Title
Smart Grid Cyber and Physical Security

Permalink
https://escholarship.org/uc/item/03t3b2m6

Author
Amini, Sajjad

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Smart Grid Cyber and Physical Security

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Sajjad Amini

September 2017

Dissertation Committee:

Dr. Hamed Mohsenian-Rad, Chairperson
Dr. Fabio Pasqualetti
Dr. Amit Roy-Chowdhury

The Dissertation of Sajjad Amini is approved:

____

____

____

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor Dr. Hamed Mohsenian-Rad, whose collective support, guidance, and expertise have made this dissertation a reality. Many thanks to Dr. Fabio Pasqualetti and Dr. Amit Roy-Chowdhury for their great feedback and being in my thesis committee. Thank you all for sharing your knowledge and passion with me during my Phd program.

I am also thankful of all my lab mates in Smart Grid research Lab. Without their help and support, I couldn't complete my research. I had a wonderful time by working beside you during these four years.

I also extend my gratitude to my family. I specifically, thank my beloved parents for their unconditional kindness, love, patience, and support throughout my life. They have always been and continue to be my shelter and comfort in this life and especially during my academic career. I also thank my wonderful sisters for their soothing words and endless encouragement which comforted me during many hard moments along the way.

And finally, I am really lucky that have awesome friends. They are not only a friend but like a family member to me. They stood beside me in any situations. Thanks to all of you as you make the life easier for me.

This work was supported in part by three NSF grants. The content of this thesis is a reprint of the material that are appeared in the following publications:

- S. Amini, F. Pasqualetti, M. Abbaszadeh, and H. Mohsenian-Rad, "Hierarchical Location Identification of Destabilizing Faults and Attacks in Power Systems: A Frequency-Domain Approach", *IEEE Trans. Smart Grid*, under second revision.

- S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic Load Altering Attacks Against Power System Stability: Attack Models and Protection Schemes", *IEEE*

Trans. Smart Grid, vol. PP, no. 99, Oct. 2016.

- S. Amini, F. Pasqualetti, M. Abbaszadeh, and H. Mohsenian-Rad, "Hierarchical Anomaly Identification in Power Systems Using Synchrophasor Data", *accepted for publication in InfoSymbiotics/DDDAS*, Cambridge, MA, USA, Aug. 2017.

- M. Izbicki, S. Amini, C. Shelton, and H. Mohsenian-Rad, "Identification of Destabilizing Attacks in Power Systems", *in Proc. of IEEE American Control Conference (ACC)*, Seattle, WA, USA, May. 2017.

- S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Detecting Dynamic Load Altering Attacks: A Data-Driven Time-Frequency Analysis", *in Proc. of IEEE SmartGridComm'15 Symposium*, Miami, FL, USA, Nov. 2015.

- S. Amini, H. Mohsenian-Rad, and F. Pasqualetti, "Dynamic Load Altering Attacks in Smart Grid", *in Proc. of IEEE PES Conference on Innovative Smart Grid Technologies (ISGT)*, Washington, D.C, USA, Feb. 2015.

To my beloved parents for all of their patience, kindness, and support.

ABSTRACT OF THE DISSERTATION

Smart Grid Cyber and Physical Security

by

Sajjad Amini

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, September 2017
Dr. Hamed Mohsenian-Rad, Chairperson

Anomalies in the form of natural faults or malicious attacks can affect the *dynamics* of power systems. They can be physical or cyber-physical, and can affect the generation side or the load side. Most changes in power system dynamics that are caused by faults and attacks are damped and do not cause any major harm. However, some faults and attacks may make the system *unstable*. The focus in this thesis is on such *destabilizing* faults and attacks. In a destabilizing attack against a power system, the adversary hacks into generators or load control mechanisms to insert positive feedback into the power system dynamics. In this thesis, Dynamic Load Altering Attack is introduced as a new class of cyber-physical destabilizing attacks against smart grid demand response programs and its fundamental characteristics are investigated.

It is crucial to detect presence of anomaly in power system and identify the location(s) of the affected generators and/or loads. In this thesis, the focus is on the problem of data-driven anomaly detection in power systems from measurement data provided by Phasor Measurement Units and without knowledge of the power system dynamics. It is shown that a destabilizing anomaly is detectable through a frequency-domain analysis of measurements. As for the location identification problem, an optimization-based approach in frequency domain is proposed to identify the unknown location(s) of the destabilizing faults and attacks in power systems. The proposed approach does not require prior knowledge about the number of affected location(s). It is fast and computationally more efficient than its time-domain counterparts. Importantly, it is well-suited to be implemented in a hierarchical fashion, with applications such as in Wide Area Monitoring Systems. It is also observed in this thesis that destabilizing anomalies can be modeled as a reparameterization of the power system's dynamical model. Therefore, an identification method that uses the unscented Kalman filter to jointly estimate both the system states and parameters of the anomaly is developed. A low-rank modification to the Kalman filter is also proposed that improves computational efficiency while maintaining the identification accuracy.

Finally, a protection and mitigation scheme is designed to protect vulnerable loads against destabilizing anomalies by formulating and solving a non-convex pole-placement optimization problem.

Various case studies are presented in this thesis to assess performance of the proposed detection, identification, and protection approaches in standard IEEE 9 and 39 bus test systems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The development of distributed intelligence technologies have introduced new opportunities to enhance efficiency and reliability of power grid. However, if these technologies are not accompanied with appropriate security enforcements, they may also create new vulnerabilities in power networks, leaving them open to a wide range of cyber-physical anomalies [52, 74, 76].

Anomalies in the form of natural faults or malicious attacks can affect the *dynamics* of power systems. They can be physical or cyber-physical, and can affect the generation side or the load side. Most changes in power system dynamics that are caused by faults and attacks are damped and do not cause any major harm. However, some faults and attacks may make the system *unstable*. The focus in this thesis is on such *destabilizing* faults and attacks.

1.1 Motivation

Power system destabilization is the result of creating *oscillations* or *positive feedback* within the power system. This can happen, in particular, due to natural faults or intentional attacks at power system inputs, i.e., power generation levels or power consumption levels, e.g., see [6,26,27,45,63,75]. Therefore, there is a need to devise methods to not only protect power systems against such faults or attacks, i.e., take preventive actions, but also detect and identify the fault/attack location(s) in order to take timely diagnostics and corrective actions. In this regard, the focus in this thesis is to develop accurate and computationally efficient methods to detect and identify the location(s) of destabilizing faults and attacks in power systems. Also, designing protection schemes that assure power system stability in presence of such destabilizing anomalies.

1.2 Related Work

The literature related to destabilizing faults and attacks in power systems can be divided into *protection*, *detection*, *identification*, and *mitigation*. Different methods have been developed to protect power systems against destabilizing faults/attacks, e.g., in [6,38, 58]. For example, in [58], a protection and control system mechanism is designed against frequency instability. In [38], a measurement-based online load identification approach is proposed to assess the margin of voltage instability in order to prevent voltage collapse. In this thesis, a protection scheme for protecting vulnerable portion of the loads is designed by formulating and solving a non-convex pole-placement optimization problem.

Detection and mitigation are also studied in many papers, e.g., in [5,31,64]. For example, an attack-mitigation model, based on a game-theoretic analysis, is proposed in [64] to effectively reduce the impact of attack and to maintaine physical stability of the power system. Also, in [31], an algorithm is proposed to automatically detect the fast separation of phase angles among the critical areas in the power system by using synchrophasor data, and by triggering suitable control actions. While most approaches use model-based techniques, in this thesis we focus on the problem of anomaly detection in power systems from power grid measurement data and without knowledge of the power system dynamics. Specifically, we use data of *smart meters* in this thesis. Smart meters are advanced measuring equipment that are used to measure electrical energy consumption at much higher time-resolutions than conventional meters [18]. They are also capable of two-way communications with utility companies. Currently, there are over 45 million smart meters installed in the U.S. that generate more than one billion data points every day [1]. Our results in this thesis show that it is indeed possible to detect destabilizing anomalies via purely data-driven approaches. Specifically, we apply the fast Fourier transform (FFT) to system measurements. It detects frequencies that a destabilizing anomalies adds that are not present during normal operation. The presence of such new frequencies beyond certain pre-specified magnitude thresholds indicate the presence of a destabilizing anomaly.

Location identification problem is addressed, e.g., in [79] using pattern recognition, in [31] using Kalman Filters, in [66] using observer design, and in [54] using state fault diagnosis matrix. In this thesis, we identify destabilizing attacks against power systems by only monitoring state variables. That is, we devise a method that examines the state-variable data from power system sensors such as phasor measurement units (PMUs), to indicate at which exact power system buses (i.e., nodes) the load and/or generation are compromised. The proposed location identification approach operates in *frequency-domain* and is *customized* to work well against a class of destabilizing faults/attacks in power systems, whether in generation or load side. It has several advantages over the existing methods that operate in time-domain. Additionally, we use the unscented Kalman filter (UKF) to perform dual state estimation to estimate fault/ attack parameters. We then identify the location(s) through a proper thresholding mechanism. In terms of the methodology used in this thesis, the UKF has been used before for power system problems, e.g., to estimate the rotor angle and speed in synchronous generators [21]. A more recent analysis estimates the parameters of the motor controller and bus loads [22]. However, no joint estimation is used, and the system is not under fault/attack. Note that, our results show that the current power system monitoring systems would require calculation of a large Jacobian matrix if one wants to apply the UKF without modification for the purpose of identifying destabilizing anomalies. In this regard, the current study is related also to a thread of work, such as in [57, 59], that similarly have to deal with the computational issues that arise when applying the UKF to power systems.

1.3 Dynamic Load Altering Attack

In this thesis, we introduce a new type of destabilizing anomalies that target the *consumption* sector of the power grid. Specifically, we are concerned with attacks that seek to compromise the demand response (DR) and demand side management (DSM) programs. DR programs are used by utilities to control the load at the user side of the meter in response to changes in grid conditions [61]. In a related field, DSM techniques seek to exploit the load flexibility in different load sectors, e.g., by using automated energy consumption scheduling [65].

An important class of cyber-physical attacks against DR and DSM systems is load altering attack (LAA) [49]. LAA attempts to control and change a group of remotely accessible but unsecured controllable loads in order to damage the grid through circuit overflow or other mechanisms. There is a variety of load types that are potentially vulnerable to LAAs, e.g., remotely controllable loads [34], loads that automatically respond to price or Direct Load Control (DLC) command signals [48,50,78], and frequency-responsive loads [51, 81]. Some of the recent studies that address modeling, detection, and prevention of LAAs include [36,39,43].

