sqrt{(z_i^{(1)})^2 + (z_i^{(2)})^2 ... (z_i^{(M)})^2} \tag{6.15}$$

Since other values are available from previous time steps, only the $w_i$ corresponding to the latest incoming sample needs to be calculated, giving a complexity of $O(M)$ at this stage. Since $M \ll L$, the net computation complexity for the detection phase, $O(L^2 + M)$, is approximately the same as in the single variate case.

The computational complexity can be further reduced using the analysis that follows.

### 6.6.1 Performance Enhancement of Detection Algorithm

The performance of the algorithm can be further improved using the theorem below.

**Theorem 6.1** *Let $U'$ be a linear transformation from $\mathbb{R}^L \longrightarrow \mathbb{R}^r$. Then, the norm of the projection of any arbitrary vector $z \in \mathbb{R}^L$ onto the subspace $S^r$ is equivalent to the norm of the transformed vector $\mathbf{U}'z \in \mathbb{R}^r$.*

The proof for the above theorem is given in Appendix D.

Thus, to find the distance $D_j$, it is sufficient to find the $||\mathbf{U}'(c - \mathbf{w}_j)||$ without the explicit projection of $\mathbf{P}(c - \mathbf{w}_j)$. Thus, the computation complexity in the detection phase is finally reduced to linear in $L$ or $O(L)$.

**Remark**: The computational complexity of MSSA detection at the control center may be further improved by using parallel computing-based high-performance computing devices.

## 6.7 Multi-level Attack Detection Results

### 6.7.1 Example 6.1: Multi-level Attack Detection on 39-Bus System

The detection algorithm is implemented on the IEEE 39-bus 3 area New England Test system. The attack surface consists of the tie-line powers and the frequency. The system and the attack surface are shown in Fig. 6.3. The MSSA detection in figure is at the control center. For process level detection, the detection module will be at the sensor points as shown in Fig. 6.1.

The unified multi-area frequency control attack model and the detection algorithm were carried out in MATLAB. The steps followed to obtain results for attack detection are as follows:

**Step 1**: The simulation uses load forecast data from the New England ISO website for the training phase. The actual load data are then used for the testing phase.

**Step 2**: Attack simulation is conducted by injecting the different attacks mentioned in the previous section at a time of $t_{att}$. Attack values are selected such that frequency remains within the prescribed limits to maintain the stealthiness of the attack. The Frequency control simulation is performed using the load data from the New England ISO website, adding white gaussian noise, and the attacks to obtain the study dataset. The noise has a variance of $10^{-8}$ for frequency and $10^{-6}$ for power measurements.

**Step 3**: In the detection phase, we first determine the threshold using the data till a time $t_{th}$ such that $0 < t_{th} < t_{att}$ (The method for determination of an adaptive threshold will be discussed in Chapter 6).

**Step 4**: The algorithm raises the alarm if the distance $D_j$ goes beyond the threshold for any incoming measurement.

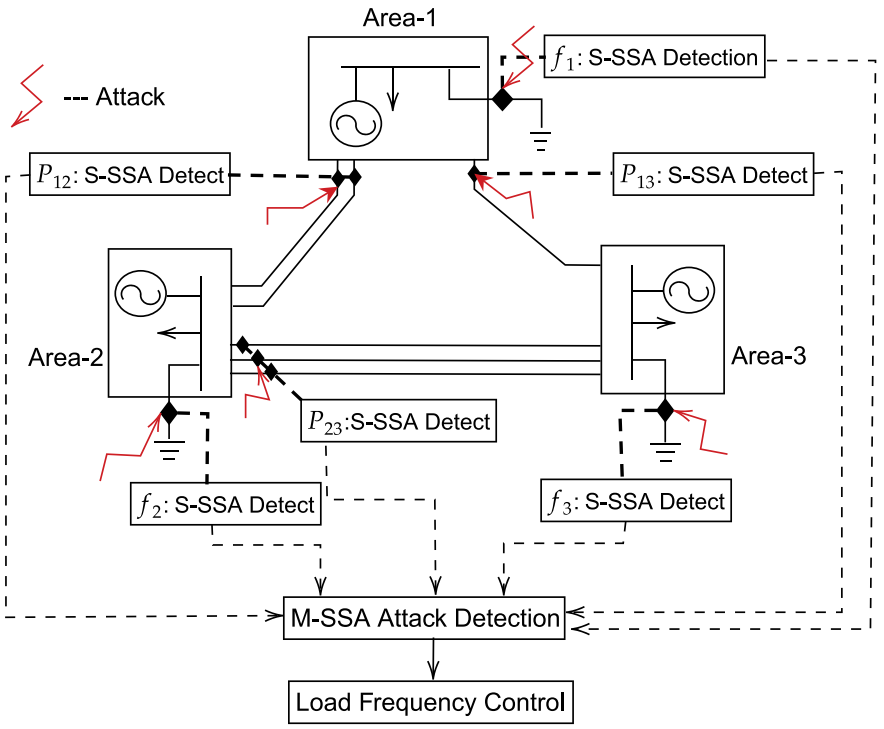



**Fig. 6.3** Attack surface and detection for IEEE 39-bus system LFC

**Step 5**: False positive and false negative values are finally calculated to obtain the detection accuracy and to compare the algorithm with existing algorithms.

The time of detection is defined as the difference between the sample at which the attack begins and the sample at which the distance value crosses the threshold.

The accuracy can be calculated using false positive and false negative values. These parameters are determined as follows:

1. False Positives(FP): number of attacked measurements being detected as normal ones
   a. False Positive rate(FPR)=$\frac{FP}{N_s}$, where $N_s$=Total samples

2. False Negatives(FN): number of normal measurements being detected as attacked.
   a. False Negative rate(FNR): $\frac{FN}{N_s}$

3. True Positives(TP) and True negatives(TN): Number of attacked measurements (TP) and normal measurements(TN) being detected correctly.

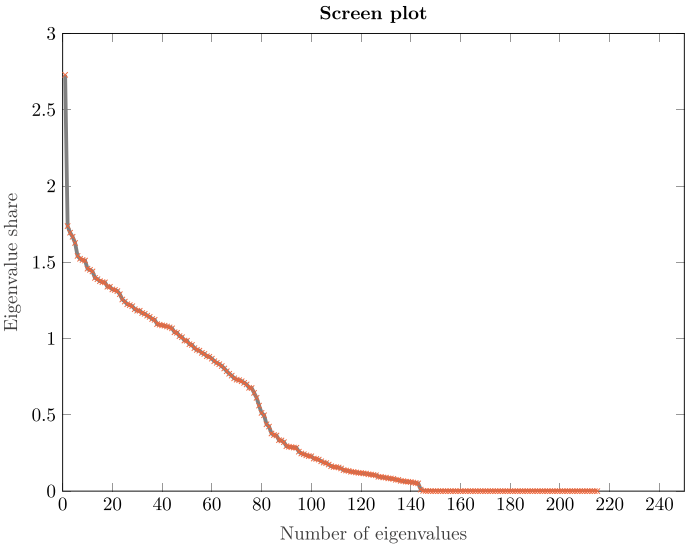



**Fig. 6.4** Screen plot

4. Accuracy of detection

$$\text{Accuracy} = \frac{TP + TN}{N_s} \tag{6.16}$$

### *6.7.2 RTU/IED Level Detection Results*

We consider an attack compromising the frequency and tie-line sensors of balancing area-3. The screen plot is first used to choose the value of '$s$' which is obtained as shown in Fig. 6.4.

It can be observed that a significant drop in the eigen value occurs around component 75, which could be interpreted as the start of the noise floor. Therefore, we choose a value of 75 for $s$.

Figures 6.5 and 6.6 depict the detection of stealthy attacks at the frequency and tie-line sensors of area-2. The top plot shows the actual sensor value and the bottom plot shows the distance variation. The horizontal line indicates the detection threshold. In the current results, the threshold is taken as the maximum value of distance. An adaptive threshold strategy will be proposed in Sect. 6.10.

The attacks are detected in the frequency sensor within 7 samples and in the tie-line sensor within 9 samples as shown in Table 6.3. Since the sampling is at 1 s interval, the detection time translates to 7–9 s. Therefore, the detection algorithm effectively works against different test scenarios with an acceptable time for detec-

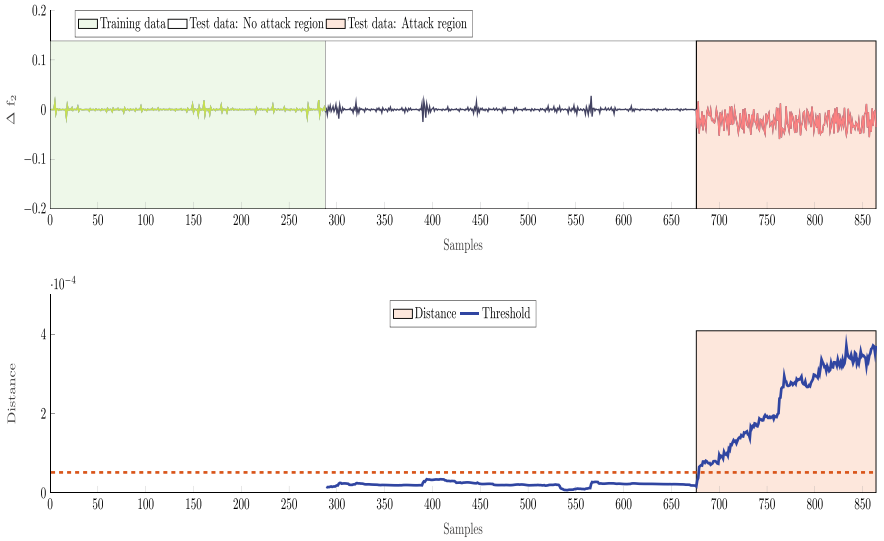



**Fig. 6.5** Attack detection at the frequency sensor with N = 288, L = 10, s = 13

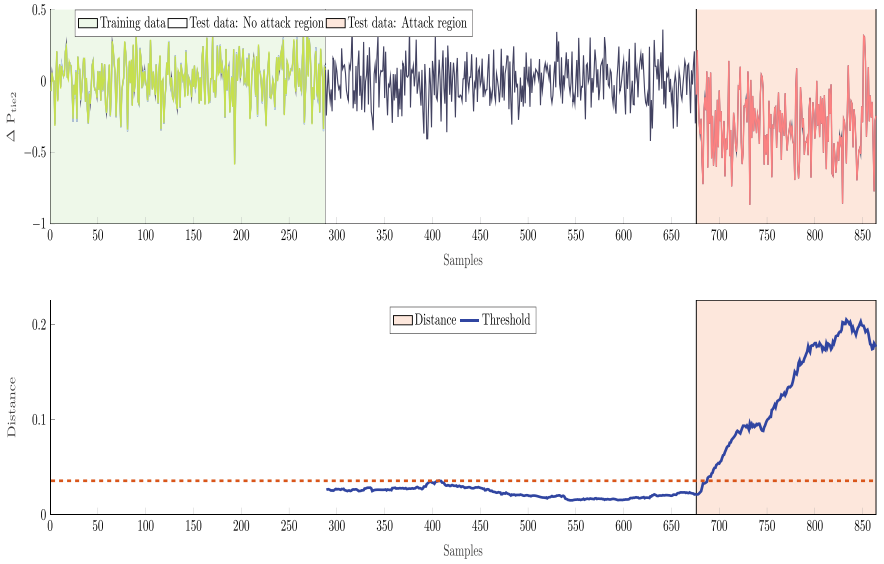
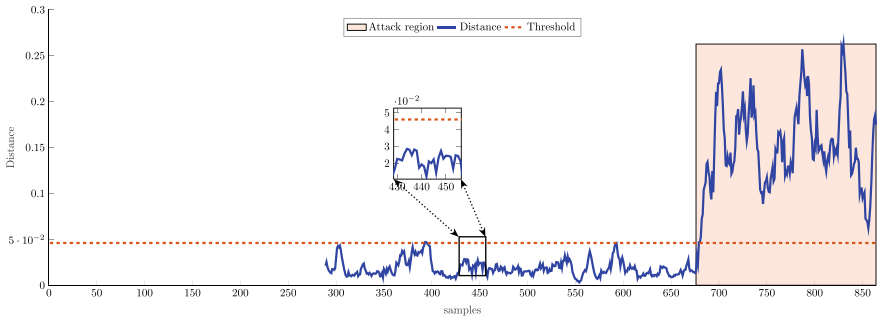



**Fig. 6.6** Attack detection at the tie-line sensor with N = 288, L = 10, s = 13


(a) $\Delta f$, $\Delta P_{tie}$, ACE Signal Dynamics




(b) Distance and threshold

**Fig. 6.7** Multivariate stealth attack detection with N = 288, L = 10, s = 10

tion. Additionally, attacks in one area can be detected by sensors belonging to other areas. This is because the dynamics of the entire system is connected.

### *6.7.3 Control Center Level Detection Results*

Figure 6.7 shows the detection at the control center using the multivariate algorithm with a constant threshold (As shown in the zoomed-in graph, Fig. 6.7b).

It can be seen that the plots look similar to that of the single variate algorithm. However, when we analyze the time of detection and the accuracy levels, it can be observed that the multivariate counterpart detects the attacks with better accuracy and in less time as indicated in Table 6.3. This improved accuracy is because the multivariate algorithm exploits the relation between different signals in addition to the dynamic variations existing in a single signal.

Table 6.3 compares a single variate algorithm and different combinations used for the multivariate analysis of an attack on the tie-line sensor of balancing area-1. For a better comparison, the values of N and L used in both S-SSA and M-SSA are

**Table 6.3** Effect of signal selection on accuracy and time of detection

| Algorithm | Signals used | Accuracy (%) | Time of detection (samples) |
|---|---|---|---|
| S-SSA | Frequency | 97.57 | 7 |
| | Tie-line | 98.78 | 9 |
| | ACE | 99.48 | 6 |
| M-SSA | Frequency, Tie-line, ACE | 99.85 | 2 |
| | Frequency, Tie-line | 99.48 | 3 |
| | Frequency, ACE | 99.48 | 4 |
| | Tie-line, ACE | 99.83 | 3 |

kept constant as N = 288 and L = 10. The accuracy is calculated using (6.16) for the complete range of detection, i.e., with $N_s = 576$.

The S-SSA and M-SSA are implemented at different levels of the grid. For the S-SSA, it can be observed that an attack on any sensor measurements is also reflected in other sensors with a decent detection time. For M-SSA, it can be observed from the Table 6.3 that the accuracy and the detection time are best when a combination of frequency, tie-line, and ACE is used. This better accuracy is because the ACE is a combination of frequency and tie-line, and thus the relation between them is robust. Any deviations will thus be immediately detected. If a grid control system has several measurements, these changes will become more evident using various combinations.

## 6.8 Hypothesis Testing-Based Attack Detection

The conclusions presented in the above sections are presented asymptotically. Based on a measure of distance from the centroid, they define the attacks that can be detected. These asymptotic descriptions may be further developed to create statistical tests that detect malicious activity with an acceptable false alarm rate in a finite amount of time. Various tests may be applied to describe the problem at hand formally. Mehra and Peschon (1971), use whiteness, mean, and covariance in the data sequence to detect control system faults. The sequential probability ratio test devised by Wald (1945) is one of the most popular and widely used tests. This test selects either the null or alternate hypothesis or continues testing based on comparison with specific threshold values.

The conventional statistical hypothesis testing methods can fail in the attack detection method proposed in this thesis for the following reasons:

1. It is not possible to define a distribution from which the observations would arise (the system is under attack). Therefore, it is impossible to define a likelihood

ratio, which means that the conventional sequential probability ratio test cannot be applied.
2. The sequential probability ratio test should be applied continuously, and the threshold established should constantly adapt to the system changes.
3. No two exact distributions differentiate between the null and alternate hypotheses. The problem is to reject the null hypothesis when there is an attack.

Based on the above observations, we define a new hypothesis test to define our problem and determine the parameters of the detection test based on the proposed hypothesis testing framework.

In this chapter, we propose a formal description for the above detection as a hypothesis testing framework and propose a method for threshold determination based on Information theory.

## 6.9 SSA Hoeffding Test-Based Hypothesis Testing

We will treat the problem of deciding whether a new set of measurements $\mathbf{w}_i$ denotes an attack or are normal. We treat this as the composite hypothesis testing problem between the hypothesis $\mathcal{H}_0$ and the complement of $\mathcal{H}_0$ denoted by $\mathcal{H}_1$. We call this test composite since the exact probability distribution of the measurements is unknown.

The space of equivalent measurements $\mathcal{W}= \{w_l; l = 1, 2, ...\}$ is a Hilbert space since the $w_i$ represents a Euclidean norm of measurements $z_i$ derived from a generative process. We use the principles of Quantum hypothesis testing (Nagaoka and Hayashi 2007) to define the test and the threshold parameter. The distance $D_l$ is a measure of distance between a set $\mathcal{S}_n$ generated by a normal system operation and a set $\mathcal{S}_a$ generated by an operation under attack. Thus $D_l$ is equivalent to the relative entropy $H(\upsilon|\mu)$, which is the divergence between two probability laws $\upsilon$ and $\mu$ (Jaksic et al. 2012). We propose a composite hypothesis test based on the Hoeffding inequality (Amir and Ofer 2010).

**Definition 6.1** (*SSA Hoeffding Test*) The SSA Hoeffding Test is a hypothesis test that rejects the hypothesis $\mathcal{H}_0$ when $w_l \in \mathcal{S}_{SHT}$ where

$$\mathcal{S}_{SHT} = \{w_l | D_l \geq \tau, \forall\, l = 1, 2, ..., D_l = ||\tilde{c} - \mathbf{Pw}||\} \tag{6.17}$$

It can be shown that the above hypothesis test satisfies the Neyman-Pearson lemma, i.e., it is the test with maximum power or minimum detection error.

**Theorem 6.2** *The SSA Hoeffding Test satisfies the Neyman-Pearson lemma*

**Definition 6.2** (*Neyman-Pearson Lemma*) Consider a binary hypothesis test and the distance measure:

$$d(x) = ||c - Px||_2 \underset{H_0}{\overset{H_1}{\gtrless}} \tau \tag{6.18}$$

with a probability of false alarm given by

$$P_{FA} = P(d(x) \geq \tau | H_0) = \beta \tag{6.19}$$

There does not exist another test with $P_{FA} = \beta$ and a detection problem larger than $P(d(x) \geq \tau | Ho)$. That is, the SSA-HT is the most powerful test with $PF_A = \beta$.

The proof for Theorem 6.1 is given in Appendix D.

## 6.10 Adaptive Threshold Selection

The false Positive rate is used to select a suitable threshold $\tau$. The false positive rate is the probability that an attack is declared as detected when there is actually none. The threshold $\tau$ can be tuned such that the SSA Hoeffding Test has a high detection rate and a low false positive rate. The theoretical false positive rate is given by

$$\beta = \mathcal{P}_{\mathcal{H}_0}[D_l \geq \tau] \tag{6.20}$$

Since the number of training data is large enough and the probability $\beta$ corresponds to that of a rare event probability, we can use the large deviation principles and Sanov theorem (Amir and Ofer 2010) to approximate the threshold. Large deviation principles provide asymptotic estimates for rare events' probabilities. Sanov theory can be used to determine the minimum value of $\tau$ that can bring the false positive rate below $\beta$.

For a given false positive rate $\beta$, an optimal threshold for the SSA Hoeffding test is obtained using the Sanov theorem as given in (6.21)

$$\tau \geq \frac{-1}{N} \log \beta \tag{6.21}$$

As new data come in and are classified as normal, the threshold can be updated using (6.21). Thus the detection process becomes adaptive.

Equation (6.21) can be used to determine the threshold until a certain finite number of data points ($N$). As $N$ keeps increasing, the threshold $\tau$ keeps decreasing, i.e., as $N \longrightarrow \alpha$, $\tau \longrightarrow 0$, which is not a realistic assumption. Thus, we can further use the large deviation principles and empirical Cumulative Distribution Function (eCDF) to approximate the threshold. eCDF can be used to derive a $\tau$ that can bring the false positive rate below $\beta$.

**Definition 6.3** (*Empirical Cumulative Distribution Function (eCDF)*) Let $(X_1, \ldots, X_n)$ be independent, identically distributed real random variables with the common cumulative distribution function $F(t)$. Then the empirical distribution function is defined as

$$\hat{F}(\tau) = \frac{\text{Number of elements in sample} \leq \tau}{n} \tag{6.22}$$

Thus, if the eCDF of distance is known, we can estimate the $\tau$ using inverse eCDF as in Algorithm 6.1.

**Algorithm 6.1:** Threshold estimation for the SSA Hoeffding test using Empirical Cumulative Distribution Function (eCDF)

**Input:** Sample size ($n$), Target False positive rate ($\beta$), Distance ($D_l, \forall l = 1, 2, ..n$)
1. Choose the first $n$ samples of the test phase.
2. Based on the $n$ samples obtained in Step 1, estimate an empirical CDF of $D_l$, denoted $\hat{F}_{emp}(.; n)$
3. Obtain an estimated value for $\tau$ using the inverse eCDF, $\hat{F}^{1}_{\text{emp}}(.; n)$ and (6.23)

$$\tau = F^{-1}_{emp}(1 - \beta; n) \tag{6.23}$$

4. Use the $\tau$ in previous step to detect attacks for next $n$ samples.
5. If next $n$ samples are also normal measurements, include them and determine the new $\tau$
6. Repeat steps 1 to 5 until attack is detected.

## 6.11 Adaptive Attack Detection Results

### *6.11.1 Example 6.2: Adaptive Attack Detection*

The detection is first analyzed on the 39-bus 3 area test system as shown in Fig. 6.8. The complete data for the system are given in Appendix A. The attack is on the frequency and tie-line sensors and the detection is carried out at the control center.

We use MATLAB on a Core i5 processor system to implement the unified multi-area frequency control attack model and the MSSA detection algorithm. The steps followed to obtain results for attack detection are similar to those given in this chapter. The threshold, however, is determined adaptively. The steps followed to obtain results for attack detection are as follows:

**Step 1**: Use load forecast data from New England ISO in the LFC to generate $\Delta f$, $\Delta P_{tie}$, and ACE training data.

**Step 2**: Project the data and determine the centroid using the multivariate SSA-based detection and (6.6).

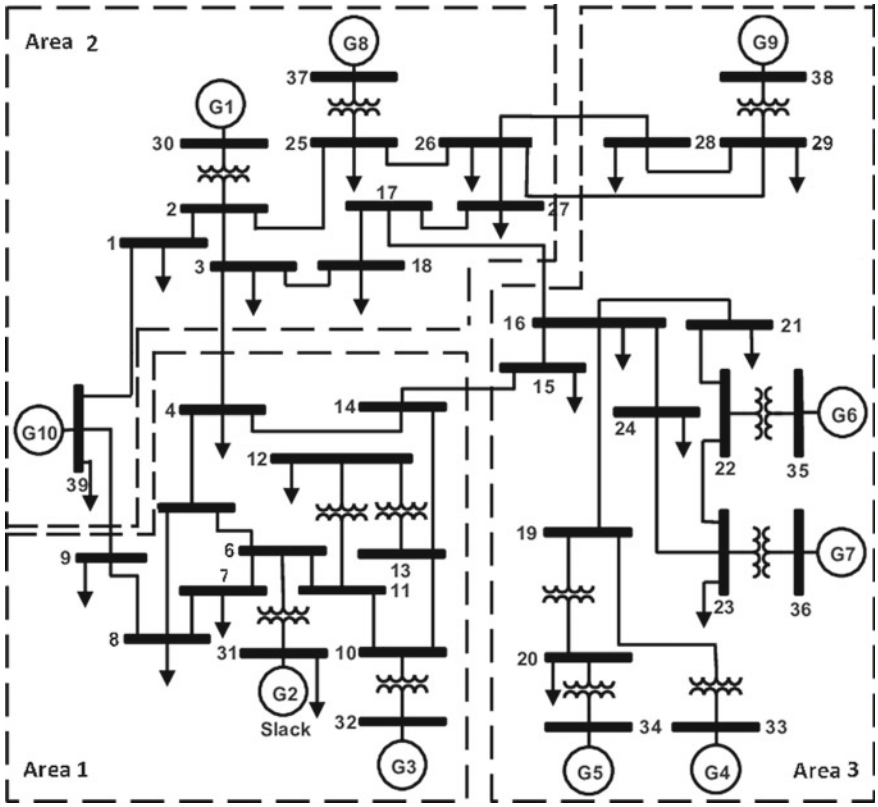



**Fig. 6.8** IEEE 39-bus 3 area new England test system

**Step 3**: Inject attacks and noise into the system at time $t_{att}$ to generate the test data.

**Step 4**: In the detection phase, determine the threshold using the data till time $t_{th}$ such that $0 < t_{th} < t_{att}$.

**Step 5**: The algorithm raises the alarm if the distance $D_j$ goes beyond the threshold for any incoming measurement.

**Step 6**: Use Algorithm 6.1 for each window to change the threshold adaptively. A window of 150 samples is chosen to change the threshold adaptively.

The window for the threshold change is currently selected randomly. In the future, the window size selection can also be made optimal using learning algorithms.

In the detection phase, we first determine the threshold using Algorithm 6.1 with a significance value or false positive rate of $10^{-4}$. An alarm is triggered if the distance $D_j$ goes beyond the threshold for any measurement and sustains there.

Figure 6.9 shows the adaptive change in threshold during attack detection. The first window for determining $\tau$ considers the samples from 290 to 440. Once the next dataset is classified as normal, the threshold value is modified by including

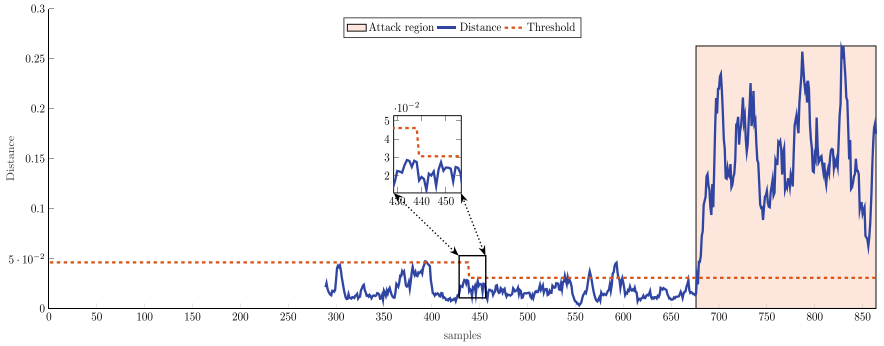



**Fig. 6.9** Adaptive threshold detection

the following dataset. The threshold keeps changing as more data are classified as normal.

### *6.11.2 Performance Under Load Variations*

In this section, we show that the proposed method works well even if there are sudden load changes compared with forecast loads or faults in the system. These are normal operating conditions of the system that can occur even in the absence of an attack. Any detection scheme should be able to discriminate between these and attack conditions.

To analyze the effect of high load variations, let us consider that the load in balancing area-1 undergoes a load shedding of 500 MW and is restored after 8 h. Figure 6.10 shows the actual and forecast load in area-1. Thus, there is a considerable difference between the forecast load used for training and the actual load during detection.

The system conditions in Fig. 6.11 are similar to Fig. 6.7; except that, in balancing area-1, there is a load drop.

It can be seen that the distance calculated during the attack state is much higher than that in the load shedding state. This distance value proves that the proposed scheme works effectively even if the forecast and actual load values are different. The detection performance is because the changes in load contribute to normal dynamic variations in the system, which are represented by the subspace $S^s$. Additionally, the above results suggest that MSSA-based detection cannot detect malicious direct tripping of loads.

In the presence of sensor and actuator faults, the control system may make wrong decisions leading to system instabilities. Thus, the proposed algorithm considers such faults as attacks, so appropriate mitigation steps may be taken. The difference between the estimated and actual value historical information is used to identify

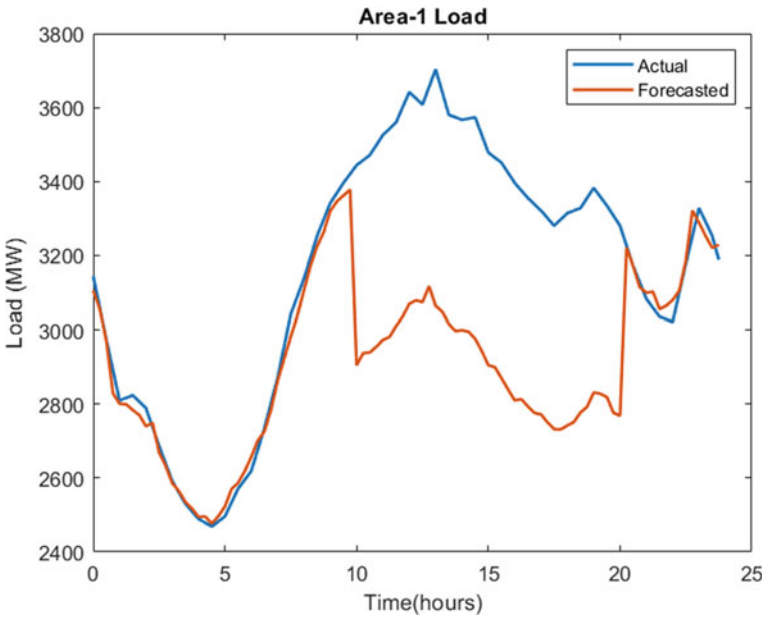



**Fig. 6.10** Actual and Forecast Load In Balancing Area-1

meter offset in the control center. Also, the proposed approach and State Estimation can be used to identify sensor faults.

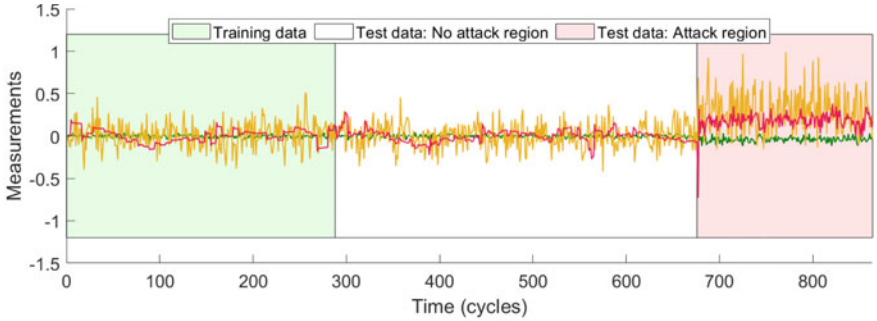
### *6.11.3 Comparison with Existing Detection Strategies*

Some of the advantages of the proposed method have been discussed in Chap. 1. In this section, we compare the performance of the proposed detection technique with three different types of existing techniques concerning the 3 area system above. We compare the proposed detection with a model-based, data-based, and a machine learning- based algorithm to show its superior performance when compared to various types of methods. The different techniques compared are as follows:

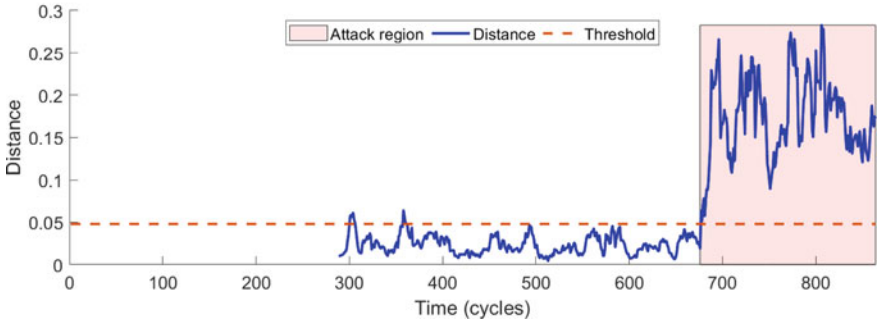
1. Kalman Filter: Model-Based technique (Khalaf et al. 2019)
2. Dynamic Characteristics Analysis: Data-Based technique (Bi et al. 2019b)
3. One Class Support Vector Machine (OC-SVM): Machine Learning-Based technique (Demetriou et al. 2017)

Table 6.4 gives a comparison between these methods based on FPR, FNR, Accuracy of detection, and average cycles of LFC needed for detection.

The Kalman Filter-based detection cannot detect attacks in the noise region, due to which the false negative rate is 100%. The dynamic analysis-based method and




(a) $\Delta f$, $\Delta P_{tie}$, ACE Signal Dynamics




(b) Distance and threshold

**Fig. 6.11** Attack detection in the event of load sheds

**Table 6.4** Comparison with other detection methods

| Parameter | MSSA based (Proposed) | Kalman Filter | Dynamic analysis | OC-SVM |
|---|---|---|---|---|
| False positive rate(%) | 2.00 | 0 | 15.00 | 1.70 |
| False negative rate(%) | 0.01 | 100 | 18.00 | 0.78 |
| Accuracy | 98.96 | 0 | 98.8 | 98.79 |
| Detection time (samples) | 3 | 140 | 60 | 3 |

OC-SVM-based methods give good detection accuracy. However, the computation time is high in the Dynamic Analysis method, and OC-SVM is not adaptive to the system changes. Thus, the proposed method has better accuracy and computation burden than the existing methods from the comparison.

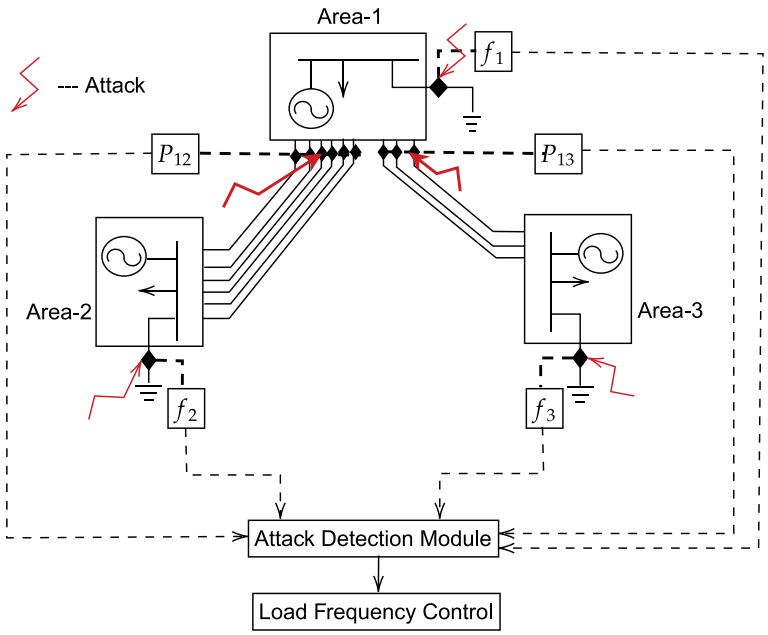



**Fig. 6.12** Attack Surface and Detection for IEEE 300-Bus system

### *6.11.4 Scalability Evaluation*

Practical power systems are very large compared to test systems. It is important to study the performance of detection algorithms for large-scale implementations. To establish the scalability of the proposed method, the algorithm is implemented on the 300-bus system (Demetriou et al. 2017), 1888-bus RTE (Réseau de Transport d'Électricité, France) system (RTE 2021).