So far, the focus in the LAA literature has been mainly on *static* load altering attacks, where the attack is concerned with changing the volume of certain vulnerable loads, in particular in an *abrupt* fashion. In contrast, in this thesis, we address *Dynamic* Load Altering Attacks (D-LAA), where we are concerned with not only the amount of the change in the compromised load but also the *trajectory over time* at which the load is changed. Unlike in [36, 39, 43, 49], the analysis in this thesis is based on power system dynamics. Accordingly, we use feedback control theory as the main analytical tool to model or prevent the attack. In this regard, we take into account not only the cyber security challenges but also the physics of the power system.

This study complements and merges two generally independent lines of research in the literature. First, it benefits the recent efforts in designing efficient and practical demand response and demand side management programs [34, 48–51, 61, 65, 78, 81] by increasing awareness about potential vulnerabilities in these programs, not only to consumers, but also to grid as a whole. Second, it also adds to the existing results on control-theoretic study of cyber-physical attacks, c.f. [14, 47, 53, 54].

1.3.1 Attack Classification

A D-LAA can be *open-loop* or *closed-loop*. In an open-loop D-LAA, see Fig. 1.1(a), the attacker tends to manipulate some vulnerable load *without* monitoring the grid conditions in real-time or monitoring the impact that its load manipulation may cause on the power grid while the attack is being implemented. Accordingly, an open-loop D-LAA relies on some historical data that it may collect prior the attack to impose a *pre-programmed* trajectory to the compromised load. In contrast, in a closed-loop D-LAA, the attacker constantly monitors the grid conditions, e.g., through the attacker's installed sensors or via hacking into an existing power system monitoring infrastructure, such that it can control the load trajectory at the victim load bus(es) based on the grid operating conditions. An adversary can conduct a successful D-LAA only if it compromises sufficient amount of vulnerable loads. That is, D-LAA is meaningful only if there is enough flexible and vulnerable (not secured) load to potentially compromise.

The feedback in a closed-loop D-LAA can be based on different types of power grid measurements. For example, the grid conditions can be monitored by measuring voltage magnitude or frequency, aiming for various malicious goals.

The D-LAAs can be classified also based on their scope. Specifically, D-LAAs can

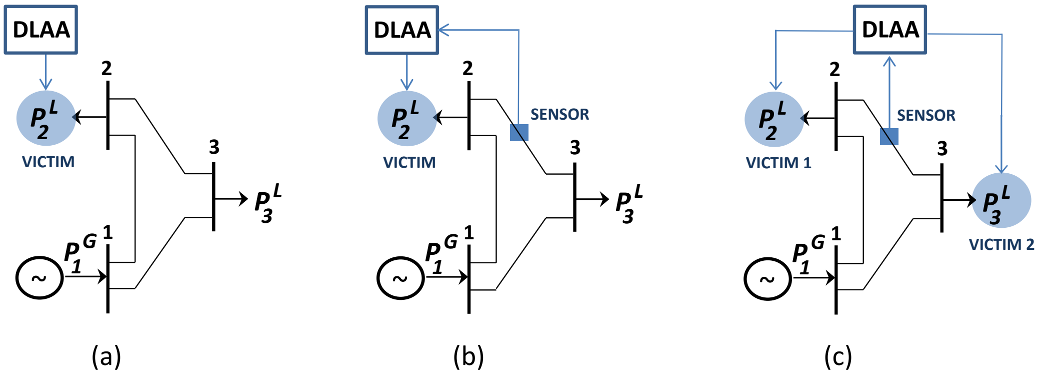


Figure 1.1: Three examples of dynamic load altering attacks: a) open-loop D-LAA, b) single-point closed-loop D-LAA, c) multi-point closed-loop D-LAA.

be *single-point* or *multi-point*. In a single-point D-LAA, the attacker seeks to compromise the vulnerable load at *one* victim load bus. In a multi-point D-LAA, the attacker seeks to compromise a group of vulnerable loads at *several* victim load buses. The vulnerable loads at different load buses are compromised in a *coordinated* fashion. Examples of single-point and multi-point closed-loop D-LAAs are shown in Figs. 1.1(b) and (c), respectively.

Finally, one can classify D-LAAs also based on the type of controller being used in order to manipulate the control variable which is load consumption of victim bus(es), whether through a feed-forward controller in case of an open-loop attack or a feedback controller in case of a closed-loop attack. For example, if the D-LAA is closed-loop, then the attacker may use a bang-bang, P, PI, or PID controller [15], or any other more complex feedback control system mechanism.

1.3.2 Attack Adverse Impacts

Load altering attacks may seek to cause different adverse impacts. For example, a static load altering attack may involve abruptly increasing the load at the most crucial locations in the grid in order to cause *circuit overflow* on distribution or transmission lines that can cause significant damage to the utility company and/or user equipment. Such attacks may also seek to disturb the balance between power supply and demand during peak-load hours. Please refer to [49] for more details about the possible impacts of static load altering attacks.

As for dynamic load altering attacks, the attack objective depends on the type of attack. For example, for a closed-loop D-LAA where the feedback is based on power grid frequency, the attack may seek to deviate the frequency from its nominal value. Note that, an entire interconnected power grid operates at or around a nominal frequency. For example, the nominal frequency in North America is 60 Hz and regional transmission system operators are required to maintain and stabilize frequency at or very closely around this level. Accordingly, a D-LAA may try to damage the grid by *destabilizing* the frequency away from its nominal value.

For a D-LAA against power system stability, an attack may be considered *successful* once it trips one or more over/under frequency relays, c.f. [70], e.g., to force at least one generator go offline, causing a major disturbance to the normal operation of power grid. Such disturbance can potentially trigger *ripple-effects* across the interconnected power system. In fact, due to the connectivity of the grid, small localized perturbations can reach far away regions, and in a disruptive fashion. See for instance the *Nature* paper in [60] for a characterization of certain cascading effects across interconnected networks. Alternatively, if the size of the compromised load is small, it is also possible that triggering the relays and protection systems rather confines the compromised load area, avoiding the attack to spread out to other regions of the power grid, c.f. [68]. But even in that case, the attack is considered successful because it cuts off service for a subset of loads, even though the impact is not catastrophic as in case of an attack with cascade effects.

1.3.3 Closed-loop Attack Implementation

In this thesis, we are interested in closed-loop D-LAAs because they can potentially affect power system *stability*. We assume that the attack feedback based on measuring power grid *frequency*. This setup is of practical importance also due to its link to the concept of frequency-responsive loads [51, 81]. Note that, if a frequency-responsive load is compromised, then power system frequency is already available to the attacker through local measurements. The frequency sensor can be either co-located with the victim load bus, or it can be placed at some other bus but on the same interconnected network. We refer to the bus where the frequency sensor is located as the *sensor bus s*. While D-LAAs take place at the customer and distribution level, their impact is understood only when the system dynamics are studied at the transmission level. Because it is at the transmission level where the area frequency is affected due to an aggregate impact of compromised loads. Nevertheless, the adversary does *not* need access to the transmission-level SCADA/EMS system to implement the attack. All that he/she needs is to hack into the remote load control systems that often exist in demand response programs to adjust the power consumption trajectory.

To implement a D-LAA, the adversary must undergo two major tasks: 1) changing load, and 2) sensing feedback.

Changing Load

The adversary must alter the energy consumption of target vulnerable loads by breaking into the smart grid communications, monitoring, or control infrastructure. This can be done in different ways depending on the type of attack, type of load, or type of the communications infrastructure. In particular, an attack may target compromising *price signals* in price-based demand response programs or *command signals* in Direct Load Control (DLC) programs [16, 33, 78]. For example, the communications infrastructure vulnerability in price-based demand response is discussed in details in [30, 77]. Compromising the command signals in DLC programs is also directly related to D-LAAs, because DLC programs allow remote and direct access to and control over the load without the need to bypass an intermediate or local load control mechanism. Fig. 1.2 shows how an adversary may generate its desired aggregated load profile by sending a carefully selected sequence of DLC

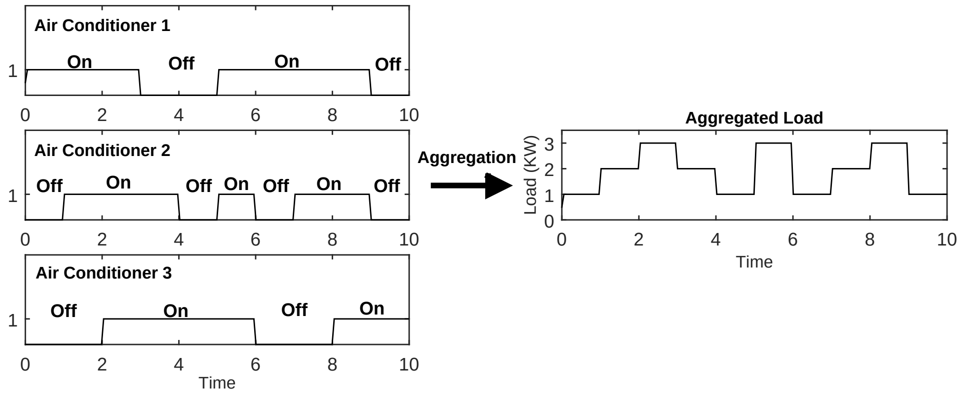


Figure 1.2: An example on how an adversary may achieve its desired aggregated load by sending proper DLC command signals to individual vulnerable loads.

signals - in form of simple on/off commands - to three air conditioners.

In [34], the authors proposed a remote load control mechanism that works over the Internet. Hacking into this Internet-based system may allow taking simultaneous control over several small controllable loads, see Fig. 1.3. Other communications infrastructures, such as cellular or other wireless networks that are used in advanced metering infrastructures (AMIs) [67], may also be vulnerable to various intrusion attempts.

Some load types that could potentially be vulnerable to load altering attacks due to their major role in demand response and DLC include: vacuum cleaners, e.g., Roomba, smart washing machines, e.g., Miele, smart ovens, e.g., LG Thinq, [46], air conditioners [78], water heaters [69], irrigation pumps [42], electric vehicles [62], and computation equipment [23].