Figure 6.12 shows the IEEE 300-bus system. It consists of three areas and thus there are three frequency measurements and there are 9 tie-lines. The detailed system parameters are given in Appendix A.

The 1888 RTE system is divided into 5 areas to study the detection algorithm. As shown in Fig. 6.13, the attack surface is very large for the 1888-bus system as there are more number of tie-lines between the areas and also 5 different frequency measurements.

Thus the above two systems are good candidate systems to evaluate the practical applicability of the algorithm.

It can be observed from Table 6.5 that the time required for the proposed detection is less than five cycles which is acceptable in the system. The variation in accuracy and number of detection cycles are minimal compared to the change in system size. Therefore, the proposed detection strategy applies to attack detection in practical power systems.

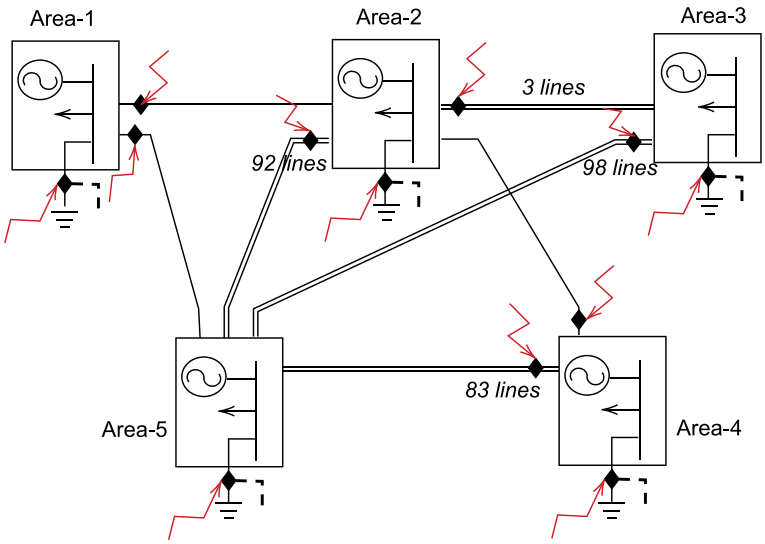



**Fig. 6.13** Attack surface for 1888 RTE system

**Table 6.5** Scalability analysis

| | 39-Bus NE (Bevrani 2014) | 300-Bus (Demetriou et al. 2017) | 1888-Bus RTE (RTE 2021) |
|---|---|---|---|
| Accuracy (%) | 98.96 | 97.85 | 97.42 |
| Detection time (samples) | 3 | 3 | 4 |

The bus and line data are obtained from the MATPOWER database for all the systems. The machine dynamic data for the 300-bus system is obtained using (Demetriou et al. 2017), and the load data is assumed. The machine data, load data, and forecast loads for the 1888-bus system are obtained from the RTE website. As the size of the system increases, the attack surface also proportionally increases. All the systems' data are given in Appendix A.

## 6.12 Summary

This chapter proposed a spectral analysis-based algorithm at two different system levels using S-SSA at the RTU/IED level and M-SSA at the control center level to detect attacks with a very low computation burden. The proposed method can successfully identify multiple coordinated and stealth attacks using measurements with noise with a high accuracy level since it is a data-based algorithm. The method's

significant advantages are that it can be implemented on limited hardware, and the time taken for attack detection is small such that it is detected before any significant impact is caused on the grid.

Multivariate detection can also be implemented within RTUs or IEDs, which have the capabilities of processing multiple signals. It is possible to extend the method to other power system applications with minimal hardware.

From the above results, the proposed algorithm has the following advantages:

1. The attack is detected within a span of 3 cycles. For an LFC system, the acceptable detection time is usually under five cycles, and thus the proposed algorithm can be effectively applied for fast attack detection. If the samples are available at the control center at a faster rate, the detection becomes faster. Fast detection can provide sufficient time to implement mitigation strategies to defend against attacks.
2. The algorithm is not dependent on the exact value of the predicted load, and any variations in the system do not degrade the algorithm's performance.
3. The proposed algorithm has a better performance than existing detection algorithms used for LFC.
4. The performance of the proposed algorithm is independent of the size of the system and thus is suitable for practical grid systems.
5. The computation burden is very low; thus it can be implemented even inside an existing IED with significantly less hardware requirement.
6. Power system topology is continuously changing. Algorithm performance is not impacted by system topology
7. Measurements have noise. The algorithm can work well with noisy measurements.

## References

Akbarian F, Ramezani A, Hamidi-Beheshti M-T, Haghighat V (2020) Advanced algorithm to detect stealthy cyber attacks on automatic generation control in smart grid. IET Cyber Phys Syst Theory Appl 5(4):351–358

Ameli A, Hooshyar A, Yazdavar AH, El-Saadany EF, Youssef A (2018) Attack detection for load frequency control systems using stochastic unknown input estimators. IEEE Trans Inf Forensics Secur 13(10):2575–2590

Amir D, Ofer Z (2010) Large deviations techniques and applications. Springer, Berlin Heidelberg

Ayad A, Khalaf M, El-Saadany E (2018) Detection of false data injection attacks in automatic generation control systems considering system nonlinearities. In: 2018 IEEE electrical power and energy conference (EPEC). ON, Canada, Toronto, pp 1–6

Bevrani H (2014) Robust power system frequency control. Springer

Bi W, Chen C, Zhang K (2019a) Optimal strategy of attack-defense interaction over load frequency control considering incomplete information. IEEE Access 7:75342–75349

Bi W, Zhang K, Li Y, Yuan K, Wang Y (2019b) Detection scheme against cyber-physical attacks on load frequency control based on dynamic characteristics analysis. IEEE Syst J 13(3):2859–2868

Chen C, Zhang K, Yuan K, Zhu L, Qian M (2018) Novel detection scheme design considering cyber attacks on load frequency control. IEEE Trans Ind Inf 14(5):1932–1941

Darwish I, Igbe O, Saadawi T (2015) Experimental and theoretical modeling of dnp3 attacks in smart grids. In: 2015 36th IEEE Sarnoff symposium. Newark, NJ, USA, pp 155–160

Demetriou P, Asprou M, Quiros-Tortos J, Kyriakides E (2017) Dynamic IEEE test systems for transient analysis. IEEE Syst J 11(4):2108–2117

East S, Butts J, Papa M, Shenoi S (2009) A taxonomy of attacks on the dnp3 protocol. In: Palmer C, Shenoi S (eds) Critical infrastructure protection III. Springer, Berlin, Heidelberg, pp 67–81

Hossein Hassani RM (2018) Singular spectrum analysis using R. Palgrave Macmillan

Jaksic V, Ogata Y, Pillet C-A, Seiringer R (2012) Quantum hypothesis testing and non-equilibrium statistical mechanics. Rev Math Phys 24(06):1230002

Khalaf M, Youssef A, El-Saadany E (2019) Joint detection and mitigation of false data injection attacks in agc systems. IEEE Trans Smart Grid 10(5):4985–4995

Klema V, Laub A (1980) The singular value decomposition: its computation and some applications. IEEE Trans Autom Control 25(2):164–176

Li Y, Huang R, Ma L (2021) False data injection attack and defense method on load frequency control. IEEE Internet of Things J 8(4):2910–2919

Malioutov D, Cetin M, Willsky AS (2005) A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans Signal Process 53(8):3010–3022

Mehra R, Peschon J (1971) An innovations approach to fault detection and diagnosis in dynamic systems. Automatica 7(5):637–640

Nagaoka H, Hayashi M (2007) An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses. IEEE Trans Inf Theory 53(2):534–549

RTE (2021). Eco2mix-electricity consumption in France

von Bünau P, Meinecke FC, Király FC, Müller K-R (2009) Finding stationary subspaces in multivariate time series. Phys Rev Lett 103:214101

Wald A (1945) Sequential tests of statistical hypotheses. Ann Math Stat 16(2):117–186

Wang P, Govindarasu M (2018) Anomaly detection for power system generation control based on hierarchical dbscan. In: 2018 North American power symposium (NAPS), pp 1–5

# Chapter 7
# Machine Learning-Based Attack Detection


**Abstract** This chapter presents an innovative approach to enhance the cybersecurity of smart grid systems through the utilization of machine learning techniques. It commences with a comprehensive introduction and then delves into the pivotal role of machine learning in the context of smart grid attack detection. A focal point emerges in the form of Support Vector Data Description (SVDD) for online attack detection, elucidating its core components. Simulating the application of SVDD, the chapter meticulously details the results and engages in insightful discussions. Furthermore, a comparative analysis with other classifiers is presented, shedding light on the strengths of the SVDD approach. In summary, this chapter offers a comprehensive exploration of machine learning-based attack detection in smart grids, featuring practical simulation results and discussions. It underscores the effectiveness and adaptability of the SVDD methodology while providing valuable insights into its application in real-world scenarios.

**Keywords** Support vector data description · Machine learning · Attack detection · Zero-day attacks

## 7.1 Introduction

The present-day grid control systems can use data from sensors and actuators and artificial intelligence algorithms to perform timely qualitative and quantitative analysis to understand the dynamics of the control system under various system operating conditions and fault conditions. In the past few years, various algorithms have been developed to detect attacks using the known attack semantics and RTU and IED data.

For attack detection at various levels of the power system, various strategies have been presented in the literature. Denial of Service (DoS) has been extensively investigated because they are one of the most accessible forms of attack (Liu et al. 2019; Cheng et al. 2020). However, attack patterns have evolved, and attempts to



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_7

modify or append the sensor and actuator measurements Sridhar and Manimaran (2010) require further investigation because such data integrity attacks can have a direct and considerable impact on the system's economy and stability (Tan et al. 2017).

Ideally, every system should have some basic attack-defense capabilities. Attackers attempt to breach these capabilities, and the system defends itself, giving rise to a game-theoretic model of cyber attack-defense interaction. The authors of Bi et al. (2019a) focus on the limitations of knowledge available to attackers and defenders, and they propose a game-theoretic approach to model attack detection. Attack-specific detection strategies imply that analysis of attack techniques is required for detection. Chen et al. (2018) presents a unified model consisting of detection corresponding to exogenous and scaling attacks on tie-line and frequency measurements, as well as discussions of their effects on frequency and tie-line power. Bi et al. (2019b) discusses Fixed and Variable attacks, as well as the differences in their impacts and detection.

The Kalman filter (Akbarian et al. 2020) and Stochastic Unknown Input Estimators (Ameli et al. 2018) can be employed to estimate LFC states using outputs and initial states. Attack detection is achieved by comparing the estimates with the measurements. The accuracy of system models used for the estimate is critical for model-based detection tactics. The authors of Wang and Govindarasu (2018) derive conformity measures by observing the behavior of generators in the same balance area. These measurements are then employed with a semi-supervised clustering approach called Hierarchical Density-based Spatial Clustering of Application with Noise to detect aberrant generation controls caused by cyberattacks (HDBSCAN). The model is trained using a series of attack templates. It is a data-based algorithm because it uses raw data.

In this chapter, we discuss the general framework for using machine learning methods for smart grids for attack detection. This is followed by a Support Vector Data Description (SVDD) based attack detection strategy. An adaptive support vector data description-based attack detection strategy is developed to detect zero-day attacks for fast and reliable real-time attack detection. The detection has two SVDD modules: SVDD-A for attack detection and SVDD-Z to classify zero-day attacks from known attacks.

## 7.2 Machine Learning in Smart Grid Attack Detection

There are several machine learning and artificial intelligence-based techniques that have the required capabilities to be applied to the detection of attacks from smart grids. The key steps involved in implementing an attack detection strategy using machine learning involve some common steps, such as effectively training models, deploying them in a production environment, and continuously monitoring for threats.

Here are the essential steps in the implementation of such a strategy:

1. **Define Objectives and Scope**
   - Clearly define the objectives of your attack detection strategy. What types of attacks are you aiming to detect? What is the scope of your detection efforts (e.g., network, endpoint, application layer)?
2. **Data Collection and Preprocessing**
   - Collect high-quality, labeled data that includes both normal and malicious activities. Data sources may include logs, network traffic, or endpoint telemetry.
   - Preprocess and clean the data to remove noise, handle missing values, and transform it into a suitable format for machine learning algorithms.
3. **Feature Engineering**
   - Identify relevant features (attributes or variables) from the data that can aid in distinguishing between normal and malicious behavior.
   - Extract, select, or engineer features that capture meaningful information about the system's behavior.
4. **Data Splitting**
   - Divide the labeled data into training, validation, and test sets. The training set is used to train machine learning models, the validation set helps in hyperparameter tuning, and the test set evaluates model performance.
5. **Select Machine Learning Algorithms**
   - Choose machine learning algorithms suitable for your detection problem. Common choices include decision trees, random forests, support vector machines, neural networks, and anomaly detection methods.
   - Consider ensemble methods for combining multiple models to improve accuracy and robustness.
6. **Model Training**
   - Train the selected machine learning models on the training dataset. Fine-tune hyperparameters to optimize model performance.
   - Experiment with different algorithms and configurations to find the best-performing models.
7. **Validation and Cross-Validation**
   - Validate model performance using the validation dataset. Employ cross-validation techniques to assess model stability and generalization.
   - Evaluate metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and others relevant to your use case.

8. **Model Deployment**

- Deploy the trained models in a production environment where they can monitor real-time or near-real-time data for signs of attacks.
- Ensure scalability, fault tolerance, and low-latency processing in the deployment infrastructure.

9. **Continuous Monitoring**

- Implement continuous monitoring of model performance in production. Set up alerting mechanisms to notify security teams of potential issues or model degradation.
- Regularly retrain models with fresh data to adapt to evolving attack techniques and maintain high detection accuracy.

## 7.3 Support Vector Data Description Based Online Attack Detection

This section proposes a cyber attack detection framework that utilizes Support Vector Data Description (SVDD). Figure 7.1 shows the detailed system model of one balancing area of a multi-area power system (Wood et al. 2013) along with the attack surface and the detection algorithm implementation.

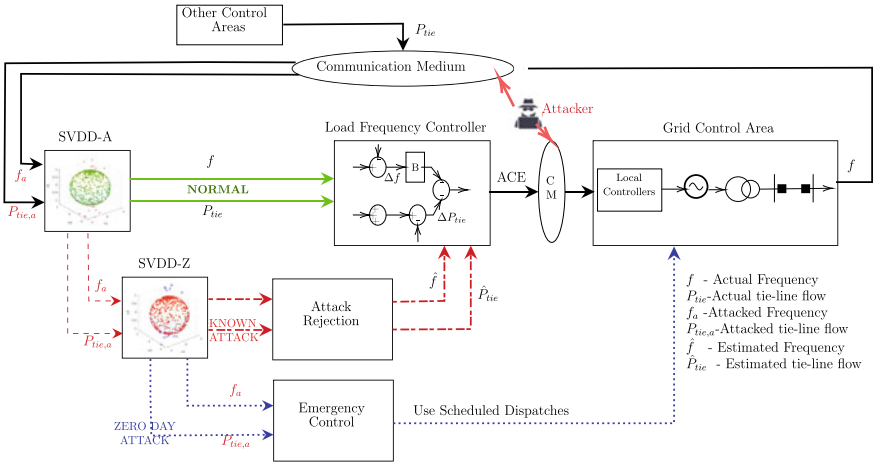



**Fig. 7.1** Load frequency control with attacks and detection

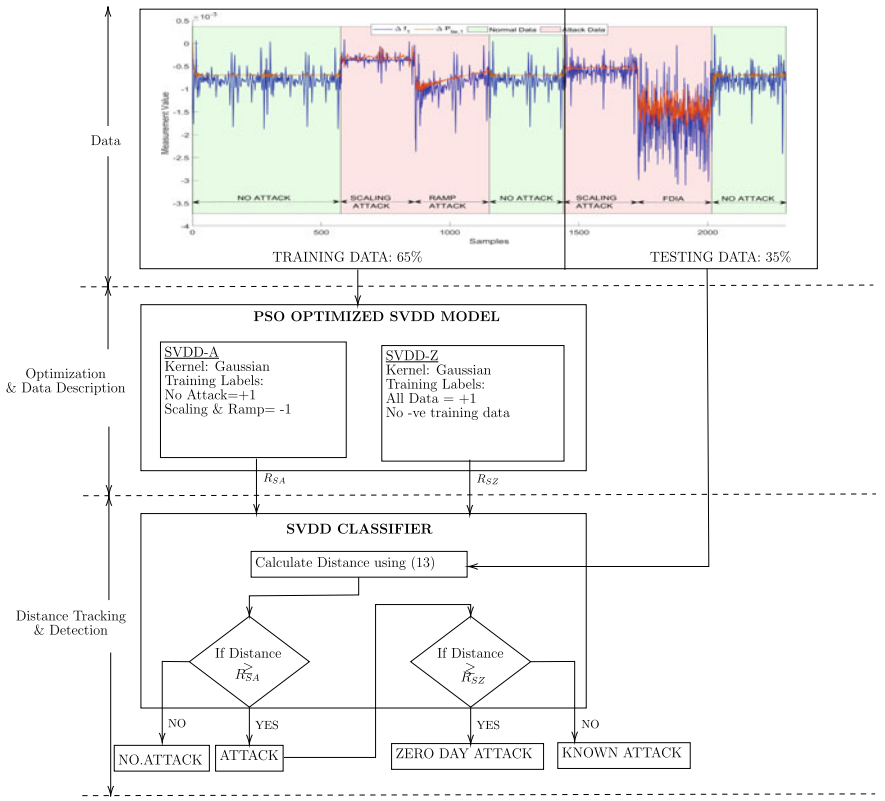



**Fig. 7.2** Intelligent attack detection using SVDD

SVDD obtains a spherically shaped boundary around a data set. We make the spherical boundary more flexible in the proposed strategy by using appropriate kernel functions. Thus, it is an efficient tool to detect unknown or zero-day attacks.