Sensing Feedback

In a closed-loop D-LAA, the energy consumption of vulnerable loads is changed according to a feedback signal, such as power system frequency. While a single-point attack requires installing one frequency sensor, a multi-point attack may need one or multiple frequency sensors, depending on how the attack is designed and implemented.

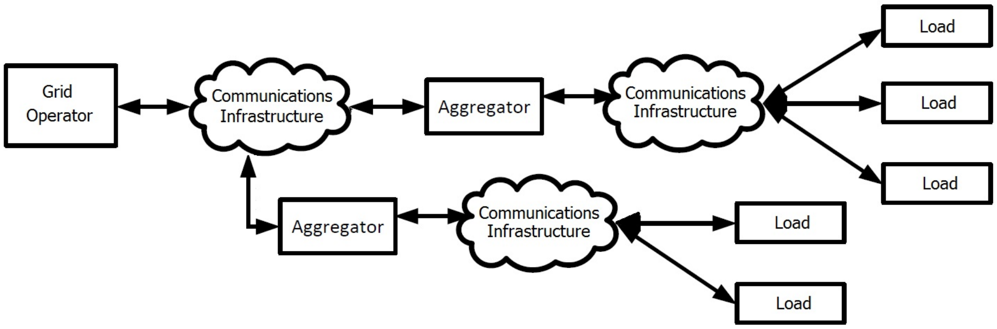

Grid Operator
Communications Infrastructure
Aggregator
Communications Infrastructure
Load
Load
Load
Aggregator
Communications Infrastructure
Load
Load

Figure 1.3: A demand response program may involve two-way communications between grid operator and aggregators and between aggregators and loads. An intrusion may occur in any of these communications infrastructures.

In general, measuring frequency of power grid is not difficult as it can be done at any power outlet using an inexpensive commercial sensor [2]. In fact, such sensing mechanism is already embedded in *frequency responsive* loads that control power usage to contribute, e.g., to frequency regulation [81].

Attack Steps

In summary, an adversary may undergo the following three main steps to implement a D-LAA:

1. Monitor frequency at sensor bus(es) and constantly send measurements to the D-LAA controller. For the special case where the D-LAA controller measures frequency locally, i.e., when the sensor bus and the victim bus are the same, frequency can be measured without the need to intrude into any cyber or physical system.

2. Calculate the amount of vulnerable load P^{LV} that needs to be compromised at victim bus(es) according to the feedback signal and based on the attack control mechanism. This step is done inside the D-LAA controller; therefore, no intrusion is needed in this step.

3. Remotely control the victim load at the amount that is calculated in Step 2. This is

the only step which requires an intrusion mechanism in order to remotely control the load.

Assessing the vulnerability of Supervisory Control and Data Acquisition (SCADA) systems in smart grids, i.e., the focus of Step 3 above, is also discussed in [20, 67].

1.4 Main Contributions

The main contributions of this thesis are summarized as follows:

1.4.1 Destabilizing Anomalies

- A comprehensive model for power system dynamics in presence of destabilizing anomalies is developed.
- Dynamic Load Altering Attacks (D-LAA) are introduced, characterized, and classified as a new form of cyber-physical attacks against smart grid.
- A closed-loop D-LAA against power system stability is formulated and analyzed, where the attacker controls the changes in the victim load based on a feedback from the power system frequency. System vulnerabilities and the impacts of single-point and coordinated multi-point attacks are assessed.

1.4.2 Data-Driven Detection Methods

- This thesis introduces the problem of detecting destabilizing anomalies from measurement data only, and without knowledge of the power system dynamics. The data-driven detection problem is addressed for smart meters readings only, and for smart meter readings together with frequency measurements.

- The detection with smart meter readings only is addressed in the frequency domain. We show that a destabilizing anomaly is detectable through a frequency domain analysis, and that the attack signature corresponds to the system poles that are relocated by the adversary. We provide conditions on the time resolution of the smart meters to ensure anomaly detection, and we highlight the potential interference from instrumentation and communication devices.

- For the case when smart meter readings and frequency measurements are both available, we show that a cross-correlation analysis allows to detect anomaly, and to distinguish between anomalies and the effect of benign frequency responsive loads.

1.4.3 Location Identification Approaches

- The proposed location identification approaches do not require prior knowledge of the number of buses that are compromised. That is, as in practice, we assume that the grid operator is not initially aware of how many buses are compromised. Nevertheless, our methods can identify which buses are compromised. In fact, we provide a means to effectively estimate the unknown number of affected fault/attack location(s).

- They are capable of distinguishing destabilizing attacks, i.e., load or generation control loops that are malicious and based on positive feedback, from the many load and generation control loops that exist in a power system that are benign and based on negative feedback.

- The frequency-domain location identification approach makes direct use of the information that is obtained during the detection phase. In particular, it uses the frequency at which the fault/attack *signature* was detected.

- Compared to its time-domain counterparts, such as unknown input observers, it needs a lower time resolution for measurements, because it does *not* need to reconstruct the entire unknown input signals before it can identify the location(s) of affected power system inputs.

- The proposed optimization-based location identification approach is computationally efficient.
- The proposed approach is well-suited to be deployed in wide area monitoring systems (WAMS) to do fault/attack location identification in a *hierarchical* fashion.

1.4.4 Protection and Mitigation Schemes

- A protection scheme is designed against destabilizing anomalies by formulating and solving a non-convex pole placement optimization problem. It seeks to minimize the total vulnerable load that must be protected to assure power system stability under destabilizing anomalies against the remaining unprotected vulnerable loads. Designing under uncertainty with respect to the exact anomaly location is also taken into consideration.

The techniques that are developed in this thesis are tested and verified on illustrative examples based on an IEEE 9 bus test system, and on a large multi-area IEEE 39 bus test system.

Chapter 2

Power System Dynamics in Presence of Anomalies

2.1 Power System Dynamics

Consider a power transmission system with $\mathcal{B} = \mathcal{G} \cup \mathcal{L}$ as the set of buses, where $\mathcal{G}$ and $\mathcal{L}$ are the sets of generator buses and load buses, respectively. An example is shown in Fig. 2.1. The linear power flow equations at each bus $i \in \mathcal{B}$ can be written as [24]:

$$P_i^E = \sum_{j \in \mathcal{G}} H_{ij}(\delta_i - \delta_j) + \sum_{j \in \mathcal{L}} H_{ij}(\delta_i - \theta_j), \forall i \in \mathcal{G}, \tag{2.1}$$

$$-P_i^L = \sum_{j \in \mathcal{G}} H_{ij}(\theta_i - \delta_j) + \sum_{j \in \mathcal{L}} H_{ij}(\theta_i - \theta_j), \forall i \in \mathcal{L}, \tag{2.2}$$

where P_i^E is the injected electrical power of the generator at bus i, P_i^L is the power consumption of the load at bus i, δ_i is the voltage phase angle at generator bus i, θ_i is the voltage phase angle at load bus i, and H_{ij} is the admittance of the transmission line between buses i and j. If there is no transmission line between buses i and j, then we have $H_{ij} = 0$.

We adopt the linear swing equations, c.f., [35], to model the generator dynamics at each generator bus $i \in \mathcal{G}$, that is,

$$\dot{\delta}_i \ = \omega_i, \tag{2.3}$$

$$M_i \dot{\omega}_i = P_i^M - D_i^G \omega_i - P_i^E, \tag{2.4}$$

where ω_i is the rotor angular frequency deviation at the generator bus i, $M_i > 0$ is the inertia of the rotor, $D_i^G > 0$ is the damping coefficient, and P_i^M is the mechanical power input. We assume two controllers that affect the mechanical power input: turbine-governor controller and load-frequency controller [24]. The turbine-governor controller compares the rotor frequency with a base frequency, for instance 377 rad/s, to determine the amount of mechanical power that is needed to compensate the generated electrical power at steady state. The load-frequency controller, which has a slower dynamic, aims to maintain the rotor frequency at its nominal level by pushing the frequency deviation ω_i back to zero. The two controllers can together be modeled as a proportional-integral (PI) controller, that is,

$$P_i^M = -\left(K_i^P \omega_i + K_i^I \int_0^t \omega_i + P_i^G \right), \tag{2.5}$$

where P_i^G denotes the constant power generation at the generator bus i which is zero for generators with Automatic Generation Control (AGC) and non-zero for generators without AGC. Also, $K_i^I > 0$ and $K_i^P > 0$ are the proportional and integral controller coefficients, respectively. These coefficients are zero for the generators without AGC. Equation (2.4) can be rewritten by combining (2.1) and (2.5) as

$$-M_i \dot{\omega}_i = \left(K_i^P + D_i^G\right) \omega_i + K_i^I \delta_i + \sum_{j \in \mathcal{G}} H_{ij} \left(\delta_i - \delta_j\right) + \sum_{j \in \mathcal{L}} H_{ij} \left(\delta_i - \theta_j\right) + P_i^G, \quad \forall i \in \mathcal{G}. \tag{2.6}$$

Three load types are considered in this system [80]: (i) uncontrollable, (ii) controllable but frequency-insensitive, and (iii) controllable and frequency-sensitive. For notational convenience, at each load bus i, we represent the type (i) and type (ii) loads with term P_i^L in (2.2), and represent the type (iii) loads with term $D_i^L \varphi_i$, where $\varphi_i = -\dot{\theta}_i$ is the frequency deviation at load bus i. The power flow equation in (2.2) becomes

$$-D_i^L \varphi_i - P_i^L = \sum_{j \in \mathcal{G}} H_{ij}(\theta_i - \delta_j) + \sum_{j \in \mathcal{L}} H_{ij}(\theta_i - \theta_j), \tag{2.7}$$

and the overall power system dynamics can be conveniently written as the following linear state-space descriptor system:

$$\underbrace{\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & -M & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{E} \begin{bmatrix} \dot{\delta} \\ \dot{\theta} \\ \dot{\omega} \\ \dot{\varphi} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & -I \\ K^I + H^{GG} & H^{GL} & K^P + D^G & 0 \\ H^{LG} & H^{LL} & 0 & D^L \end{bmatrix}}_{A} \overbrace{\begin{bmatrix} \delta \\ \theta \\ \omega \\ \varphi \end{bmatrix}}^{x} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix}}_{B} \overbrace{\begin{bmatrix} P^G \\ P^L \end{bmatrix}}^{u}, \tag{2.8}$$

where H^{GG}, H^{GL}, H^{LG}, and H^{LL} are derived from the imaginary part of the Y-bus admittance matrix , i.e., we have:

$$Y_{bus} = \begin{bmatrix} H^{GG} & H^{GL} \\ H^{LG} & H^{LL} \end{bmatrix}.$$

In practice, several sensors, such as Phasor Measurement Units (PMU) [56], can be used to measure the system states. We denote such measurement outputs by y. We may have:

$$y = Cx, \tag{2.9}$$

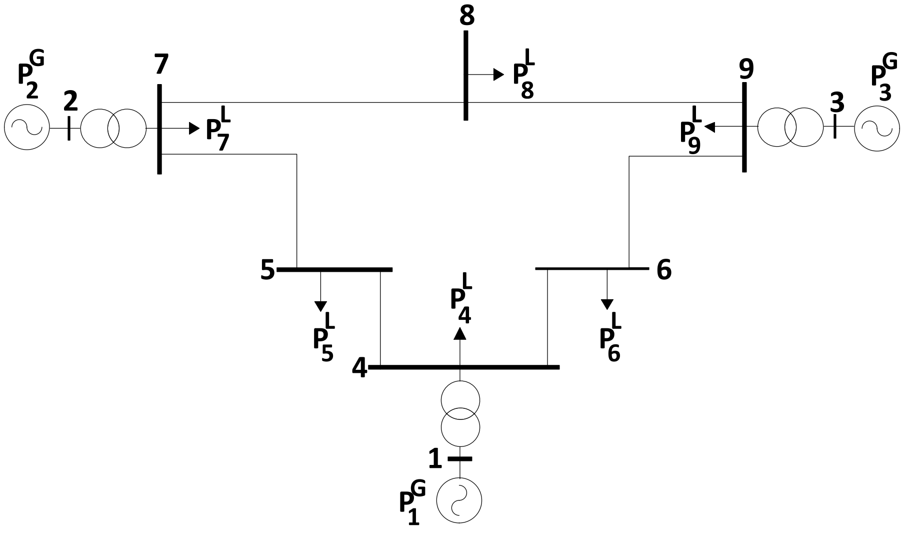


Figure 2.1: The IEEE 9 bus test system. $\mathcal{G} = \{1, 2, 3\}$ and $\mathcal{L} = \{4, \ldots, 9\}$.

where C is the measurement matrix.

2.2 Destabilizing Anomalies

The focus in this thesis is on cases where one or more power system inputs, i.e., the power generation level of generators and/or the power consumption level of loads, are either faulty due to natural causes, or compromised by adversarial actions. We are concerned with those cases where the faulty or compromised inputs have the potential to destabilize the power system at certain operating conditions. In this setup, we model faults and attacks using the following general expression:

$$u^c = u + f, \tag{2.10}$$

where u^c denotes the new input vector under faults and/or attacks, and f denotes the fault and/or attack vector.

We shall point out four notes with respect to (2.10). *First*, the faults and attacks in this thesis are related to physical quantities of the power system inputs. For example, in case of a faulty generator, either there is a fault in choosing the set point or there is a fault in following the set point. In either case, the physical generation output is affected. *Second*, without loss of generality, here we assume that faults and attacks are additive. In principle, the analysis in this thesis is applicable also to multiplicative faults and attacks. *Third*, if an input is neither faulty nor compromised, then the corresponding entry in f is zero. *Fourth*, the fault and attack vector f is essentially a signal. In order to cause destabilization, it must demonstrate certain dynamics. In practice, e.g., when it comes to implementing a destabilizing attack, vector f is likely to be constructed through a *positive feedback* mechanism, see [6], and also the illustrative example in Section 2.3.1.

Once we substitute u with u^c in (2.8), the power system dynamics under destabilizing faults or attacks is read as

$$\begin{aligned} E\dot{x} &= Ax + Bu^c, \\ y &= Cx. \end{aligned} \tag{2.11}$$

The dynamics in (2.11) are different from those in (2.8). The reason is the fact that u^c is not an exogenous signal vector; rather it includes intrinsic positive feedback from system states, as we explained in the forth item in the previous paragraph.

2.2.1 Dynamic Load Altering Attack

Based on the system model in (2.8), a dynamic load altering attack can be characterized based on how it affects the *vulnerable* portion of the load vector P^L, i.e., the input signal in (2.11). Accordingly, at each load bus i, we define

$$P_i^L = P_i^{LS} + P_i^{LV}, \tag{2.12}$$

where P_i^{LS} denotes the *secure* and P_i^{LV} denotes the *vulnerable* portion of the load at bus i, respectively. An attack may compromise only the vulnerable part of a victim load bus.

Now, consider a single-point closed-loop D-LAA that is implemented at victim load buses $\mathcal{V} \subseteq \mathcal{L}$. Suppose a proportional controller is used by the attacker. Let $K_{vs}^{LG} \geq 0$ denote the attack controller's gain at bus $v \in \mathcal{V}$ if the sensor bus s is a generator bus. Similarly, let $K_{vs}^{LL} \geq 0$ denote the attack controller's gain at bus v if the sensor bus s is a load bus. Note that, for each victim load bus v, only one of the two parameters K_{vs}^{LG} and K_{vs}^{LL} can be non-zero, depending on the choice of sensor bus. We can write

$$P_v^{LV} = -K_{vs}^{LG}\omega_s - K_{vs}^{LL}\varphi_s. \tag{2.13}$$

Note that, since K_{vs}^{LG} and K_{vs}^{LL} are positive valued, P_v^{LV} is updated in opposition to the values of ω_s and φ_s. For example, if ω_s decreases, i.e., the frequency drops from its nominal value, then the attack controller increases the load at bus v. This is exactly the opposite of how a frequency-responsive load would react to frequency lag in a DR program, c.f. [81]. The power system dynamics subject to the above D-LAA becomes

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & -M & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\delta} \\ \dot{\theta} \\ \dot{\omega} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & -I \\ K^I + H^{GG} & H^{GL} & K^P + D^G & 0 \\ H^{LG} & H^{LL} & -K^{LG} & -K^{LL} + D^L \end{bmatrix} \begin{bmatrix} \delta \\ \theta \\ \omega \\ \varphi \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P^G \\ P^{LS} \end{bmatrix}. \tag{2.14}$$

From (2.14), the attacker is capable of affecting the system dynamics. Specifically, the attacker can affect the system matrix and the system poles by adjusting its controller matrices K^{LG} and K^{LL}. If the size of the vulnerable load is large enough, then the attacker can render the system dynamics unstable by moving the system poles to the right-half complex plane [15]. Of course, in practice, since the generators are equipped with over- and under-frequency relays as part of their protection systems, c.f. [70], a D-LAA may ultimately force certain generators to disconnect from the main grid, possibly triggering cascading effects or blackouts.

Next, we investigate sufficient conditions for making (2.14) unstable. To do so, we modify the system model in (2.14) into a regular, i.e., non-descriptor state-space model. This is done by eliminating the power flow equations and integrating them into the swing equations. Suppose the sensor bus s is a generator bus, i.e., $s \in \mathcal{G}$. Accordingly, we have $K^{LL}_{vs} = 0$ for all victim load buses v. From this, and the last row in (2.14), we have:

$$\varphi = -\left(D^L\right)^{-1}\left(\begin{bmatrix} H^{LG} \\ H^{LL} \\ -K^{LG} \end{bmatrix}^T \begin{bmatrix} \delta \\ \theta \\ \omega \end{bmatrix} + P^{LS}\right). \tag{2.15}$$

If we substitute (2.15) with φ in (2.14), the equivalent non-descriptor / regular state-space model under attack becomes:

$$\begin{bmatrix} \dot{\delta} \\ \dot{\theta} \\ \dot{\omega} \end{bmatrix} = A \begin{bmatrix} \delta \\ \theta \\ \omega \end{bmatrix} + B\left(-\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & K^{LG} \end{bmatrix}^T \begin{bmatrix} \delta \\ \theta \\ \omega \end{bmatrix} + \begin{bmatrix} P^G \\ P^{LS} \end{bmatrix}\right), \tag{2.16}$$

where

$$A = \begin{bmatrix} I & 0 & 0 \\ 0 & (D^L)^{-1} & 0 \\ 0 & 0 & -M^{-1} \end{bmatrix}\begin{bmatrix} 0 & 0 & I \\ H^{LG} & H^{LL} & 0 \\ K^I + H^{GG} & H^{GL} & K^P + D^G \end{bmatrix}$$

and

$$B = \begin{bmatrix} 0 & 0 \\ 0 & (D^L)^{-1} \\ -M^{-1} & 0 \end{bmatrix}.$$

Note that, we have $K^{LG}_{ij} = 0$ for any $i \notin \mathcal{V}$ and any $j \neq s$.

The state-space model in (2.16) represents the system dynamics in *presence* of a closed-loop D-LAA, where A and B are the system and input matrices in the corresponding open-loop system in *absence* of the D-LAA. The instability of this linear system can be analyzed using the Linear Quadratic Lyapunov Theory that is overviewed in the Appendix A.1. Specifically, the closed-loop system in (2.16) is unstable if there exists a *symmetric negative definite* matrix X such that

$$\left(A - B \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & K^{LG} \end{bmatrix}^T \right)^T X + X \left(A - B \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & K^{LG} \end{bmatrix}^T \right) < 0. \tag{2.17}$$

This Nonlinear Matrix Inequality (NLMI) can be changed to Linear Matrix Inequality (LMI) by applying linear fractional transformation [11]. Specifically, if we define $Y \triangleq X^{-1}$ and $W \triangleq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & K^{LG} \end{bmatrix} X^{-1}$, we can rewrite (2.17) as

$$(A - BWY^{-1})^T Y^{-1} + Y^{-1}(A - BWY^{-1}) < 0. \tag{2.18}$$

If we multiply both sides by Y, we obtain [11]:

$$YA^T - W^T B^T + AY - BW < 0, \tag{2.19}$$

which is an LMI in Y and W . If this LMI has a solution over $Y < 0$, then the Lyapunov function $V(z) = z^T Y^{-1} z$ proves the *instability* of the closed loop system under attack.