The detection algorithm is to be implemented in real time. The real-time grid frequency and tie-line data sent to the LFC control center is first sent to the SVDD-A module. This module is trained to differentiate attack data from normal data. Thus it is called the attack detection SVDD module.

Once the attack is detected, it is required to identify the attacks to implement suitable mitigation strategies. The data is thus sent to the SVDD-Z module to check if the detected attack is a zero-day or known attack. If it is a known attack, the attack is identified and then a suitable planned mitigation corresponding to the identified attack is implemented. If the attack is classified as a zero-day attack, the immediate response is to implement emergency action to safeguard the LFC system. Then the SVDD-Z model is trained with the new attack and updated. At the same time, appropriate mitigation steps are devised corresponding to the new attack. Attacks

whose signatures are not available to the defenders can also be detected using the proposed method.

An overview of the steps involved in each SVDD module is shown in Fig. 7.2. The detection algorithm is explained in words in Algorithm 7.1.

**Algorithm 7.1:** SVDD Based Zero Day Attack Detection

**DATA GENERATION**

**Input:** Historical Load Data $\Delta P_l$

**Output:** $\Delta f$, $\Delta P_{tie}$

1. Generate the $\Delta f$, $\Delta P_{tie}$ using $\Delta P_l$ as input to the LFC model.
2. Fix the $\beta_s$, $\beta_p$, $\beta_r$ values and generate data with attack using (3.10).
3. Divide the data into training and test data.

**TRAINING PHASE**

**Input:** $\Delta f$, $\Delta P_{tie}$ training data, Kernel type, Data Labels

**Output:** Trained SVDD Model, $R_{SA}$ and $R_{SZ}$

1. Use the training data to find optimal value of $P$ and $b_f$ of hypersphere (Detailed in 7.3.3).
2. Using obtained hyperspheres, find the radii $R_{SA}$ and $R_{SZ}$ using (7.7)

**DETECTION PHASE**

**Input:** SVDD Model, $R_{SA}$ and $R_{SZ}$

**Output:** Attack: Yes or No, Zero Day Attack: Yes or No

1. Calculate the distance of each test data using (7.6)
2. If Distance $\geq R_{SA}$, attack detected.
3. If Distance $\geq R_{SZ}$, zero-day attack.

### *7.3.1 Normal Data Description in SVDD*

SVDD models a hypersphere, with center '$c$' and radius '$R$', which includes all the training data to produce a description of the normal data. To obtain the parameters of the hypersphere, we minimize the volume of the sphere subject to the inclusion of all the data points ($y_1$, $y_2$, ..$y_i$...),

$$Min \;\; F(R, c, \varepsilon_i) = R^2 + P \sum_i \varepsilon_i \tag{7.1a}$$

$$Subject\ to, \; ||y_i - c||^2 \leq R^2 + \varepsilon_i, \; i = 1, 2, ..., n, \varepsilon_i \geq 0. \tag{7.1b}$$

$P$ and $\varepsilon$ are penalty coefficient and relaxation variables, respectively, that handle the possibility of outliers in the training set. By using Lagrange multipliers ($\alpha_i$, $\gamma_i$), the above equation can be simplified as

$$L = R^2 + P\sum_i \varepsilon_i - \sum_i \alpha_i[R^2 + \varepsilon_i - (||y_i||^2 - 2c.y_i + ||c||^2)] - \sum_i \gamma_i \varepsilon_i. \tag{7.2}$$

Since the Lagrange multipliers should be positive and $L$ should be minimized, the dual of Eq. 7.1b is obtained as

$$Max\ L = \sum_i \alpha_i(y_i, y_i) - \sum_{i,j} \alpha_i \alpha_j (x_i, x_j) \tag{7.3a}$$

$$s.t\ 0 \leq \alpha_i \leq P,\ i = 1, ..., n. \tag{7.3b}$$

The algorithm is further improved by considering Kernel functions in place of the dot products. Thus (7.3b) can be written as,

$$Max\ L = \sum_i \alpha_i K(y_i, y_i) - \sum_{i,j} \alpha_i \alpha_j K(y_i, y_j) \tag{7.4a}$$

$$s.t\ 0 \leq \alpha_i \leq P,\ i = 1, ..., n. \tag{7.4b}$$

Several kernel functions can be used. It is observed that the Gaussian kernel function gives the best result for our application, and thus we use the Gaussian kernel.

Using the KKT conditions, solving Eq. 7.4b gives three conditions on the variable $\alpha_i$

$$||y_i - c||^2 < R^2 \Rightarrow \alpha_i = 0 \tag{7.5a}$$

$$||y_i - c||^2 = R^2 \Rightarrow 0 < \alpha_i < P \tag{7.5b}$$

$$||y_i - c||^2 > R^2 \Rightarrow \alpha_i = P. \tag{7.5c}$$

Equation 7.5b represents the data points that are on the hypersphere boundary. The center of the sphere is a linear combination of all the data points. However, only those data points which satisfy the condition $\alpha_i \geq 0$ are required to represent the boundary of the training data samples. These particular $y_i$ values are called the support vectors.

### 7.3.2 Distance Tracking and Detection

In the detection or testing phase, any test vector that falls within the hypersphere is normal. Any measurement that falls out of this boundary will be considered anomalous, and the attack alarm will go off. A test vector $yt_i$ will be within sphere if it satisfies

$$||yt_i - c||^2 = (yt_i.yt_i) - 2\sum_i \alpha_i(yt_i.y_i) + \sum_{i,j} \alpha_i \alpha_j (y_i.y_j) \leq R^2. \tag{7.6}$$

where the radius is the distance of any support vector from the center

$$R^2 = (y_k.y_k) - 2\sum_i \alpha_i (y_i.y_k) + \sum_{i,j} \alpha_i \alpha_j (y_i.y_j). \tag{7.7}$$

Each SVDD module will have a different radius which we define as follows:

1. $R_{SA}$ = Radius for model SVDD-A.
2. $R_{SZ}$ = Radius for model SVDD-Z.

### 7.3.3 Optimization-Based Parameter Selection

The major challenge in the SVDD algorithm is designing the SVDD parameters ($P$ and kernel parameters). The selection of optimal parameters can significantly improve the calculation accuracy, simplify the calculation complexity, and improve the speed of the detection process.

If the support vector description rejects an object from the target distribution, it is an error. A $P$ value of 1.0 indicates that all target data should be accepted, which is not a reasonable assumption, and thus selection of $P$ determines the number of outliers. The kernel parameter determines the generalization ability. To set the kernel parameters (for example, width of the Gaussian kernel), we consider the target acceptance rate since as the width increases, the number of target data included in the description becomes larger.

The Particle Swarm Optimization (PSO) chooses the optimal value of the parameters $P$ and width $b_f$. PSO is selected due to its reduced implementation complexity, accuracy, and simplicity in finding optimal solutions. Moreover, PSO has shown excellent performance in parameter optimization for various nonlinear and real-world applications. Thus, the PSO algorithm is a near-ideal option for choosing the SVDD parameters since its structure allows the particles to preserve the best previous experiences over multiple generations.

The results in the upcoming sections will show that the algorithm gives a good accuracy of classification and can be implemented fast. The limitation of the proposed algorithm is that the performance can be affected depending on the choice of Kernel. For the given system and data, the Gaussian kernel was shown to give good results. However, for a different system, a suitable choice of kernel is highly important.

**Table 7.1** Decision table for the SVDD modules

| Conditions | Attack type | Output: SVDD-A | Output: SVDD-Z |
|---|---|---|---|
| Normal data | – | Normal/No Attack | – |
| Attack data | Scaling (SA) | Attack | Known attack |
| | Ramp (RA) | Attack | Known attack |
| | Denial of service (DoS) | Attack | Known attack |
| | FDIA | Attack | Zero-day attack |

## 7.4 Simulation Results and Discussions

### *7.4.1 Example 7.1: SVDD Detection for Attacks*

The above algorithm is tested on the frequency control of a 39-bus New England test system. Simulations are carried out using MATLAB 2021a on a Core i5 processor system. The system consists of 3 areas with multiple tie-lines between the areas as shown in Fig. 6.2.

### *7.4.2 Data Preparation*

For testing the algorithm, we use load forecast data and actual load data obtained from the New England ISO website to simulate the LFC operation under normal grid load variations. Different attacks as described in (3.10) are then injected into the system. The Bernoulli variables describe the instant of these attacks. To obtain the study dataset, noise is added to the signals at different signal-to-noise ratios (SNR).

The attack types and classification are as shown in Table 7.1. The attack types are indicated as N: No attack; SA: Scaling Attack; RA: Ramp Attack; DoS: Denial of Service Attack. The random attacks are used as the unknown attack (represented by $*$) to test the algorithm for zero-day attacks. Thus the two modules classify the attacks as follows:

1. SVDD-A: Normal (N) and Attack (SA, RA, DoS,$*$).
2. SVDD-Z: Known (SA,RA,DoS), Unknown ($*$).

Figure 7.2 shows the plot of the frequency and tie-line signals of balancing area-1 under normal and attack conditions. It can be observed that the variations in the signal are negligible. Thus, it cannot be immediately identified by the operators or by using bad data detection and Kalman filter estimations, which establishes the need for a more advanced and accurate detection process.

**Table 7.2** Optimal parameter values for SVDD modules

| | $P$ | $b_f$ | False positive rate | Accuracy |
|---|---|---|---|---|
| SVDD-A | 0.728 | 0.084 | 0.041 | 92.312 |
| SVDD-Z | 0.156 | 100.180 | 0.011 | 97.443 |

### *7.4.3 Particle Swarm Optimization (PSO) Based Parameter Selection*

PSO is used to determine the fitness parameters for $P$ (penalty parameter) and $b_f$ (Kernel width). The optimal fitness and values of $P$ and $b_f$ at optimal fitness for SVDD-A and SVDD-Z are indicated in Table 7.2.

### *7.4.4 Detection Results*

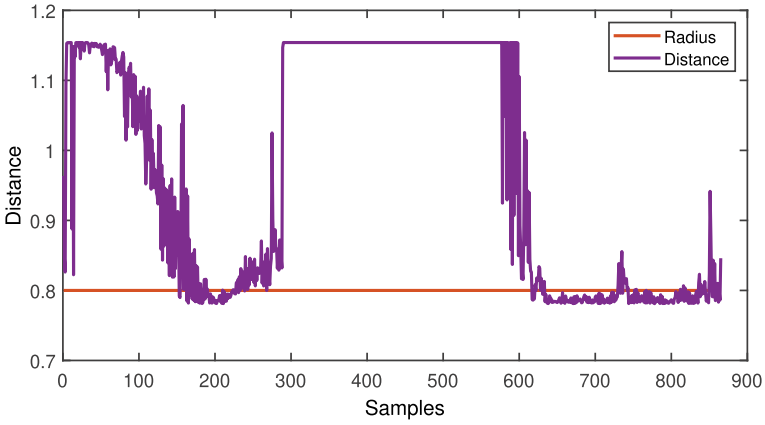
Figure 7.3a shows the results for SVDD-A, and Fig. 7.3b shows the detection results for SVDD-Z. In Fig. 7.3a, the attack is detected if the Distance value is greater than the red line, i.e., the radius. Similarly, in Fig. 7.3b, the samples with distance values greater than the radius indicate a zero-day attack.

We use the Accuracy, False Positive Rate, and the Area Under ROC Curve to evaluate the algorithm's performance.

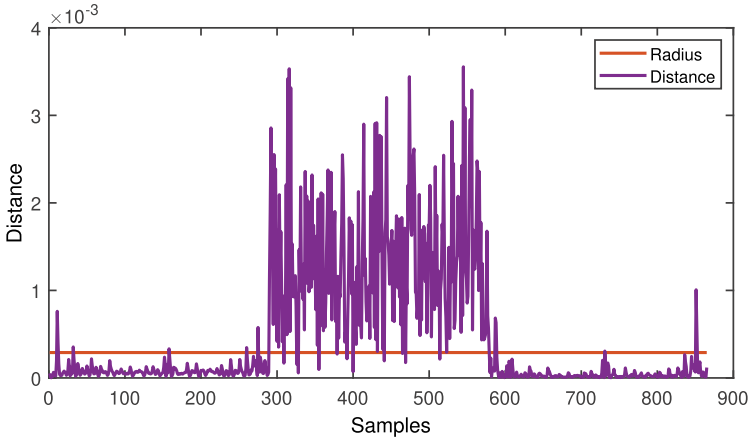
1. *Accuracy* is a measure of correct predictions for the dataset. A high level of accuracy is expected in the grid environment as it is a critical control system.
2. *False positive rate* (FPR) is the fraction of normal data that is detected as an attack. In the power system environment, the FPR must be very low since detecting normal data as an attack could lead to downtime in the system operation, which is unacceptable.

Figure 7.3 is obtained based on the test data from Fig. 7.2. The test data consists of scaling attack, FDIA, and normal data, respectively. Figure 7.3a shows the data points to be well above the threshold for the attack part and below the threshold for the no-attack part. In Fig. 7.3b, since scaling attack is a known attack, the data points for scaling attack are also below the threshold and only the FDIA is classified as zero-day attack.

The accuracy and false-positive rates are shown in Table 7.2. It can be observed that the proposed algorithm gives a very low FPR for the SVDD-A algorithm, which is acceptable. The high accuracy makes the algorithm suitable for implementation in smart grid controls.




(a) SVDD-A




(b) SVDD-Z (Unknown FDIA)

**Fig. 7.3** Test results for attack detection in SVDD-A and zero-day attack detection in SVDD-Z

### *7.4.5 Comparison with Other Classifiers*

In this section, we will evaluate the performance of the proposed SVDD-A, Neural Network (NN), K-Nearest Neighbor (KNN), and Naive Bayes (NB) Classifier and Gaussian Support Vector Machine (SVM) in LFC attack detection. Since some methods need both positive and negative values for training, the step attack, which is the simplest attack type, will be used to generate the negative training samples. All the algorithms are trained on this dataset.

Table 7.3 gives the results of the classifications, and Fig. 7.4 shows the ROC curves for the different methods along with the Area under them.

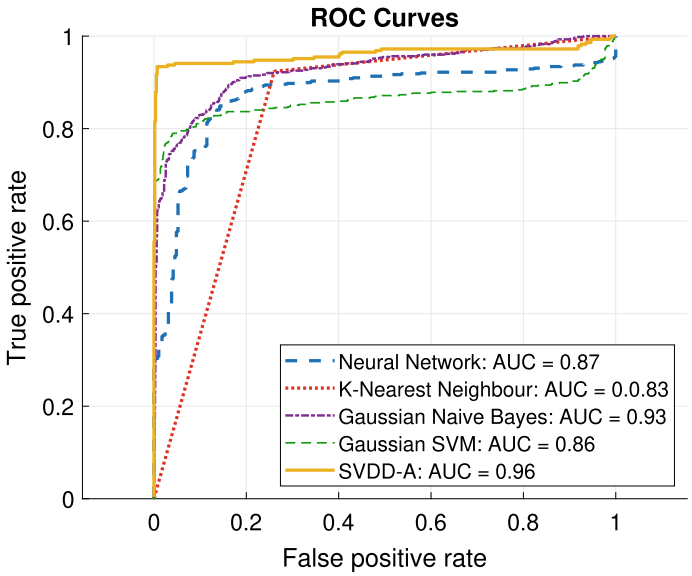



**Fig. 7.4** Receiver operator characteristics for various algorithms

**Table 7.3** Performance of ML algorithms in attack detection

| ML algorithm | False positive rate (%) | Accuracy (%) | Training time (s) | Testing time (per sample) (ms) |
|---|---|---|---|---|
| SVDD-A | 0.425 | 97.332 | 2.000 | 0.025 |
| Neural network | 7.243 | 90.762 | 1.496 | 0.002 |
| KNN | 6.576 | 86.212 | 1.675 | 0.032 |
| Gaussian SVM | 0.373 | 76.239 | 0.780 | 0.021 |
| Naive Bayes | 6.182 | 87.814 | 2.760 | 0.005 |

The accuracy and FPR are dependent on the threshold selected for the algorithm. One threshold can give better result than the others, and thus the accuracy and FPR cannot be used as good measures for comparing different algorithms.