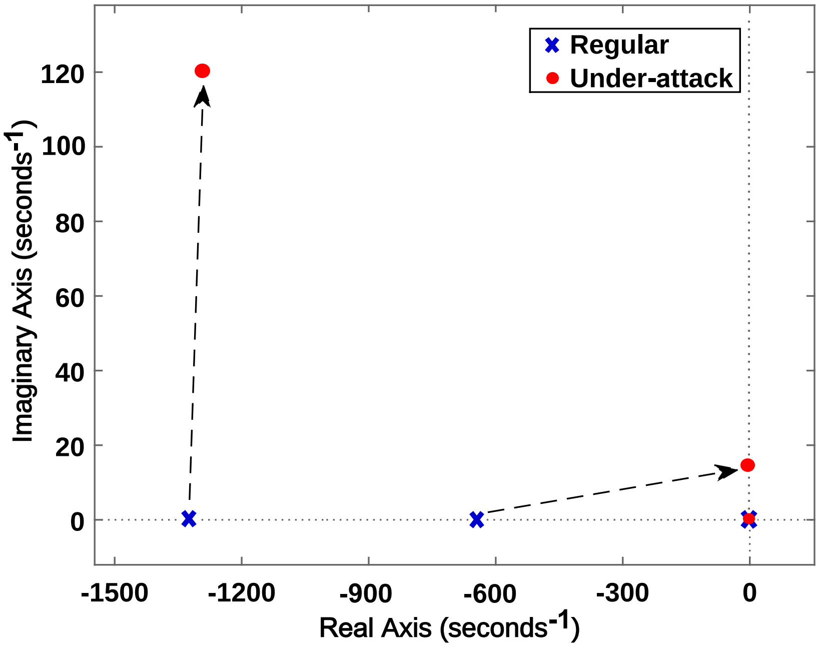


Figure 2.2: An example on how a destabilizing fault or attack can move the dominant eigenvalues of the power system matrix towards the $j\omega$ axis.

2.3 Case Studies

2.3.1 Impact of Destabilizing Anomalies on System Dynamics

Consider the IEEE 9 bus network in Fig. 2.1. Suppose the power system is under a Dynamic Load Altering Attack. The dynamics of the system in Fig. 2.1 under a D-LAA can be described by (2.11), where parameters of matrices E and A are as in [3]. We assume that all three generators have AGC, i.e., $P^G = 0$. Also, P_6^L and P_9^L are affected by D-LAA through the adversary's proportional controllers, with gains 25, and 24, respectively, by taking feedback from ω_3 according to (2.13). Accordingly, the entries in rows 6 and 9 in attack vector f are non-zero. All buses are equipped with PMUs, except for bus 7.

Fig. 2.2 shows how a destabilizing D-LAA changes the power system dynamics by moving the dominant eigenvalues of its system matrix towards the $j\omega$ axis. The power system frequencies at different buses start deviating from their nominal value, i.e., 60 Hz,

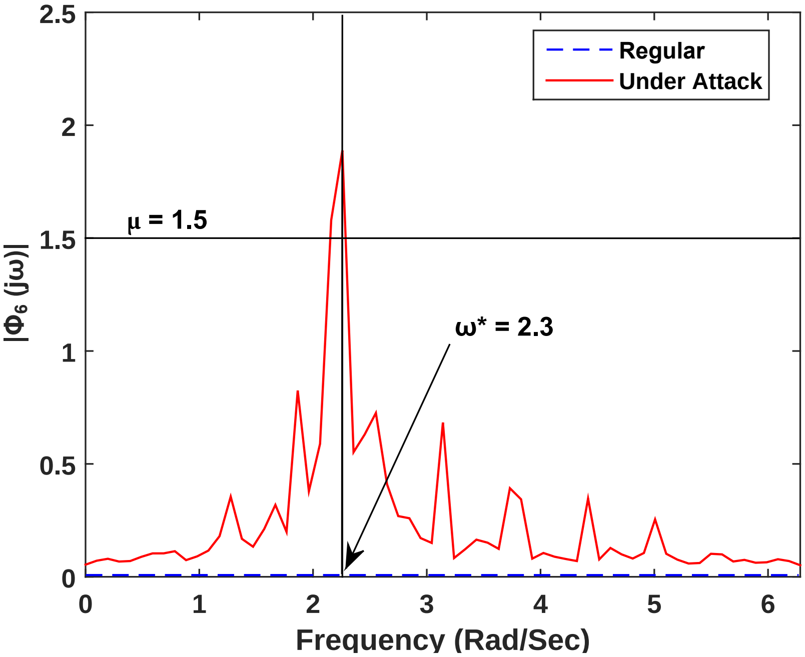


Figure 2.3: The attack in Section 2.3.1 creates a clear signature on frequency ω^* of the FFT magnitude of frequency deviation signal at load bus 6.

putting the system at the margin of instability.

One can evaluate the destabilizing impact of the attack by performing a *frequency-domain* analysis. This requires taking the Fast Fourier Transform (FFT) [17] of the measurement outputs. The results are shown in Fig. 2.3 for both the regular and under-attack scenarios. Here, $\Phi_6(j\omega)$ denotes the FFT representation of the power system frequency deviation at bus 6, i.e., φ_6. The magnitude of $\Phi_6(j\omega)$ is shown by $|\Phi_6(j\omega)|$. We can see a clear *signature* and a large beam at frequency ω^* in this figure for the case with the presence of the attack. The above aforementioned fault/attack signature in frequency domain provides the grid operator with an effective tool to *detect* the fault/attack through a proper data-driven analysis, as explained in [5] or Section 3.2. Accordingly, to monitor the signature at ω^*, sampling frequency of sensors must be at least two times of ω^*. In this example, since $\omega^* = 2.3$, the minimum sampling frequency must be 4.6 radian per second or 0.73 Hz.

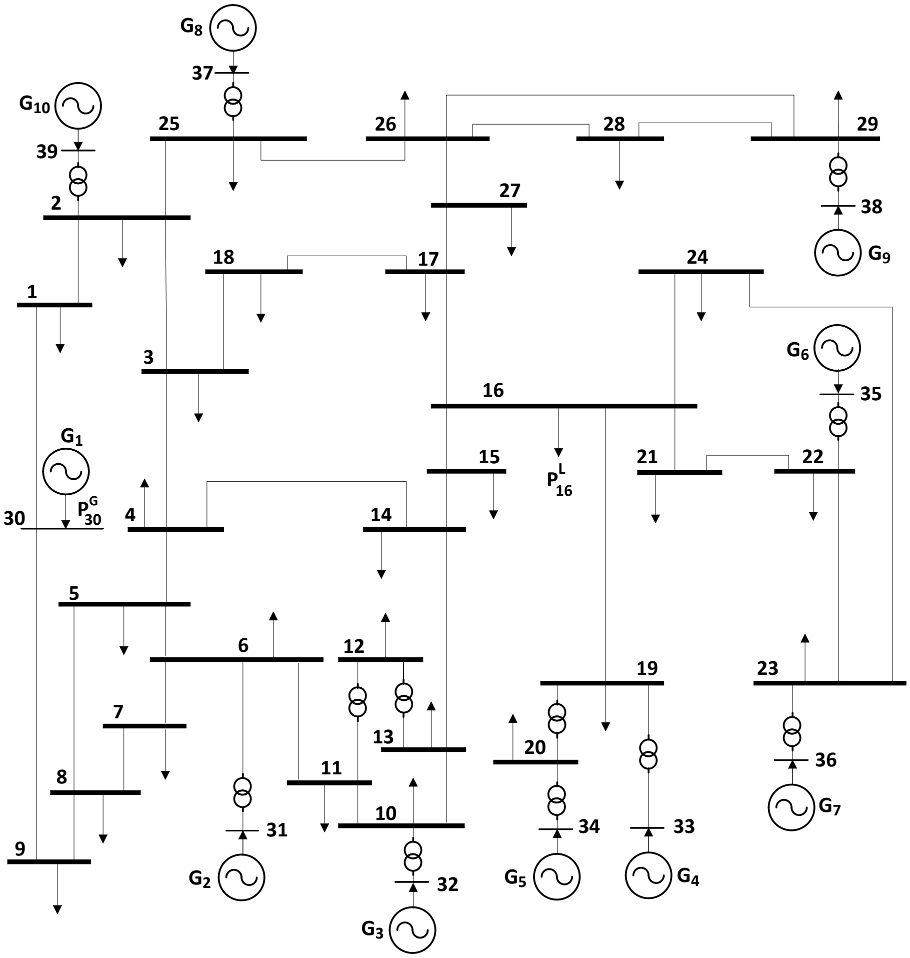


Figure 2.4: The IEEE 39 bus test system based on the 10-machine New-England power network, where $\mathcal{L} = \{1, \ldots, 29\}$ and $\mathcal{G} = \{30, \ldots, 39\}$.

2.3.2 Dynamic Load Altering Attack

Consider the IEEE 39 bus power system in Fig. 2.4. Suppose the parameters of the transmission lines and the inertia and damping coefficients of generators are as in [1]. Secure loads and vulnerable loads at each load bus are as in Table 2.1. Generator controller parameters are $K^P_1 = 100$, $K^P_2 = K^P_3 = 45$, $K^P_4 = 10$, $K^P_5 = K^P_{10} = 50$, $K^P_6 = K^P_9 = 40$, $K^P_7 = 30$, $K^P_8 = 20$, and $K^I_1 = \ldots = K^I_{10} = 60$. The damping coefficient for each fixed dynamic load is 10. Controller parameters are set so as to keep the system stable during normal operations, i.e., in absence of an attack. The system is initiated to run with P^L

Table 2.1: Total loads and vulnerable loads

Load Bus	P^{LS} **(p.u)**	P^{LV} **(p.u)**	**Load Bus**	P^{LS} **(p.u)**	P^{LV} **(p.u)**
1	4	0	**16**	7.8	3.1
2	4	0	**17**	4	0
3	7.2	0	**18**	5.6	0
4	9	0	**19**	4	1.6
5	4	0	**20**	10.3	0
6	5	2	**21**	6.7	0
7	6.3	0	**22**	4	0
8	9.2	0	**23**	7	2.8
9	4	0	**24**	7	0
10	4	0	**25**	6.2	0
11	4	0	**26**	5.4	0
12	4.1	0	**27**	6.8	0
13	4	0	**28**	6.1	0
14	4	0	**29**	10.8	4.3
15	7.2	0	—	—	—

being equal to $P^{LS} + P^{LV}/2$.

We assume that only five load buses have vulnerable loads. They can potentially become victim buses, i.e., we can have $\mathcal{V} = \{6, 16, 19, 23, 29\}$. These victim load buses are highlighted using color gray in Table 2.1. Sensor buses are assumed to be placed only at $\mathcal{S} = \{31, 33, 36, 38\}$. The nominal system frequency is 60 Hz. The generator's over-frequency relays trip at 62 Hz and the under-frequency relays trip at 58 Hz.

Assessing System Vulnerabilities

The attacker can assess the vulnerability of the loads at each load bus to see the possibility of conducting D-LAA in the power system, also the type of attack. Fig. 2.5 shows how the root locus [15] analysis helps the attacker to find the minimum attack gain $K^{LG}_{19,33} = 15$ to conduct a single-point D-LAA when $v = 19$ and $s = 33$. If we multiply the minimum attack gain by two times the frequency deviation threshold $\omega_s^{\max} = 2/60$ at which the generators frequency relays trip, then we can conclude that at least $2K^{LG}_{19,33}\omega_s^{\max} =$

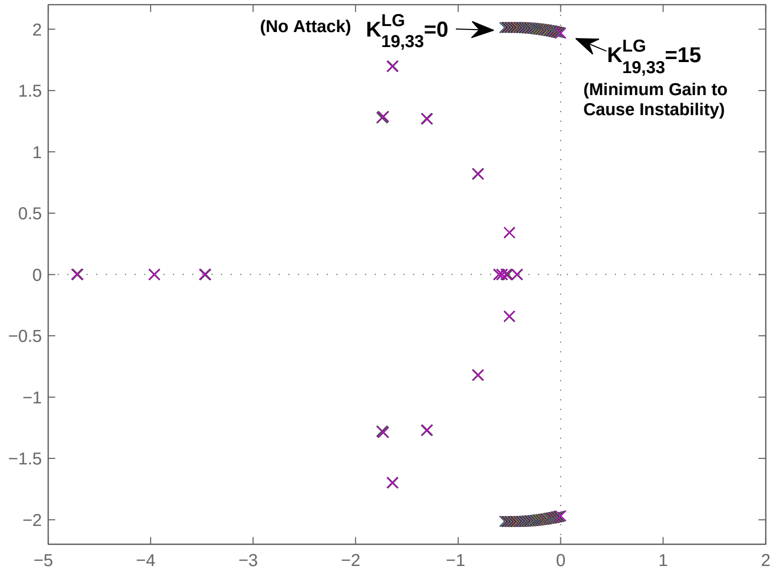


Figure 2.5: Power system poles versus the attack gain $K^{LG}_{19,33}$.

$15 \times 2 \times 2/60 = 1$ p.u. of the total 1.6 p.u. vulnerable load at victim bus 19 must be compromised when the frequency sensor is at bus $s = 33$ in order to have a successful single-point D-LAA. Note that the compromised load consumption must follow the frequency signal by a proportional controller. Also, the frequency signal deviates around its nominal value. Hence, the multiplication by two in $2K\omega_s$ is due to the fact that the compromised load must provide enough room to allow both over and under frequency fluctuations before the attack makes the frequency relays tripped. Similarly, we can calculate the minimum portion of vulnerable load that must be compromised for having successful single-point D-LAAs for all victim and sensor bus scenarios to find the vulnerabilities of the power system. The results are shown in Table 2.2. We can see that only two successful single-point attacks are feasible: a single-point attack at victim bus $v = 19$ with sensor bus $s = 33$, and a single-point attack at victim bus $v = 29$ and sensor bus $s = 38$. No other single-point attack is feasible due to lack of sufficient vulnerable load. Another implication of the results in Table 2.2 is with respect to the coordinated multi-point attacks. For example, based on the column with $s = 33$, although hacking the loads individually at victim buses 16 and 23

Table 2.2: Minimum portion of vulnerable load that must be compromised to assure a successful D-LAA

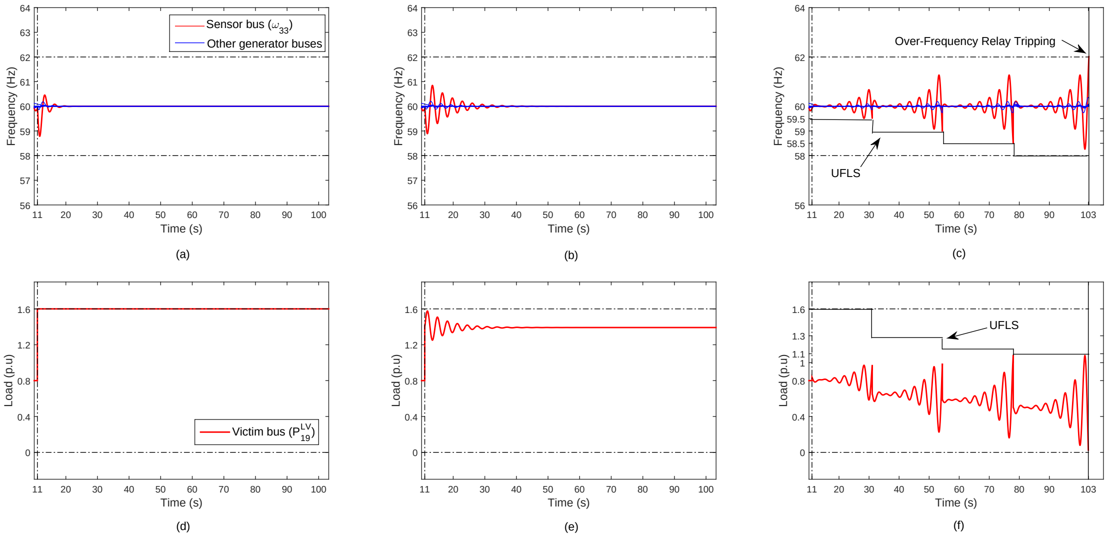
Sensor Bus **Victim Bus**	**31**	**33**	**36**	**38**
6	4.9	18.4	81.2	128.1
16	24.7	1.2	6.5	23.3
19	69.2	0.6	15.2	48.8
23	79.1	3.2	1.9	66.8
29	92.2	8.9	46.5	0.7

cannot lead to successful single-point attacks, it might be possible to hack some loads at both buses and conduct a successful *coordinated multi-point* D-LAA.

Single-point Attack

Next, we examine three single-point attack scenarios for the case where $v = 19$ and $s = 33$. The results are shown in Fig. 2.6. First, assume that the attack is *static*, causing an abrupt change in victim load as shown in Fig. 2.6(d). The poles of the system are *not* changed under this attack. We can see in Fig. 2.6(a) that the system can easily absorb such one-time abrupt change. Second, assume that the attack is *dynamic* and $K^{LG}_{19,33} = 10$. We can see in Fig. 2.6(b) that the attack causes some relatively major over- and under-shoots in frequency. Nevertheless, the system remains stable and the frequency deviation is forced back to zero.

Finally, suppose the attack is dynamic and $K^{LG}_{19,33} = 20$. Under this third attack, two of the system poles are pushed to the right half-plane, making the system *unstable*. What is different in this case is that load Bus 19 is assumed to be equipped with a three-step UFLS protection relay [28]. This UFLS sheds only the vulnerable (but protected) portion of the load in response to frequency drop in its three sequential steps as listed in Table 2.3. Fig. 2.6(c) shows that even after the three load shedding steps by the UFLS relay, the attack can still force the frequency deviation at generator bus $s = 33$ to reach



(a)

(b)

(c)

(d)

(e)

(f)

Figure 2.6: Simulation results under various attack conditions. First row: system frequencies over time. Second row: vulnerable load changes. First column: S-LAA causing an abrupt load increase. Second column: an unsuccessful D-LAA with $K_{19,33}^{LG} = 10$. Third column: a successful D-LAA with $K_{19,33}^{LG} = 20$.

Table 2.3: Frequency settings of the UFLS relay

Step Number	**Frequency Setting (Hz)**	**Amount of P_{19}^{LV} to be shed (p.u)**
1	**59.5**	**20% = 0.32**
2	**59**	**10% = 0.13**
3	**58.5**	**5% = 0.06**

the threshold $\omega_1^{\max} = 2/60$ p.u., causing the over-frequency relay of the generator at bus 33 to trip at time $t = 103s$, pushing this generator offline, thus, concluding the attack. Interestingly, the D-LAA under this last scenario did not need to hack the entire available vulnerable load at bus $v = 19$. Instead, it only followed the *right* trajectory in response to the changes in frequency in order to be successful.

Note that, implementing a D-LAA does not require all loads to be equipped with smart meters. In fact, according to Table I, only less than *one-third* of the loads at each bus are assumed to be vulnerable. That means, at each bus, over *two-third* of the loads are traditional loads and may not even have smart meters or any demand response equipment.

Also, only a portion of the vulnerable loads needs to be compromised to conduct a successful attack. For example, according to Table II, the adversary can plan a single-point D-LAA by compromising only 60% of the total vulnerable loads at bus 19. Hence, only 60 % of smart meters at bus 19 need to be compromised.

Coordinated Multi-point Attack

Recall from Section 2.3.2 that a coordinated multi-point attack at victim buses $v = 16$ and $v = 23$ might lead to a successful D-LAA. The amount of vulnerable load that needs to be hacked at each of the two victim buses to make the system unstable can be obtained using a two-dimensional root locus analysis in form of an exhaustive search. The results are shown in Fig. 2.7. This figure shows the *attack success time*, i.e., the time that takes from the moment the attack is launched until the moment the target generator goes offline, for all possible combinations of hacking vulnerable loads at buses $v = 16$ and $v = 23$. Note that, for those combinations where a successful attack is not feasible, no point is plotted in the curve. We can conclude that, while increasing the amount of compromised loads may not always be necessary to make the system unstable, it can still be beneficial to decrease the attack success time.