The Receiver Operator Characteristic (ROC) is a probability curve that plots the TPR against FPR at various threshold values. The Area Under the Curve (AUC) measures the ability of a classifier to distinguish between classes. The higher the AUC, the better the model's performance at distinguishing between two classes. Thus, the area under the ROC curve gives a better comparison between different methods. The higher the AUC, the better the algorithm's classification performance. From Fig. 7.4, it can be observed that the AUC is maximum for the proposed SVDD-A method as compared to other existing classifiers in the literature. Thus, the proposed algorithm is a better choice for attack classification in grid control systems.

It is evident from Table 7.3 that the proposed optimized SVDD-A gives much better accuracy than other existing binary classifier machine learning algorithms. The model's training is done offline, so the higher training time is acceptable. However, the testing time is very low, making the method suitable for grid control systems.

### *7.4.6 Summary of Results*

From the obtained results of the algorithm, the following conclusions may be drawn:

1. The SVDD-A algorithm is able to successfully detect whether there is an attack or not for both known and zero-day attacks with good accuracy.
2. The SVDD-Z algorithm successfully classifies the attacks as known or zero-day attack, thus facilitating fast mitigation.
3. Both the algorithms are found to be computationally efficient and act very fast which is highly important in the power grid environment.
4. The algorithm performs better than most of the existing machine learning algorithms.
5. The results on the large-scale system show the scalability of the proposed method to practical large-scale grid systems.

In the algorithm, the normal training data is taken over a long period of time; thus encapsulating the different contingency states that could occur in the system. The algorithm has the limitation that if any new contingency state comes up that is not included in the training, these data points could be classified as attacks by the SVDD-A algorithm. This can be avoided by retraining the model at regular intervals of time by including any new contingency events thus making the algorithm adaptive. Since the training is not computationally complex, retraining the algorithm does not impose a burden on the system operation. The algorithm is thus re-trained in the following events:

1. At definite time-intervals: To encapsulate the changes in the grid.
2. When a zero-day attack is detected: To add the new attack type to the category of known attacks.

## 7.5 Summary

This chapter proposes a machine learning-based Support Vector Data Description (SVDD) model to detect zero-day attacks in Load Frequency Control of a Smart Grid. The authors use an LFC model with parametric uncertainties and nonlinearities to generate accurate training data. This modeling encapsulates the actual grid conditions and thus, improves the training accuracy.

SVDD has been a promising approach to classifying the data when only one class of training data is available. This paper proposes a two-step attack detection, namely SVDD-A, to detect attacks, and SVDD-Z to classify zero-day attacks. The SVDD-A module classifies normal data from attack data. Once attacks are detected, appropriate mitigation steps can be incorporated based on the attack type. In addition, an SVDD-Z module is incorporated to classify zero-day attacks from known attacks to avoid wrong mitigation steps.

In order to test the algorithm, a multiple and time varying attack model has been used. It accurately classifies the scaling and step attacks and distinguishes them from the random FDIA as zero-day attacks. The high accuracy and low false positive rates suggest the suitability of the algorithm for smart grid control attack detection. It is also compared with various other algorithms and the area under the ROC curve, which is better than the existing algorithms. Thus, the proposed algorithm can effectively distinguish between normal, known attack, and zero-day attack conditions.

The proposed detection strategy can be effectively implemented inside the power system control centers to detect cyber attacks in power grid control systems. Such a fast detection technique gives the system operators sufficient time to implement mitigation and response plans to protect the grid from collapse.

## References

Akbarian F, Ramezani A, Hamidi-Beheshti M-T, Haghighat V (2020) Advanced algorithm to detect stealthy cyber attacks on automatic generation control in smart grid. IET Cyber Phys Syst Theory Appl 5(4):351–358

Ameli A, Hooshyar A, Yazdavar AH, El-Saadany EF, Youssef A (2018) Attack detection for load frequency control systems using stochastic unknown input estimators. IEEE Trans Inf Forensics Secur 13(10):2575–2590

Bi W, Chen C, Zhang K (2019a) Optimal strategy of attack-defense interaction over load frequency control considering incomplete information. IEEE Access 7:75342–75349

Bi W, Zhang K, Li Y, Yuan K, Wang Y (2019b) Detection scheme against cyber-physical attacks on load frequency control based on dynamic characteristics analysis. IEEE Syst J 13(3):2859–2868

Chen C, Zhang K, Yuan K, Zhu L, Qian M (2018) Novel detection scheme design considering cyber attacks on load frequency control. IEEE Trans Ind Inf 14(5):1932–1941

Cheng Z, Yue D, Hu S, Huang C, Dou C, Chen L (2020) Resilient load frequency control design: Dos attacks against additional control loop. Int J Electr Power Energy Syst 115:105496

Liu J, Gu Y, Zha L, Liu Y, Cao J (2019) Event-triggered $h_\infty$ load frequency control for multiarea power systems under hybrid cyber attacks. IEEE Trans Syst Man Cybern Syst 49(8):1665–1678

Sridhar S, Manimaran G (2010) Data integrity attacks and their impacts on SCADA control system. IEEE PES general meeting. Minneapolis, MN, USA, pp 1–6

Tan R, Nguyen HH, Foo EYS, Yau DKY, Kalbarczyk Z, Iyer RK, Gooi HB (2017) Modeling and mitigating impact of false data injection attacks on automatic generation control. IEEE Trans Inf Forensics Secur 12(7):1609–1624

Wang P, Govindarasu M (2018) Anomaly detection for power system generation control based on hierarchical dbscan. In: 2018 North American power symposium (NAPS), pp 1–5

Wood AJ, Wollenberg BF, Sheble GB (2013) Power generation automation and control. Wiley

# Chapter 8
# Attack Mitigation and Recovery in Smart Grid Control


**Abstract** This chapter explores the critical domain of attack mitigation and recovery within the context of smart grid control systems. As smart grids become increasingly integral to modern energy infrastructure, the need for robust cyber-security measures to safeguard against malicious attacks is paramount. The chapter begins by elucidating various Attack Mitigation methods in Smart Grids. Through a comprehensive overview, readers gain insights into the diverse techniques available to protect smart grid systems against cyber threats. A focal point of this chapter is the exploration of Attack Mitigation for a 39-bus 3-area system, considering both single-step load and dynamic load scenarios. Through a detailed examination, readers are presented with a basic yet illustrative example of how mitigation strategies can be implemented within a complex smart grid control environment. Furthermore, the chapter delves into the innovative realm of IoT-based hardware models for enhancing attack mitigation and recovery capabilities. We describe the method to build a simple IoT model that can be used to launch attacks and implement detection methods using Kali Linux, Raspberry Pi and Python programming. Through a blend of theoretical frameworks and practical examples, this chapter equips readers with the knowledge and tools necessary to bolster the resilience of smart grid control systems against cyber threats.

**Keywords** Attack mitigation · Adaptive control · Hardware-in-loop · Internet of things

## 8.1 Introduction

Once attacks are detected, it is also important for the system to respond immediately to the attacks and mitigate them such that they do not cause further damage to the grid. Attack mitigation can be done in three different ways:



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8_8

1. Estimation-based: Using estimated signals (Frequency and Tie-line flow) for LFC instead of actual measurements.
2. Attack minimization/elimination: Using control strategies by including the detected attack models.
3. Bypass LFC.

Each of the above techniques has its own advantages and disadvantages. The most commonly followed method to eliminate the effect of attacks on a control system is to ignore the spoofed sensors, obtain state estimates for the missing sensors, and then use the original controller to respond to the attack.

However, this method has the following drawbacks:

1. Using the original controller cannot guarantee system safety under attacks.
2. For safety-critical systems, recovery time deadlines need to be included in the formulation of attack mitigation.
3. The common assumption that the defender knows the exact physical model of the system under attack is rare in practical systems.

Thus, when attacked, the system should employ a controller that has the ability to drive the system back to its normal state. In this chapter, we discuss a simple yet effective attack mitigation strategy.

The attack and the proposed detection strategy is then implemented into an IoT based hardware setup to illustrate the effectiveness of the complete framework discussed in this chapter.

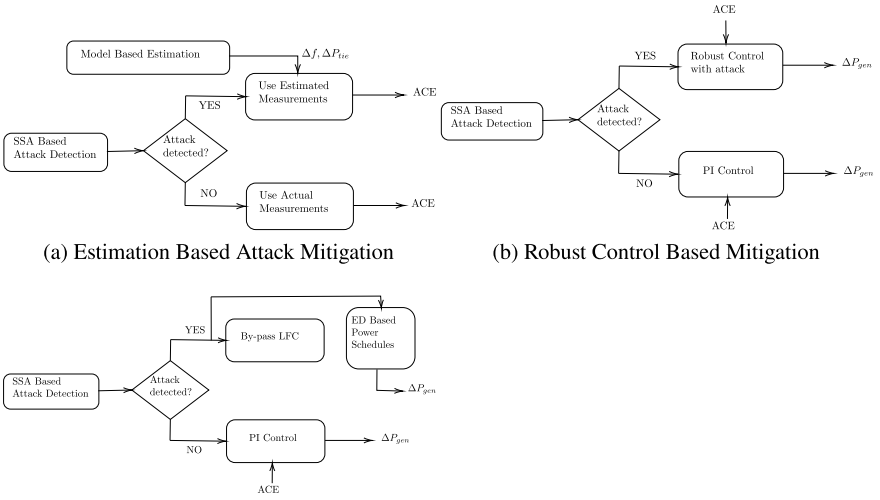
## 8.2 Attack Mitigation in Smart Grids

Attack mitigation can happen in three different ways (Fig. 8.1):

1. Estimation-based: Using estimated signals (frequency and tie-line flow) for LFC instead of actual measurements.
2. Attack minimization/elimination: Using control strategies by including the detected attack models.
3. Bypass LFC.

### *8.2.1 Estimation-Based Mitigation*

In Sridhar and Govindarasu (2014), once attacks are detected, the authors use the load forecast data to predict the ACE values which are then used in the LFC operation instead of the actual measured values because they are corrupted. The stochastic unknown input estimator in Ameli et al. (2018) can also be used to determine the states of the system without the need for load forecast data.




(a) Estimation Based Attack Mitigation

(b) Robust Control Based Mitigation

(c) By-pass LFC for attack mitigation

**Fig. 8.1** Attack mitigation methods

### *8.2.2 Attack Elimination Using Robust Control*

Robust control strategies such as model predictive control (MPC) (Liu et al. 2021), distributed event-triggered mechanism (DETM), etc., can be used to determine a control strategy in the presence of an attack. In this type of mitigation, a limit is placed on the attacks which are then modeled exactly or as uncertainties. Then, a controller is designed by including these models. Lyapunov stability margins are established to ensure system stability.

### *8.2.3 Bypass LFC*

When an attack is detected, the emergency control actions are implemented and the generation schedules are adjusted by operators using the results of the economic dispatch solution. This method is the easiest and is similar to having a system without load frequency control.

Table 8.1 gives in detail the different methods of attack mitigation.

As seen from the table above, different types of mitigation algorithms have been applied for the attack mitigation in LFC. Table 8.2 gives the advantages and disadvantages of different mitigation methods.

As seen from the table above, even though robust control-based mitigation is most effective, its stability criteria depends on the particular attacks and can fail if the type of attack is different. Estimation-based algorithms perform well except when there is an emergency or contingency condition. Therefore, it is best to choose a combination of different methods for different grid conditions.

**Table 8.1** Attack mitigation techniques

| S. No. | Method | Model | Attacks | Stability criterion | References |
|---|---|---|---|---|---|
| *Estimation-based mitigation* | | | | | |
| 1 | Load forecast based ACE | Linearized LFC | FDIA | – | Sridhar and Govindarasu (2014) |
| 2 | Stochastic unknown input estimator | Linearized LFC | FDIA | – | Ameli et al. (2018) |
| *Robust control based mitigation* | | | | | |
| 3 | Dynamic event-based model predictive control | Dynamic event-triggered (DETM) linear LFC | Deception | $H_2/H_\infty$ Performance | Liu et al. (2021) |
| 4 | Resilient load frequency design | Non-linear and uncertain model | DoS | Lyapunov-Krasovski | Cheng et al. (2020) |
| 5 | Resilient distributed co-ordination control | DETM LFC-Virtual Inertia Control | Alternating deception and DoS | $H_\infty$ Performance | Cheng et al. (2021) |
| *Bypass LFC* | | | | | |
| 6 | Use Economic Dispatch (ED) based dispatches | Power system ED | FDIA, DoS | – | Bi et al. (2019) |

**Table 8.2** Comparison of attack mitigation

| Method | Advantages | Disadvantages |
|---|---|---|
| Estimation-based mitigation | Attack models need not be considered for estimations. High performance estimation methods are available considering nonlinearities and noises | Events like faults can not be considered in the estimation. Highly dependant on load forecasts |
| Robust control-based elimination | Provides efficient control to eliminate attacks and use the measurements | Highly dependant on attack and system models considered during control system modeling |
| Bypass LFC | Easiest mitigation method | Slow response. Depends on efficiency of ED results |

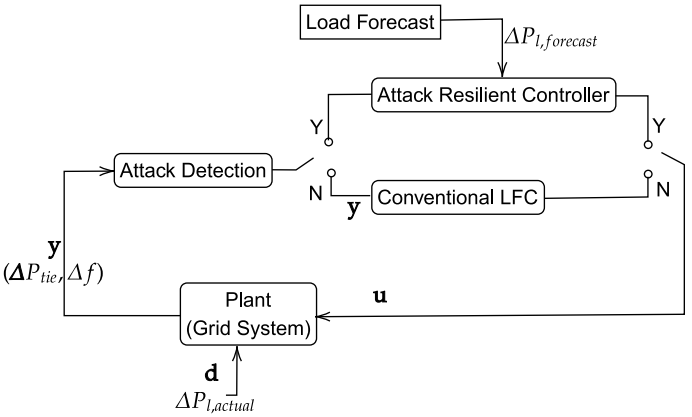



**Fig. 8.2** Attack mitigation/recovery

## 8.3 Adaptive Control-Based Attack Mitigation

Figure 8.2 gives an overview of the attack mitigation/recovery strategy.

As shown in Fig. 8.2, while the system is working in normal conditions (no attack detected), the controller action is via the PI controller present in the LFC system.

If an attack is detected, the control shifts from normal operation to a special control action that uses the system model to generate the control input or ACE. Thus, the control input is independent of the measurements obtained at the control center and only depends on the load data and the model.

Load forecasts are used for generating the control input during an attack. There are well-established methods for obtaining them accurately using load data, weather data, and user behavior analysis. Thus, load forecast-based control input determination has the capability to provide good results for sustaining the grid during an attack. The LFC system can be restored once the system is safe from attacks.

## 8.4 Attack Mitigation for 39-Bus 3 Area System

To study the attack mitigation strategy, we use the 39-bus 3 area test system as shown in Fig. 8.3.

For this system, the load forecast and actual data are available from the New England ISO website. For better understanding of the attack-resilient LFC, we first look at results for a single step load change followed by the actual load variation results.

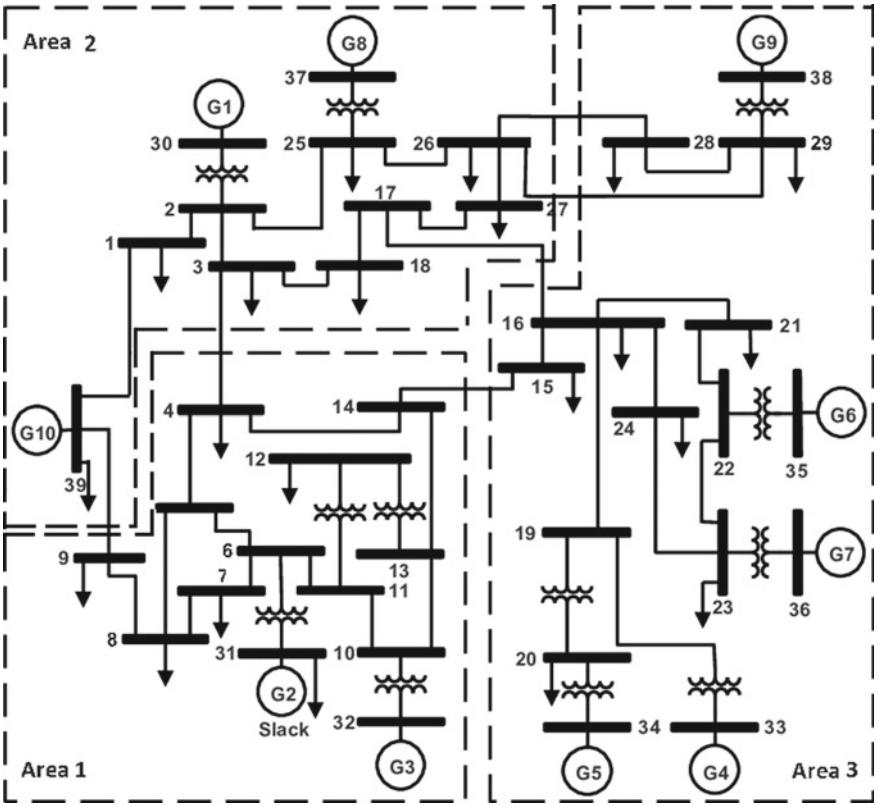



**Fig. 8.3** 39-bus 3 area new England test system (Bevrani 2014)

### *8.4.1 Example 8.1: Single Step Load Change Results*

The attack-resilient control is implemented on a 3 area system with only a single load change. Figure 8.4 shows the actual load and generation changes in the absence of any attack.

It can be seen that the generation follows the load as expected. We next introduce a step attack into the frequency and tie-line as shown in Fig. 8.5.

As soon as the attack is detected, the attack-resilient control-based ACE values are used in the LFC control, and the above values of frequency and tie-line measurements are discarded.

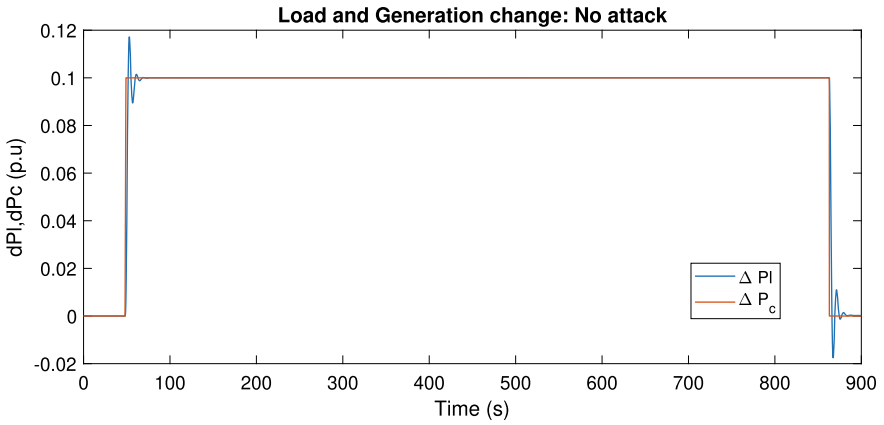



**Fig. 8.4** Single step load and generation change without attack

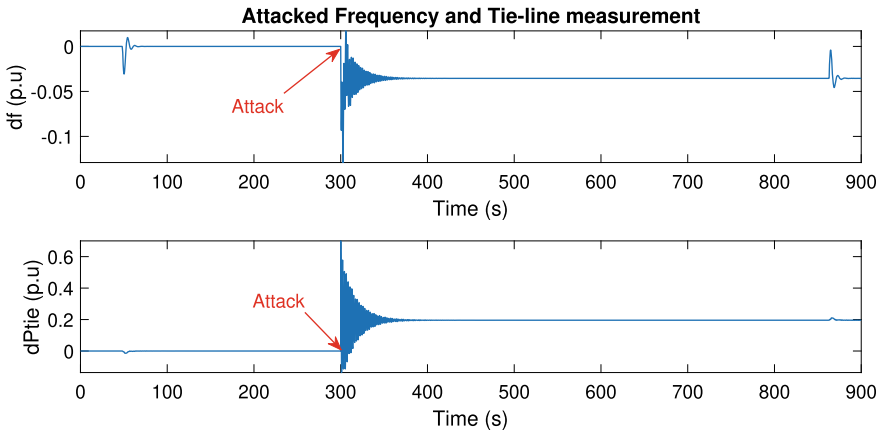



**Fig. 8.5** Attacked frequency and tie-line measurements

The actual and attack-resilient ACE values are shown in Fig. 8.6 The output of the LFC system with and without resilient control is as shown in Fig. 8.7.

It can be seen that the ACE without resilient control shows a variation when there is an attack. When we switch to resilient control, the ACE dies down to zero during the attack period and changes again only when there is a load change.

In Fig. 8.7, the generation follows the load exactly. This is because the forecast and actual load values are the same. However, in practical situations, there will be some difference in the actual and forecast load. This change will be analyzed in the next subsection where we use real load values.

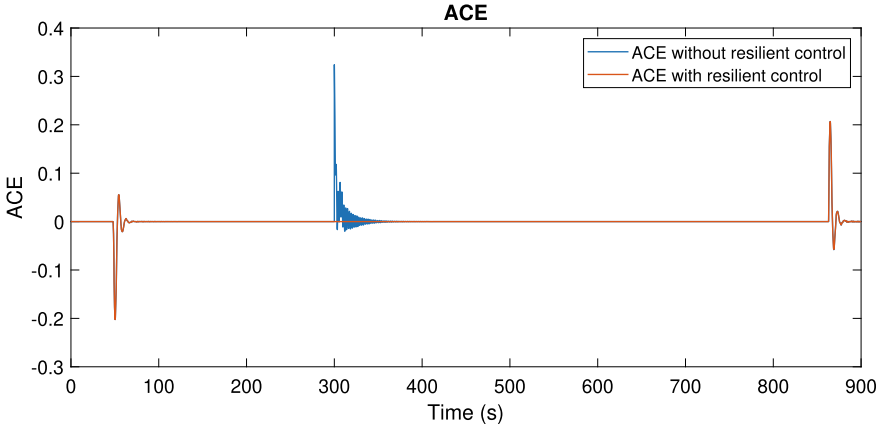



**Fig. 8.6** Actual and resilient ACE values: step load change

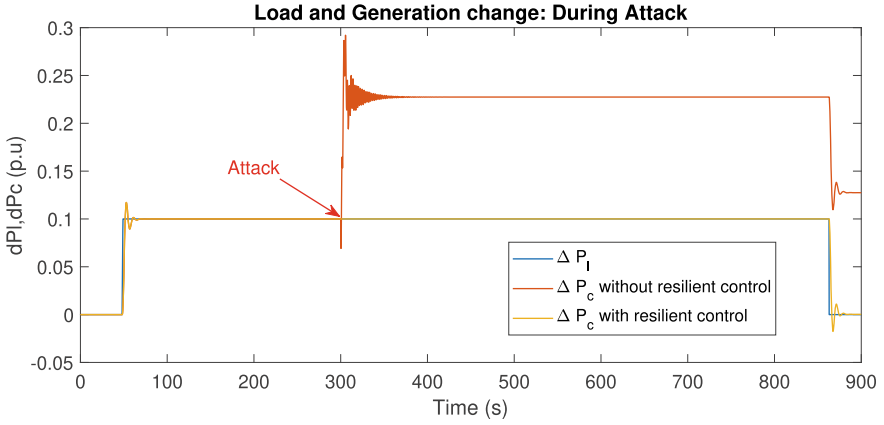



**Fig. 8.7** Attack-resilient control for step load change

### *8.4.2 Example 8.2: New England ISO Load Data Results*

For simulating actual load data, an error value is added to the actual load that is used as the disturbance input. This produces a good model to analyze the actual system behavior. Figure 8.8 shows how the generation exactly follows the load when there is no attack.

We analyze the frequency, ACE, and change in generations during an attack without attack-resilient control and with the attack-resilient control. Figure 8.9 shows the attacked frequency and tie-line measurements.

To study the effect of attack-resilient control on real data values, we use the load forecast data to obtain the ACE values for attack- resilient control, and actual load values are used to test the control algorithm.

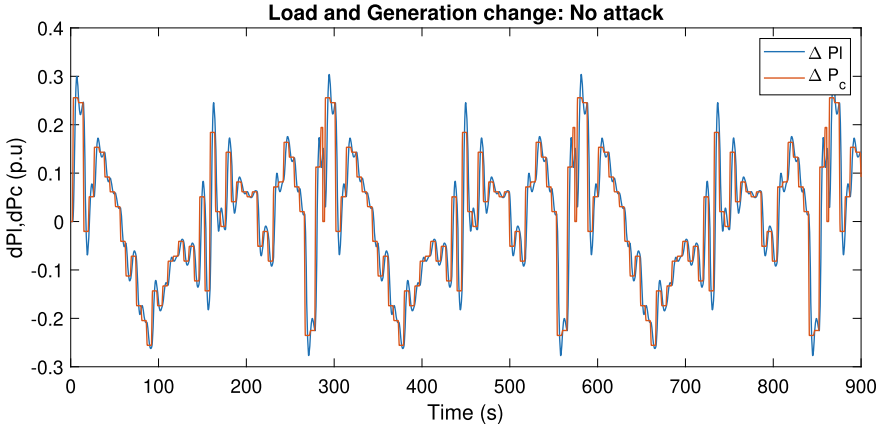



**Fig. 8.8** Load and generation change: no attack

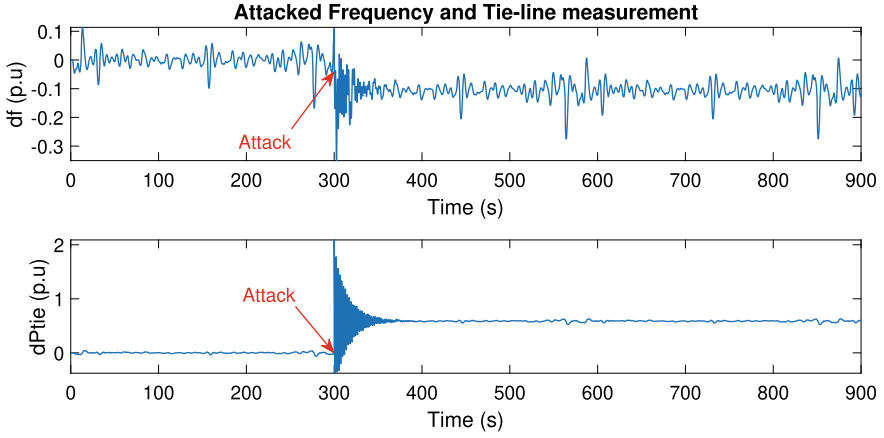



**Fig. 8.9** Attacked measurements for real load condition

As seen in Fig. 8.10, the ACE values are the same with and without control when there is no attack. During attack, the actual ACE values are replaced by the load forecast-based ACE.

As seen in Fig. 8.11, the generation values obtained during attack-resilient control do not exactly follow the load. This is because the forecast values are used to determine the ACE and the generations are according to the load forecasts. However, the difference between the load and generation is very small, and thus the algorithm can be implemented in real systems.

The above attack-resilient control algorithm could fail to provide necessary control action in the event of a fault or a large load variation from the forecast values. Under such emergency or contingency conditions, it is a better choice to bypass the LFC when an attack is detected.

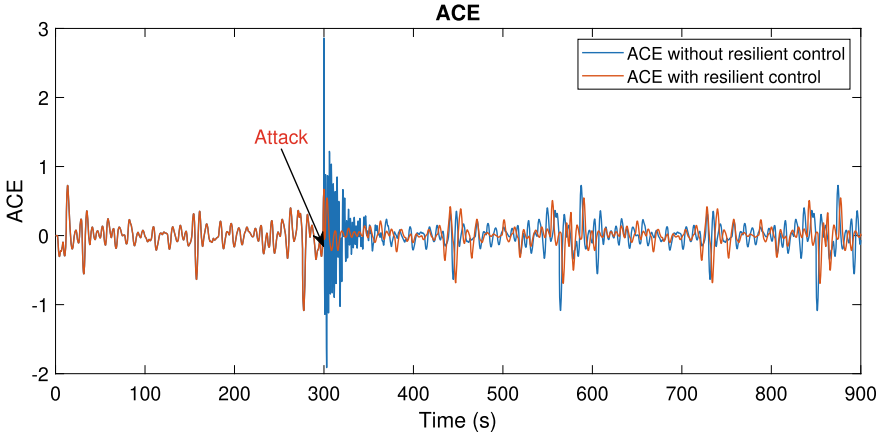



**Fig. 8.10** ACE values with and without attack-resilient control for real load condition

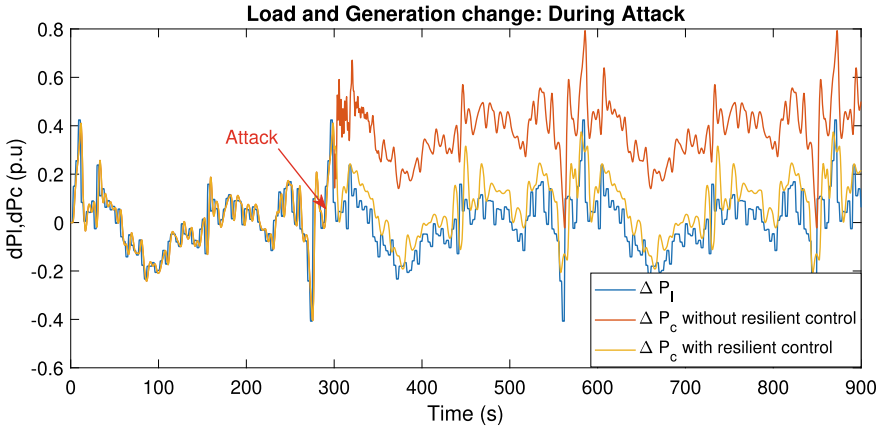



**Fig. 8.11** Attack-resilient control for real load values

## 8.5 IoT-Based Hardware Model

While cyber-security studies and research for power grid systems are important, testing and analysis cannot be performed directly on the system as it could lead to considerable downtimes which is not acceptable in OT systems. However, it is also not possible to analyze various parameters such as communication, computation, and physical dynamics by considering only simulation-based results. Hence, building cyber-security testbeds becomes essential. While implementation of a real-time cyber-physical model of a grid system is highly complex, a hardware-in-loop simulation model can be used to encapsulate the advantages of both simulation and hardware equipment (Tidball 2015; Ashok et al. 2016; Vellaithurai et al. 2017).

In this section, we build a lab setup using both simulation and hardware as shown in Fig. 8.12. The various components of the system are as follows:

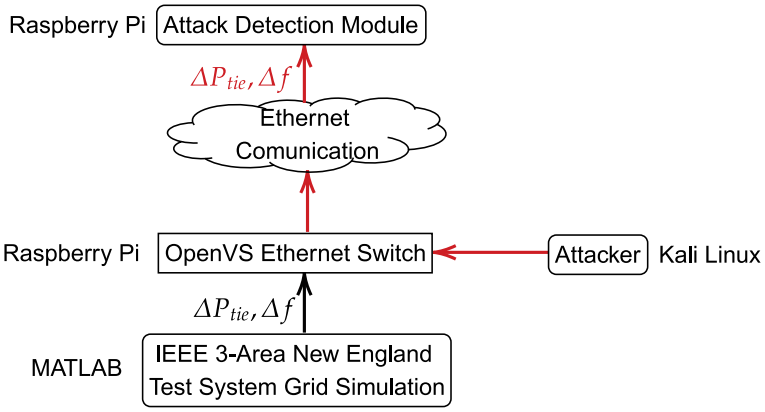



**Fig. 8.12** Lab setup for cyber-security analysis of MA-LFC

**Power Grid Simulation**

The simulation of the 39-Bus 3 area New England Test System which encapsulates the grid system is built using MATLAB on an i5 processor system. A simulation platform is used to model the system since different system models can be implemented and the data from these can be used for further processing. The output of the simulation is the measurement data ($\Delta P_{tie}$, $\Delta f$).

Python socket programming is then used to send the data over the ethernet. The measurements are sent serially over the network by specifying the address of the detection module (Raspberry Pi) as the destination.

**Communication System**

The data from the simulation is then sent over the communication channels to the attack detection module at the control center. This data transfer is achieved using an ethernet network. The ethernet switch is implemented using a Raspberry Pi unit so as to make it a programmable switch. Programmable switches are generally used in power system network so as to route the data correctly and also to provide whitelisting and data security. In the proposed model, the Raspberry Pi is programmed using Open Vswitch (OVS).

The OpenVswitch programming can be performed using Linux scripting. To program the OVS, a bridge is first created and the various ports are then added to the bridge. Further conditions can be specified to route the messages based on various parameters of the message such as source address, destination address, message type, etc. More details about the commands used to program an OVS is available on (Tutorials 2016–2023).

**Attacker System**

Attacker systems are usually implemented using Kali Linux. Kali Linux is a dedicated linux-based operating system that has built-in tools to analyze the network and perform penetration testing. In the proposed lab setup, we use a Kali Linux system as a Man-in-the-Middle attack. The attacker has access to all the DNP3 packets flowing through the switch.

During the attack, the attacker modifies the OVS such that the attacker data is routed to the control center instead of the actual frequency and tie-line data. Thus, the data received at the control center will be false data.

**Attack Detection Module**
The attack detection algorithm is then implemented on a Raspberry Pi unit. The Raspberry Pi receives the frequency and tie-line data over Ethernet and performs the detection algorithm to give out a signal to the control center if there is an attack. Based on the input from the attack detection module, the control center decides on the control input and sends it back to the test system. In the proposed lab setup, we currently do not consider the control input being sent back to the MATLAB simulation.

The hardware setup is as follows:

1. **Power Grid Simulation**: is done on a core i5 Desktop PC running MATLAB 2020a.
2. **Communication System**: Ethernet Communication cables are used for data transfer and the OpenV switch is implemented on a Raspberry Pi-4.
3. **Attacker System**: The attacker is simulated on a Kali Linux system on a core i5 laptop.
4. **Attack Detection Module**: The detection algorithm is implemented on a Raspberry Pi module whose output can be viewed on the connected screen.

All the devices are connected together using the OpenVswitch which is shown in detail along with the connections in Fig. 8.13.