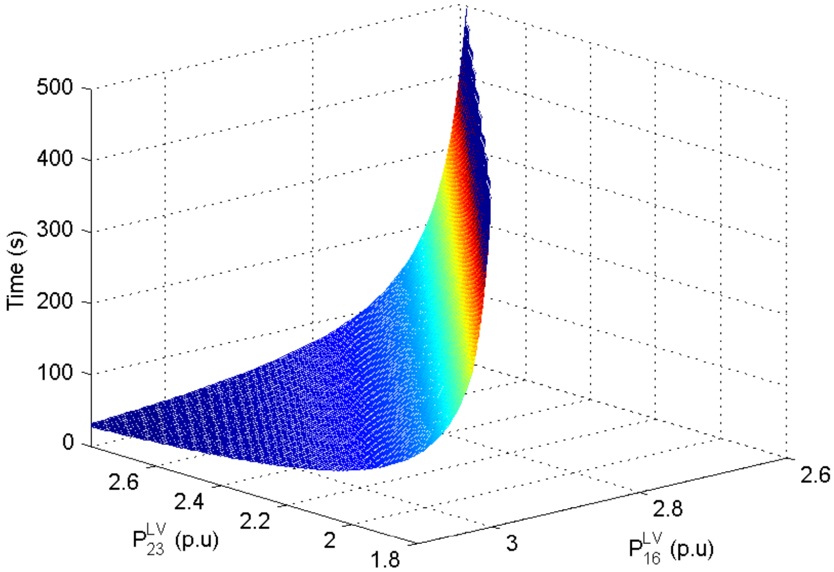


Figure 2.7: Attack success time versus the amount of compromised load at each victim load bus in a coordinated multi-point closed-loop D-LAA.

Chapter 3

Data-Driven Detection Methods

3.1 Smart Meter Data and a Case Study

The focus in this section is on data-driven detection of destabilizing anomalies in power systems. In this regard, we use the experimental data in [8], which includes the smart meter data of three different homes (A, B, and C) in Western Massachusetts; see Fig. 3.1. The time-resolution for all smart meters is *one second*. All three homes have typical household appliances such as refrigerator, washing machine, etc. Home A is further instrumented with appliance-level submeters that monitor the loads for all appliances separately. For example, about 30 of 35 wall switches have been replaced with units that transmit on-off-dim events for the switches to a gateway server at about every 2.5 seconds (on average) by using Power Line Communication (PLC) data transmissions.

To utilize the above smart meter data in a study on anomaly detection, we integrated them into the 39-bus IEEE test system in Fig. 2.4. The parameters of the transmission lines and the inertia and damping coefficients of generators are as in [1]. The generator controller parameters are chosen as $K_1^P = 100$, $K_2^P = K_3^P = 45$, $K_4^P = 10$, $K_5^P = K_{10}^P = 50$, $K_6^P = K_9^P = 40$, $K_7^P = 30$, $K_8^P = 20$, and $K_1^I = \ldots = K_{10}^I = 60$. The damping coefficient for each fixed dynamic load is 10. Note that, the generator controller parameters are set so

Table 3.1: Power consumption at load buses in per unit.

Bus	P^L	**Bus**	P^L	**Bus**	P^L	**Bus**	P^L	**Bus**	P^L
1	4	**7**	6.3	**13**	4	**19**	B	**25**	6.2
2	4	**8**	9.2	**14**	4	**20**	10.3	**26**	5.4
3	7.2	**9**	4	**15**	7.2	**21**	6.7	**27**	6.8
4	9	**10**	4	**16**	A	**22**	4	**28**	6.1
5	4	**11**	4	**17**	4	**23**	C	**29**	10.8
6	5	**12**	4	**18**	5.6	**24**	7	**-**	-

as to keep the system stable during normal operations and in the absence of attacks.

The system is initiated to run with constant P^L for all load buses as in Table 3.1. The load on buses 16, 19, and 23 change according to the smart meter data in Fig. 3.1. We assume that there are 10,000 homes at bus 16 whose load profile equals that of Home A, 30,000 homes at bus 19 whose load profile equals that of Home B, and 20,000 homes at bus 23 whose load profile equals that of Home C. A closed-loop D-LAA with $K_{vs}^{LL} = 43$ is attempted at bus 19 which is both victim and sensor bus.

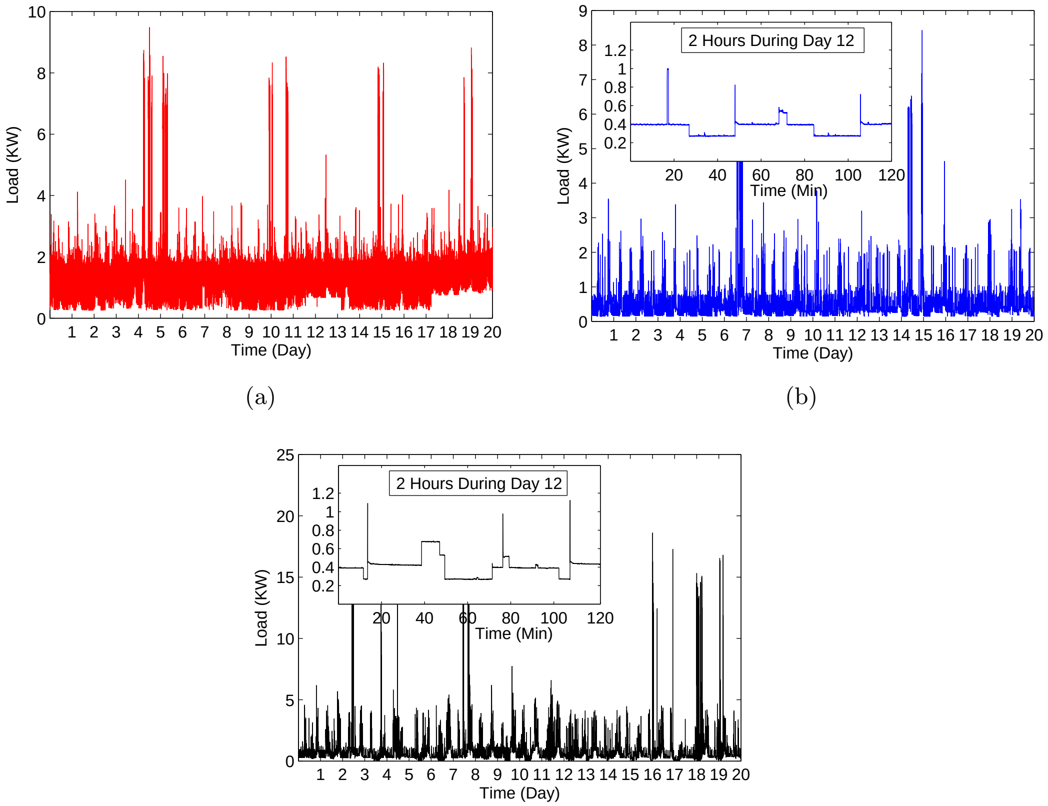
There is a frequency-responsive load with gain -2 at bus 23. Fig. 3.2 shows load and frequency at buses 19 and 23 for two hours on day 12. The D-LAA at bus 19 and the frequency-responsive load at bus 23 are activated at $t = 16.6$ min.

3.2 Detection Solely Based on Input Signal

In this section we characterize the possibility of detecting anomalies through the knowledge of the input data only.

3.2.1 Frequency Domain Analysis

Consider the typical smart meter data in Fig. 3.1. Notice that the load signal has major fluctuations during the day. Suppose that a portion of the load at a victim bus is compromised, e.g., in form of the load in Fig. 3.2(a). If the volume of the compromised



(a)

(b)

(c)

Figure 3.1: Smart meter data at second-by-second resolution over 20 days from May 1, 2012 to May 20, 2012: a) Home A, b) Home B, c) Home C.

load is high, then it can be detected by looking at the smart meter data in time domain, because the frequency-responsive behavior of the compromised load under D-LAA in this case significantly changes the shape of the total metered load. Such detection can be done automatically, e.g., by using appropriate pattern recognition algorithms, c.f. [37]. However, if the volume of the compromised load under D-LAA is low, then the attack may be difficult to detect through time-domain analysis, and a frequency-domain analysis may be preferable.

We first take the Fast Fourier Transform (FFT) [17] of the original load signals of Fig. 3.1, i.e., the second-by-second load profile in the absence of D-LAAs. Fig. 3.3 shows the results. The DC portion of the signal is omitted before applying the FFT algorithm.