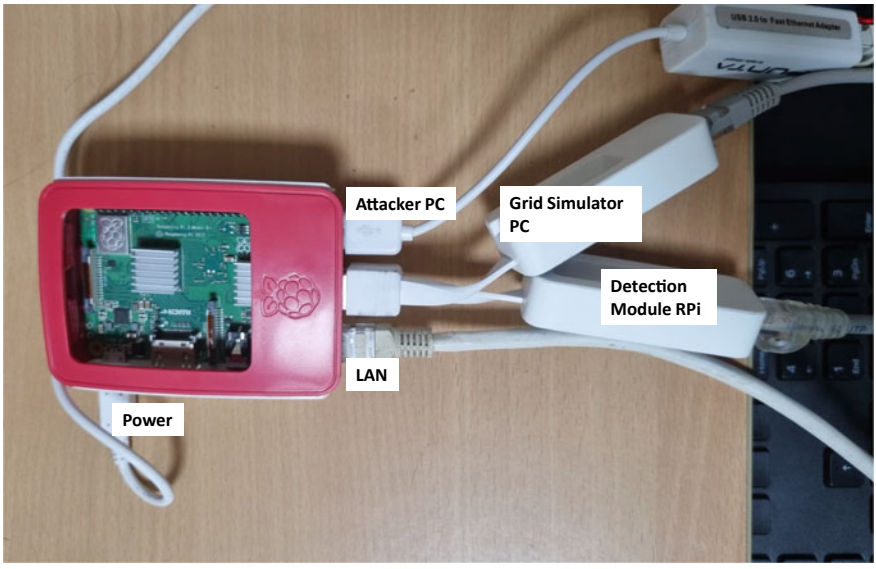



**Fig. 8.13** Detailed view of Open Vswitch connections

The Open Vswitch is connected to the local LAN through the available ethernet port on the Raspberry Pi. To connect to other devices, multiple USB to ethernet converters are used. Each of these ports is then added to a bridged network within the OVS so as to route the data correctly. The attacker modifies the switch in order to implement a False Data Injection through a Man-in-the-Middle attack.

## 8.6 Research Scope

A major scope of research in attack detection is by including the data and models on renewable generation. Renewable penetration into the grid can cause varied dynamics which could lead to failure of the detection and can lead to a high value of False Alarm Rates. Thus, more research is required in this field. Additionally, detection algorithms that combine detection and mitigation can work in better harmony and adapt to system changes leading to a complete cyber-security solution.

### *8.6.1 Research Gap*

- Lack of adaptive detection mechanisms in response to evolving cyber threats and system changes.
- Lack of interpretability and localization of detected attacks.
- Lack of analysis of cyber attacks in LFC under noisy communication networks and detection strategies capable of distinguishing between noise and strategic attacks.
- Knowledge gap between IT experts and power system experts leading to insufficient exploration of packet level data for attack detection in power system control applications.

### *8.6.2 Research Directions*

- Develop detection and control applications that can adapt to system changes. Event-triggered updates, time-based updates, or planned updates can be applied to re-learn the parameters of detection algorithm.
- A combination of clustering and estimation can be leveraged to identify exact locations of the attacks.
- Interpretable machine learning and neural networks can support in identification of attack types and distinguish attacks from system contingencies.
- Algorithms that combine packet and protocol level data with control system signal level data can be used for more accurate attack detection.

## 8.7 Summary

This chapter forms the last part of the cyber-security framework for the MA-LFC system. An attack-resilient control that uses ACE values estimated from forecast loads is used for attack mitigation. The availability of improved load forecast algorithms makes this control algorithm highly effective for real-time applications. The results suggest that the control algorithm can effectively safeguard the grid system from cyber-attacks.

Other control algorithms such as robust control, H-inf control, and model predictive control can be used to improve the attack-resilient control in the event of contingencies, uncertainties, and load drops.

This chapter also explains the development of a lab hardware-based setup that can be used to inject different attacks into the system from an external system, and thus emulate an attacker. The detection algorithm is implemented on a system with low computation and still it shows very good performance. Thus, the theoretical claims that the algorithm has a low computation burden have been practically demonstrated using the hardware setup.

The lab setup can be further enhanced to send back the control signals to the simulation platform to implement a complete loop of the system operations.

## References

Tutorials-open vswitch (2016–2023)

Ameli A, Hooshyar A, Yazdavar AH, El-Saadany EF, Youssef A (2018) Attack detection for load frequency control systems using stochastic unknown input estimators. IEEE Trans Inf Forensics Secur 13(10):2575–2590

Ashok A, Krishnaswamy S, Govindarasu M (2016) Powercyber: a remotely accessible testbed for cyber physical security of the smart grid. 2016 In: IEEE power energy society innovative smart grid technologies conference (ISGT). Minneapolis, MN, USA, pp 1–5

Bevrani H (2014) Robust power system frequency control. Springer

Bi W, Zhang K, Li Y, Yuan K, Wang Y (2019) Detection scheme against cyber-physical attacks on load frequency control based on dynamic characteristics analysis. IEEE Syst J 13(3):2859–2868

Cheng Z, Hu S, Yue D, Dou C, Shen S (2021) Resilient distributed coordination control of multiarea power systems under hybrid attacks. IEEE Trans Syst Man Cybern Syst 1–12

Cheng Z, Yue D, Hu S, Huang C, Dou C, Chen L (2020) Resilient load frequency control design: Dos attacks against additional control loop. Int J Electric Power Energy Syst 115:105496

Liu Y, Chen Y, Li M (2021) Dynamic event-based model predictive load frequency control for power systems under cyber attacks. IEEE Trans Smart Grid 12(1):715–725

Sridhar S, Govindarasu M (2014) Model-based attack detection and mitigation for automatic generation control. IEEE Trans Smart Grid 5(2):580–591

Tidball J (2015) A smart laboratory

Vellaithurai CB, Biswas SS, Srivastava AK (2017) Development and application of a real-time test bed for cyber-physical system. IEEE Syst J 11(4):2192–2203

# Appendix A
# Test Systems Data

## A.1 IEEE 9-Bus System

The IEEE 9-Bus Test system is as shown in Fig. A.1.
The bus and line data are tabulated in Table A.1.
Machine data is given in Table A.2

## A.2 39-Bus New England Test System

The one-line diagram of 39-bus New England Test system indicating the 3 areas is as shown in Fig. A.2.
The line data is as given in Table A.3.
The machine data is given in Table A.4.

## A.3 IEEE 300-Bus System

The IEEE 300-bus system is as shown in Fig. A.3.

**Reference**

Bevrani H (2014) Robust power system frequency control. Springer.



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8

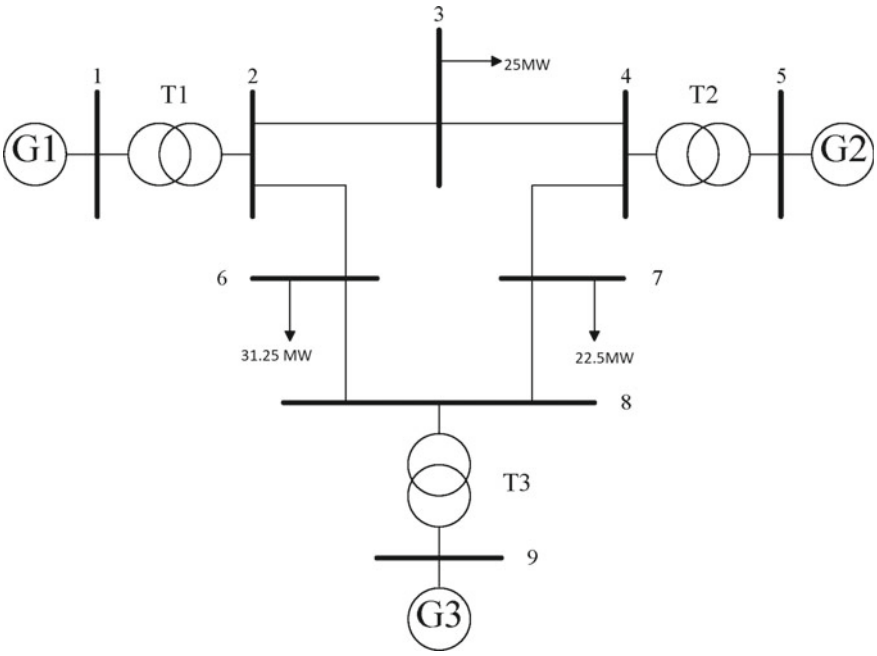



**Fig. A.1** IEEE 9-bus system

**Table A.1** Line data for 9-bus system

| From bus | To bus | Resistance (pu) | Reactance (pu) | Susceptance (pu) | Line rating (MW) |
|---|---|---|---|---|---|
| 1 | 4 | 0 | 0.0576 | 0 | 250 |
| 4 | 5 | 0.017 | 0.092 | 0.158 | 250 |
| 5 | 6 | 0.039 | 0.170 | 0.358 | 150 |
| 3 | 6 | 0 | 0.0586 | 0 | 300 |
| 6 | 7 | 0.0119 | 0.1008 | 0.209 | 150 |
| 7 | 8 | 0.0085 | 0.072 | 0.149 | 250 |
| 8 | 2 | 0 | 0.0625 | 0 | 250 |
| 8 | 9 | 0.032 | 0.161 | 0.306 | 250 |
| 9 | 4 | 0.010 | 0.085 | 0.176 | 250 |

**Table A.2** Generator parameters for 9-bus system (Bevrani 2014)

| Parameters | Generator | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Rate (MW) | 512 | 270 | 125 |
| $\beta_i$ (pu/Hz) | 0.3483 | 0.3473 | 0.3180 |
| $D_i$ (pu) | 2 | 2 | 2 |
| $R_i$ (Hz/pu) | 3.00 | 3.00 | 3.30 |
| $2H_i/f_0$ (pu.s) | 0.105 | 0.165 | 0.191 |
| $Tt_i$ (s) | 0.48 | 0.5 | 0.5 |
| $Tg_i$(s) | 0.08 | 0.06 | 0.07 |
| p.f | 0.4 | 0.4 | 0.2 |
| Ramp rate (MW/min) | 8 | 8 | 4 |

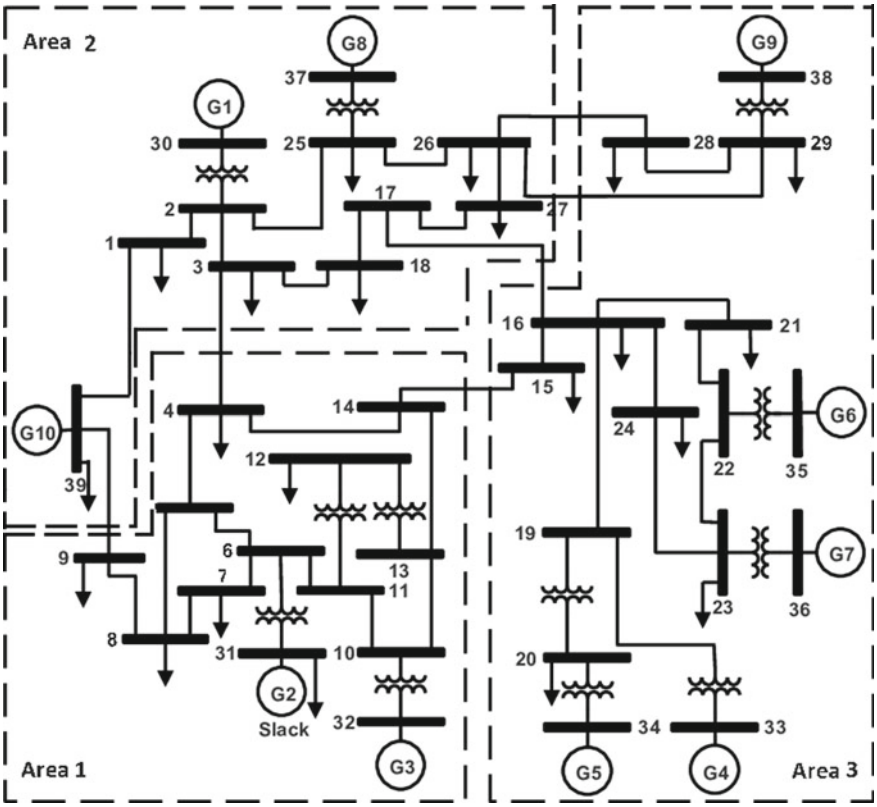



**Fig. A.2** 39-bus 3 area New England test system

**Table A.3** Line data for 39-bus system

| From bus | To bus | Resistance (pu) | Reactance (pu) | Susceptance (pu) | Line rating (MW) |
|---|---|---|---|---|---|
| 1 | 2 | 0.0035 | 0.0411 | 0.6987 | 600 |
| 1 | 39 | 0.001 | 0.025 | 0.75 | 1000 |
| 2 | 3 | 0.0013 | 0.0151 | 0.2572 | 500 |
| 2 | 25 | 0.007 | 0.0086 | 0.146 | 500 |
| 2 | 30 | 0 | 0.0181 | 0 | 900 |
| 3 | 4 | 0.0013 | 0.0213 | 0.2214 | 500 |
| 3 | 18 | 0.0011 | 0.0133 | 0.2138 | 500 |
| 4 | 5 | 0.0008 | 0.0128 | 0.1342 | 600 |
| 4 | 14 | 0.0008 | 0.0129 | 0.1382 | 500 |
| 5 | 6 | 0.0002 | 0.0026 | 0.0434 | 1200 |
| 5 | 8 | 0.0008 | 0.0112 | 0.1476 | 900 |
| 6 | 7 | 0.0006 | 0.0092 | 0.113 | 900 |
| 6 | 11 | 0.0007 | 0.0082 | 0.1389 | 480 |
| 6 | 31 | 0 | 0.025 | 0 | 1800 |
| 7 | 8 | 0.0004 | 0.0046 | 0.078 | 900 |
| 8 | 9 | 0.0023 | 0.0363 | 0.3804 | 900 |
| 9 | 39 | 0.001 | 0.025 | 1.2 | 900 |
| 10 | 11 | 0.0004 | 0.0043 | 0.0729 | 600 |
| 10 | 13 | 0.0004 | 0.0043 | 0.0729 | 600 |
| 10 | 32 | 0 | 0.02 | 0 | 900 |
| 12 | 11 | 0.0016 | 0.0435 | 0 | 500 |
| 12 | 13 | 0.0016 | 0.0435 | 0 | 500 |
| 13 | 14 | 0.0009 | 0.0101 | 0.1723 | 600 |
| 14 | 15 | 0.0018 | 0.0217 | 0.366 | 600 |
| 15 | 16 | 0.0009 | 0.0094 | 0.171 | 600 |
| 16 | 17 | 0.0007 | 0.0089 | 0.1342 | 600 |
| 16 | 19 | 0.0016 | 0.0195 | 0.304 | 600 |
| 16 | 21 | 0.0008 | 0.0135 | 0.2548 | 600 |
| 16 | 24 | 0.0003 | 0.0059 | 0.068 | 600 |
| 17 | 18 | 0.0007 | 0.0082 | 0.1319 | 600 |
| 17 | 27 | 0.0013 | 0.0173 | 0.3216 | 600 |

(continued)

**Table A.3** (continued)

| From bus | To bus | Resistance (pu) | Reactance (pu) | Susceptance (pu) | Line rating (MW) |
|---|---|---|---|---|---|
| 19 | 20 | 0.0007 | 0.0138 | 0 | 900 |
| 19 | 33 | 0.0007 | 0.0142 | 0 | 900 |
| 20 | 34 | 0.0009 | 0.018 | 0 | 900 |
| 21 | 22 | 0.0008 | 0.014 | 0.2565 | 900 |
| 22 | 23 | 0.0006 | 0.0096 | 0.1846 | 600 |
| 22 | 35 | 0 | 0.0143 | 0 | 900 |
| 23 | 24 | 0.0022 | 0.035 | 0.361 | 600 |
| 23 | 36 | 0.0005 | 0.0272 | 0 | 900 |
| 25 | 26 | 0.0032 | 0.0323 | 0.531 | 600 |
| 25 | 37 | 0.0006 | 0.0232 | 0 | 900 |
| 26 | 27 | 0.0014 | 0.0147 | 0.2396 | 600 |
| 26 | 28 | 0.0043 | 0.0474 | 0.7802 | 600 |
| 26 | 29 | 0.0057 | 0.0625 | 1.029 | 600 |
| 28 | 29 | 0.0014 | 0.0151 | 0.249 | 600 |
| 29 | 38 | 0.0008 | 0.0156 | 0 | 1200 |

**Table A.4** Generator parameters for 39-bus system (Bevrani 2014)

| Parameters | Generator | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| MVAbase (1,000 MW) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Rate (MW) | 1000 | 800 | 1000 | 1100 | 900 | 1200 | 850 | 1000 | 1020 |
| $\beta_i$ (pu/Hz) | 0.3483 | 0.3473 | 0.3180 | 0.3827 | 0.3890 | 0.4140 | 0.3692 | 0.3493 | 0.3550 |
| $D_i$ (pu MW/Hz) | 0.015 | 0.014 | 0.015 | 0.016 | 0.014 | 0.014 | 0.015 | 0.016 | 0.015 |
| $R_i$ (Hz/pu) | 3.00 | 3.00 | 3.30 | 2.7273 | 2.6667 | 2.50 | 2.8235 | 3.00 | 2.9412 |
| $2H_i/f_0$ (pu.s) | 0.1677 | 0.120 | 0.200 | 0.2017 | 0.150 | 0.196 | 0.1247 | 0.1667 | 0.187 |
| $Tt_i$ (s) | 0.4 | 0.36 | 0.42 | 0.44 | 0.32 | 0.40 | 0.30 | 0.40 | 0.41 |
| $Tg_i$(s) | 0.08 | 0.06 | 0.07 | 0.06 | 0.06 | 0.08 | 0.07 | 0.07 | 0.08 |
| p.f | 0.4 | 0.4 | 0.2 | 0.6 | 0 | 0.4 | 0 | 0.5 | 0.5 |
| Ramp rate (MW/min) | 8 | 8 | 4 | 12 | 0 | 8 | 0 | 10 | 10 |

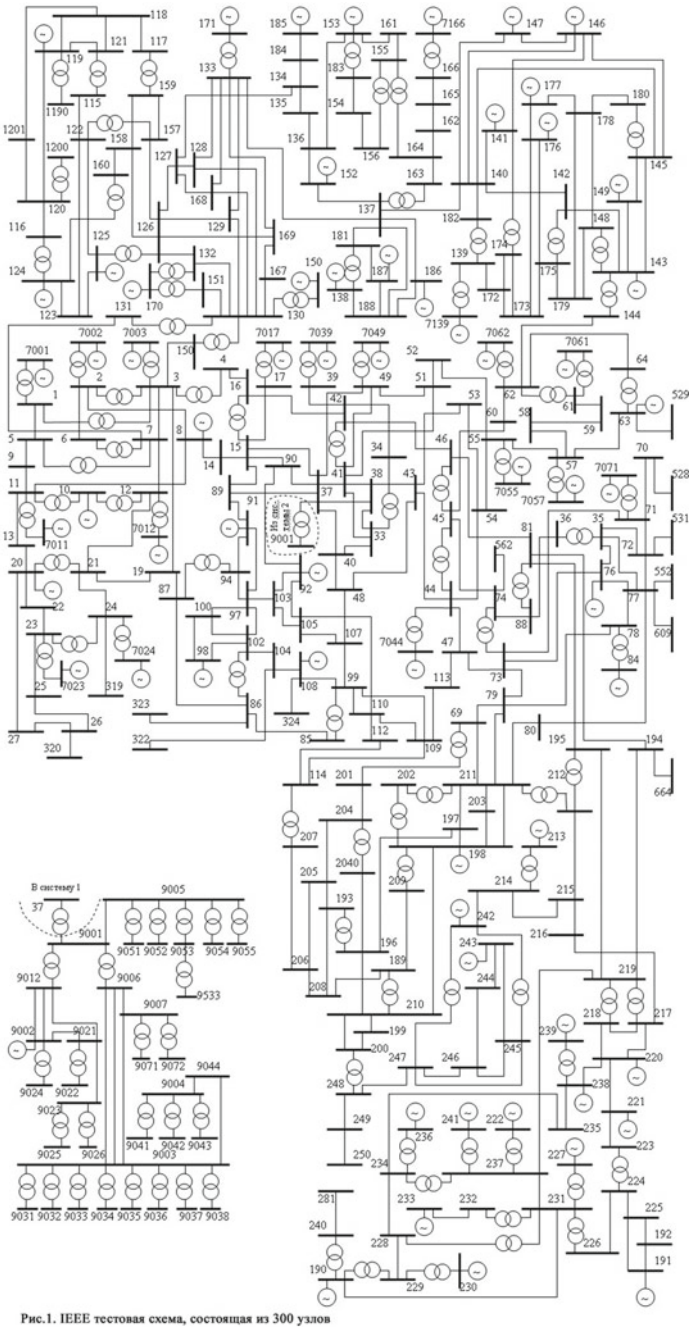



**Fig. A.3** IEEE 300-bus system

# Appendix B
# Detailed Equations for Cascading Outage Model

The cascading outage model considers various equations to create the cascading model. This appendix gives the equations used in the model. The equations are then implemented in MATLAB script and solved using Differential Algebraic Equation analysis.

The Rotor Speed ($\omega_i$)-Swing Equation is given by

$$M\frac{d\omega_i}{dt} = P_{m_i} - P_{g_i} - D\,(\omega_i - 1) \tag{B.1}$$

where $M$ is a machine inertia constant, $D$ is a damping constant, $P_{m_i}$ is the mechanical power input, and $P_{g_i}$ is the generator power output.

The equation for Rotor Angle $\delta_i$ is as given below

$$\frac{d\delta_i(t)}{dt} = 2\pi f_0\,(\omega_i - 1) \tag{B.2}$$

where $f_0$ is the base frequency.

If $X_{d,i}$ and $X'_{d,i}$ are the direct axis generator synchronous and transient reactances, respectively, the salient-pole model reactive power outputs are given by the nonlinear equations,

$$\begin{aligned} P_{g,i} &= \frac{\left|E'_{a,i}\right||V_i|}{X'_{d,i}}\sin\delta_{m,i} + \frac{|V_i|^2}{2}\left(\frac{1}{X_{q,i}} - \frac{1}{X'_{d,i}}\right)\sin 2\delta_{m,i} \\ Q_{g,i} &= \frac{\left|E'_{a,i}\right||V_i|}{X'_{d,i}}\cos\delta_{m,i} + |V_i|^2\left(\frac{\cos^2\delta_{m,i}}{X'_{d,i}} + \frac{\sin^2\delta_{m,i}}{X_{q,i}}\right) \end{aligned} \tag{B.3}$$

For the desired reference voltage $V_{ref}$, and actual terminal voltage $V_t$, the exciter equations are



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8

$$\begin{aligned}\frac{d|E_{\text{fd}}|}{dt} &= \frac{1}{T_E}\left\{K_E \cdot \text{sigm}\left[\left(1-\frac{T_A}{T_B}\right)E_1 \right.\right. \\ &\qquad \left.\left. + \frac{T_A}{T_B}(V_{\text{ref}} - V_t)\right] - E_{\text{fd}}\right\} \\ \frac{d|E_1|}{dt} &= \frac{1}{T_B}\left(V_{\text{ref}} - V_t - E_1\right)\end{aligned} \tag{B.4}$$

where $T_A$, $T_B$, and $K_E$ are the exciter time constants, and sigm$(\cdot)$ is a differentiable sigmoidal function that acts as a limiter between $E_{min}$ and $E_{max}$.

The Differential Algebraic Equations (DAE) can be solved using trapezoidal method. If $t$ is the current time step, and next step is $t + \Delta t$, the $t$, $f(t)$, and $g(t)$ can be calculated for a set of variables $x = x(t)$, $y = y(t)$, $z = z(t)$. The trapezoidal solution method solves the following nonlinear system to obtain $x_+ = x(t_+\Delta t)$ and $y_+ = y(t_+\Delta t)$:

$$\begin{aligned}\mathbf{x}_+ &= \mathbf{x} + \frac{\Delta t}{2}\left[\mathbf{f}(t) + \mathbf{f}(t_+, \mathbf{x}_+, \mathbf{y}_+, \mathbf{z})\right] \\ 0 &= \mathbf{g}(t_+, \mathbf{x}_+, \mathbf{y}_+, \mathbf{z})\end{aligned} \tag{B.5}$$

In the proposed VA, the time step used for solving the DAE is taken as 0.005 s

# Appendix C
# Information Theory and Hypothesis Testing

Information theory is a field of science that establishes a link between two different kinds of quantities. The ideal or limiting value of a specific parameter, such as the convergence rate of error probabilities, is known as an operational quantity. A measure of information such as entropy, divergence, and mutual information is the other. It should be noted that the latter's definition is more ambiguous than the former's, and that the latter's meaning is typically elucidated by relating it to the former.

The information spectrum method was initially discussed and published by Han and Verdu (1993), Han (2003). Information theory can be used to determine the relationship between the false positive rate and threshold. This appendix discusses the basic definitions used in the Hypothesis testing-based detection in Chap. 6.

## C.1 Hoeffding Test

Hoeffding test (Hoeffding 1965) is a composite hypothesis test where the test has only partial access to the distributions $P$ and $Q$.

**Definition**: In the Hoeffding test, the null hypothesis $P$ is accepted if the Kullback-Leibler (KL) divergence between the type $t_{Z^n}$ (the empirical distribution) of the observations $Z_n = (Z_1, \ldots, Z_n)$ and $P$ is below some threshold $c$. Otherwise, the alternative hypothesis is accepted. Mathematically,

$$if\, D(t_{Z^n}||P) \leq c, \text{ then accept } H_0; \text{ otherwise accept } H_1 \tag{C.1}$$

An improved form of Hoeffding test is proposed in the thesis that can be used for the proposed spectral analysis-based detection.



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8

## C.2 Neyman-Pearson Theorem

The Neyman-Pearson lemma is a part of the Neyman-Pearson theory of statistical testing (Neyman and Pearson 1933). Only one hypothesis was proposed by the earlier Fisherian theory of significance testing. The Neyman-Pearsonian form of statistical testing enables examining the two sorts of errors by introducing a competing hypothesis (Wald 1942).

Consider a test with hypotheses $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$, where the probability density function (or probability mass function) is $\rho(x \mid \theta_i)$ for $i = 0, 1$.

For any hypothesis test with rejection set $R$, and any $\alpha \in [0, 1]$, we say that it satisfies condition $P_\alpha$ iff

1. $\alpha = Pr_{\theta_0}(X \in R)\ \alpha = Pr_{\theta_0}(X \in R)$
   That is, the test has size $\alpha$.
2. $\exists \eta \geq 0$ such that

   $$x \in R \setminus A \implies \rho(x \mid \theta_1) > \eta \rho(x \mid \theta_0)$$
   $$x \in R^c \setminus A \implies \rho(x \mid \theta_1) < \eta \rho(x \mid \theta_0)$$
   where $A$ is a set ignorable in both $\theta_0$ and $\theta_1$ cases: $Pr_{\theta_0}(X \in A) = Pr_{\theta_1}(X \in A) = 0$

3. That is, we have a strict likelihood ratio test, except on an ignorable subset.

For any $\alpha \in [0, 1]$, let the set of level $\alpha$ tests be the set of all hypothesis tests with size at most $\alpha$. That is, letting its rejection set be $R$, we have $Pr_{\theta_0}(X \in R) \leq \alpha$.

### *C.2.1 Neyman-Pearson Lemma*

**Existence**:

If a hypothesis test satisfies $P_\alpha$ condition, then it is a *uniformly most powerful* (UMP) test in the set of level $\alpha$ tests.

**Uniqueness**:

If there exists a hypothesis test $R_{NP}$ that satisfies $P_\alpha$ condition, with $\eta > 0$ , then every UMP test $R$ in the set of level $\alpha$ tests satisfies $P_\alpha$ condition with the same $\eta$ .

**References**

Han T, Verdu S (1993) Approximation theory of output statistics. IEEE Trans Inf Theory 39(3):752–772.
Han TS (2003) Information-spectrum methods in information theory. Springer.
Hoeffding W (1965) Asymptotically optimal tests for multinomial distributions. Ann Math Stat 36(2):369–401.
Neyman J, Pearson ES (1933) On the problem of the most efficient tests of statistical hypotheses. Philos Trans R Soc Lond Ser A Contain Papers Math Phys Character 231(694–706):289–337.

Wald A (1942) Chapter II: The neyman-pearson theory of testing a statistical hypothesis. In: On the principles of statistical inference, vol 1. University of Notre Dame, pp 10–21.

# Appendix D
# Proofs of Theorems

## D.1 Theorem 3.1

**Theorem D.1** *For the system (3.2) with static monitor $\phi$ as in (3.5), an attack $u_a$ will be undetected if $Du_a(k) \in \{0,\ Im(C)\}$.*

***Proof*** The residue when there is no attack is given as

$$r = y(k) - Cx(k) = y(k) - CC^{\dagger}y(k) = (I - CC^{\dagger})y(k) \tag{D.1}$$

where $C^{\dagger}$ is the pseudo-inverse of $C$. The residue under an attack is obtained as

$$\begin{aligned} r_a &= y_a(k) - Cx_a(k) = (I - CC^{\dagger})y_a(k) \\ &= (I - CC^{\dagger})(y(k) + D_a u_a(k)) \\ &= r + (I - CC^{\dagger})D_a u_a(k) \end{aligned} \tag{D.2}$$

The residue $r_a = r$ iff, $(I - CC^{\dagger})D_a u_a(k)$ vanishes. This is possible only when $D_a u_a(k) \in \{0, Im(C)\}$. $D_a u_a(k) = 0$ is equivalent to a no attack condition. Thus, $D_a u_a \in Im(C)$, i.e., the attack should be in the Image space of $C$.

Thus the proof follows. □

## D.2 Theorem 6.1

**Theorem D.2** *Let $\boldsymbol{U}'$ be a linear transformation from $\mathbb{R}^L \longrightarrow \mathbb{R}^r$ .Then, the norm of the projection of any arbitrary vector $z \in \mathbb{R}^L$ onto the subspace $S^r$ is equivalent to the norm of the transformed vector $\boldsymbol{U}'z \in \mathbb{R}^r$.*



A. Sreejith and K. Shanti Swarup, *Cyber-Security for Smart Grid Control*, Transactions on Computer Systems and Networks, https://doi.org/10.1007/978-981-97-1302-8

***Proof*** **To Prove**: $||\mathbf{P}z||^2 = ||\mathbf{U}'z||^2$

$U'$ is a linear transformation from $\mathbb{R}^L \longrightarrow \mathbb{R}^r$

From (14), $\mathbf{P} = \mathbf{U}\mathbf{U}'$ Thus, the norm of projection of any vector $z$ onto the signal subspace is given by

$$\begin{aligned} ||\mathbf{P}z||^2 &= ||\mathbf{U}\mathbf{U}'z||^2 \\ &= (\mathbf{U}\mathbf{U}'z)'\mathbf{U}\mathbf{U}'z \\ &= z'(\mathbf{U}\mathbf{U}')'\mathbf{U}\mathbf{U}'z \end{aligned} \tag{D.3}$$

Since the columns of $\mathbf{U}$ are linearly independent and of unit size $\mathbf{U}$,

$$\mathbf{U}'\mathbf{U} = \mathbf{I} \tag{D.4}$$

On substituting (D.4) in (D.3),

$$\begin{aligned} ||\mathbf{P}z||^2 &= z'\mathbf{U}\mathbf{U}'z \\ &= (\mathbf{U}'z)'\mathbf{U}'z \\ &= ||\mathbf{U}'z||^2 \end{aligned} \tag{D.5}$$

Thus $||\mathbf{P}z||^2 = ||\mathbf{U}'z||^2$. Hence the proof follows. □

## D.3 Theorem 6.2

**Theorem D.3** *The SSA Hoeffding Test satisfied the Neyman-Pearson lemma*

**Definition D.1** (*Neyman-Pearson Lemma*) Consider a binary hypothesis test and the distance measure:

$$d(x) = ||c - Px|| \underset{H_0}{\overset{H_1}{\gtrless}} \tau \tag{D.6}$$

with a probability of false alarm given by

$$P_{FA} = P(d(x) \geq \tau | H_0) = \beta \tag{D.7}$$

There does not exist another test with $P_{FA} = \beta$ and a detection problem larger than $P(d(x) \geq \tau | Ho)$. That is, the SSA-HT is the most powerful test with $PF_A = \beta$.

***Proof*** The region where the SSA-HT decides $H_1$ is

$$\mathcal{R}_{SSA} = x : ||c - Px|| \geq \tau \tag{D.8}$$

Let $\mathcal{R}_T$ denote the region where some other test describes $H_1$. Define for any region $\mathcal{R}$

$$P_i(\mathcal{R}) = \int_{\mathcal{R}} P_i(x)dx \tag{D.9}$$

which is the probability of $x \in \mathcal{R}$ under hypothesis $H_i$. By our assumption above, both tests have $P_{FA} = \beta$:

$$P_0(\mathcal{R}_{SSA}) = P_0(\mathcal{R}_T) = \beta \tag{D.10}$$

The probability regions are such that

$$\begin{aligned} P_i(\mathcal{R}_{SSA}) &= P_i(\mathcal{R}_{SSA} \bigcap \mathcal{R}_T) + P_i(\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T) \\ P_i(\mathcal{R}_T) &= P_i(\mathcal{R}_{SSA} \bigcap \mathcal{R}_T) + P_i(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA}) \end{aligned} \tag{D.11}$$

From (D.10) and (D.11),

$$P_0(\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T) = P_0(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA}) \tag{D.12}$$

Now, for the alternate hypothesis,

$$\begin{aligned} P_1(\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T) &= \int_{\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T} ||c - Px||dx \\ &\geq \tau \int_{\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T} dx \\ &= \tau P_0(\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T) \\ &= \tau P_0(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA}) \\ &= \tau \int_{P_0(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA})} dx \\ &\geq \int_{P_0(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA})} ||c - Px||dx \\ &= P_1(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA}) \end{aligned} \tag{D.13}$$

From (D.13),

$$P_1(\mathcal{R}_{SSA} \bigcap \bar{\mathcal{R}}_T) \geq P_1(\mathcal{R}_T \bigcap \bar{\mathcal{R}}_{SSA}) \tag{D.14}$$

Thus, from (D.13) we see that as $\tau$ increases, $\mathcal{R}_{SSA}$ decreases, and hence the false detection probability $P_{FA}$ decreases. In other words, if $\tau_1 \geq \tau_2$, then $\mathcal{R}_{SSA}(\tau_1) \bigcap \mathcal{R}_{SSA}(\tau_2)$, and hence $\beta_1 \leq \beta_2$. □