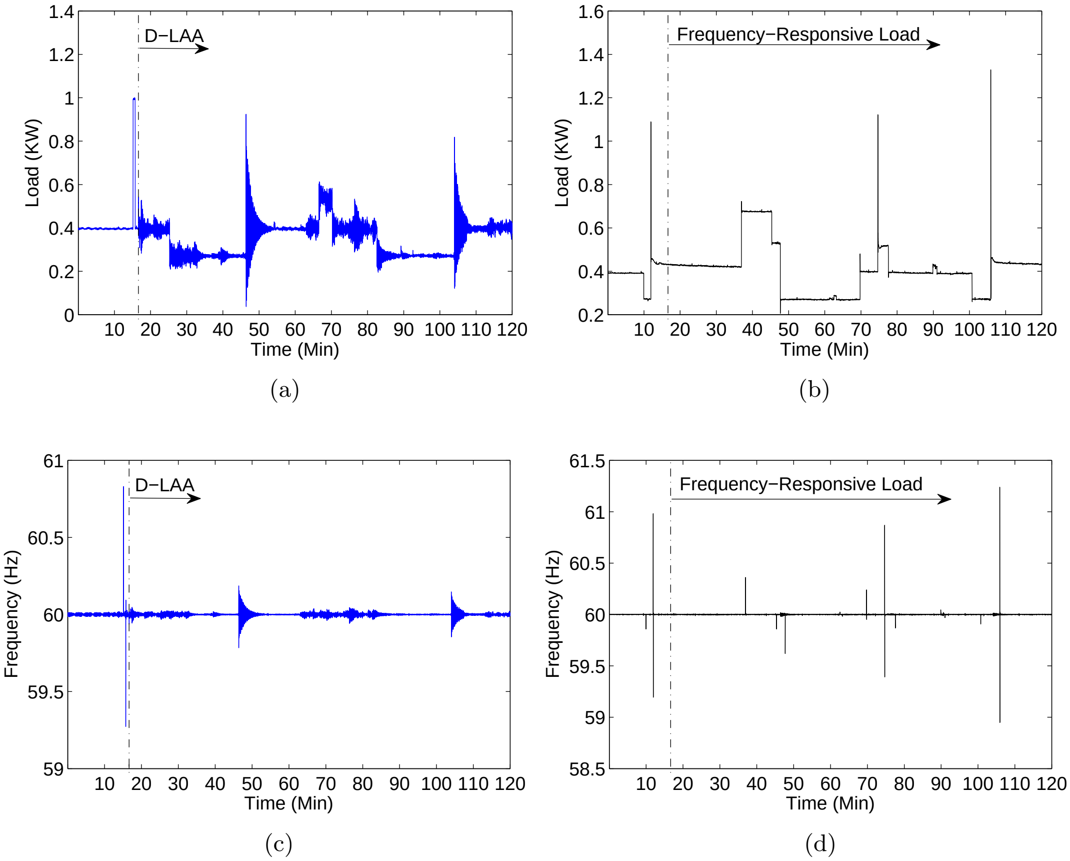
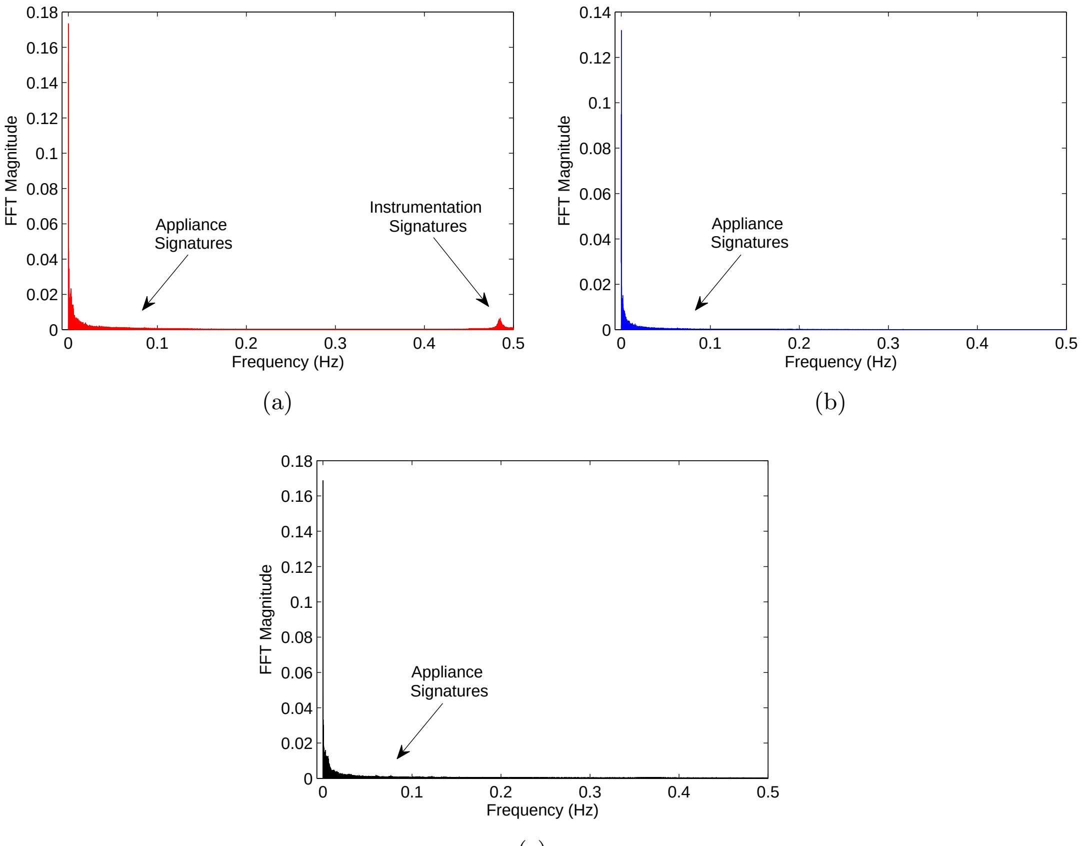


Figure 3.2: The measurements corresponding to the case study in Section 3.1: a) load signal for a home at bus 19, b) load signal for a home at bus 23, c) frequency signal a home at bus 19, d) frequency signal for a home at bus 23.

We can see that, except for some noticeable non-zero coefficients around 0.47 Hz for Home A, the FFT coefficients are negligible at frequencies above 0.05 Hz. Note that, for the non-zero coefficients around 0.47 Hz, they do not represent any residential load. Instead, they are created due to extensive instrumentation of Home A and the fact that about 30 wall switches make PLC-based transmissions of the submeter data to a gateway once roughly every 2.5 seconds, see Section 3.1 and Remark 3. From Fig. 3.3, the FFT of a typical residential load signal has non-zero coefficients only at very low frequencies.

Remark 1 *The spectrum analysis of the smart meter data in this thesis is very different from the well-studied analysis of* harmonics *for nonlinear loads in power systems and power*



(a)

(b)

(c)

Figure 3.3: The spectral analysis of the residential load signals in Fig. 3.1 over the entire 20 days period using (3.2) and (3.3): a) Home A, b) Home B, c) Home C.

electronics, c.f., [19, 44]. Let $p(t)$ denote the instantaneous power draw for a load. Note that, $p(t)$ is a continuous-time signal. In order to analyze the harmonics for nonlinear loads, one would take the following continuous Fourier transform:

$$\mathcal{F}_C \left\{ p(t) \right\}. \tag{3.1}$$

Now, consider a smart meter that reports the average power usage every T seconds. The k^{th} meter reading is calculated as

$$\bar{p}[k] = \frac{1}{T} \int_{t=(k-1)T}^{kT} p(t) dt. \tag{3.2}$$

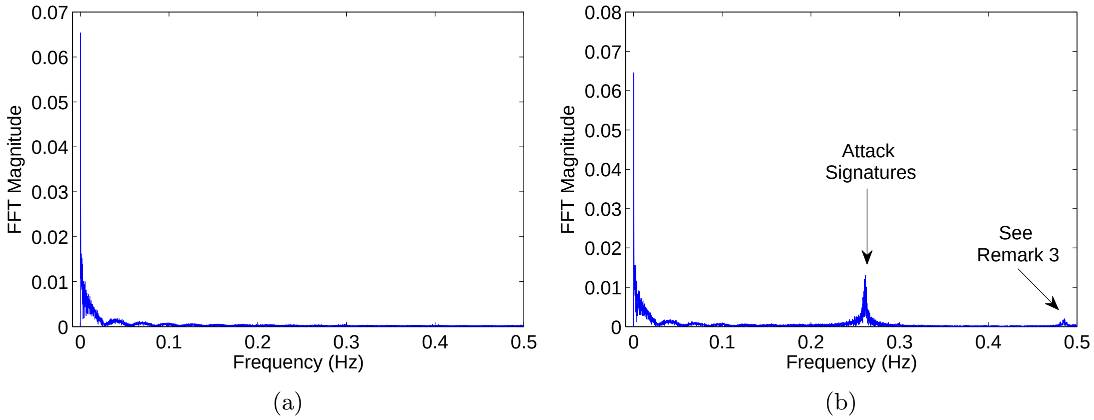


Figure 3.4: Spectral analysis of the original and compromised load signals of each home at bus 19 over a two-hour period: a) original, b) compromised.

In this paper, for the purpose of dynamic load altering attack detection, we take the following discrete Fourier transform:

$$\mathcal{F}_D \left\{ \bar{p}[k] \right\}. \tag{3.3}$$

Thus, our spectral analysis is focused on much lower frequencies than in a typical harmonics analysis of nonlinear loads.

Next, consider the two-hour zoomed-in time frame in Fig. 3.1(b). The frequency spectrum for the load signal of each home at bus 19 without and with attack is shown in Figs. 3.4(a) and (b), respectively. We see that the presence of D-LAA has created a *new signature* to the frequency spectrum at about 0.26 Hz; see Remark 2. This new signature is away from the load signatures. Hence, it can be used to detect the attack; see Section 3.2.2. The magnitude of attack signature depends on factors such as amount and location of the compromised load.

Remark 2 *The attack has moved a pair of system poles from* $-0.55 \pm 2.01i$ *to* $-0.0095 \pm 1.64i$. *Since the* real *part of these poles has increased, the poles are now much closer to the imaginary axis, making the system (almost) only marginally stable. The new poles induce slowly decaying oscillations with larger magnitudes compared to other oscillations in the* *system, creating a noticeable attack signature in frequency domain. As for the* imaginary *part of relocated poles, it highly affects the frequency at which we should see the attack frequency signature. Specifically, the attack signature in Fig. 3.4(b) has appeared at the* natural frequency *of the relocated poles [15]:*

$$\omega_n = \sqrt{-0.0095^2 + 1.64^2} = 1.64. \tag{3.4}$$

Note that, $f_n = \omega_n/(2\pi) = 0.26Hz$, which equals the central frequency of the attack signature in Fig. 3.4(b).

The above remark may also give some basic hints on how an attacker may conduct an optimal pole placement - subject to the available load vulnerabilities - in order to maximize the attack impact on the power grid while minimizing the chance of being detected through frequency-domain analysis.

Remark 3 *Besides the main attack signature at 0.26 Hz, the attack has also created a small signature at 0.47 Hz in Fig. 3.4(b). Interestingly, this signature is* indirectly *related to the instrumentation signature that we previously identified for the load at bus 16. Note that, since the grid is an* interconnected *system, the dynamics of loads/genertors at any bus may have impact on the frequency at another bus. Accordingly, since the instrumentation signature at 0.47 Hz at bus 19 has some impact on the frequency fluctuations at bus 19, and also because the compromised loads at bus 19 respond to the frequency fluctuations at bus 19, the instrumentation signature at bus 16 has now appeared, although with attenuations, at bus 19. This suggests that the communications activities of instrumentation devices can potentially* interfere *with attack detection. However, such interference is likely negligible in practice, as the power usage of instrumentation is low compared to the load of a home.*

One may ask: *is it possible to see the attack signature if the meter data is minute-by-minute instead of second-by-second?* The answer is ‘no’, as it is explained in the next remark.

Remark 4 *Based on the Nyquist-Shannon sampling theorem [25], the sampling frequency must be twice the highest frequency of the signal in order to avoid aliasing in the signal spectrum. Of course, the integral nature of energy metering operation in (3.2) is different from standard sampling. Nevertheless, the above theorem may still provide a good practical approximation for the minimum required time resolution of smart meters. Loosely speaking, for the attack signature to be observable in a frequency-domain analysis, it is required that*

$$T \leq \frac{1}{2f_n} = \frac{\pi}{\omega_n}, \tag{3.5}$$

where T is the smart meter pulse interval; see (3.2). For example, to detect the attack signature in Fig. 3.4(b), the reading interval of the smart meter needs to be roughly two seconds or less.

3.2.2 Real-time Detection in Frequency Domain

In the previous section, we studied the detectability of D-LAA via spectral analysis. In order to detect an attack in a prompt and efficient manner, in this section we employ the Windowed FFT (W-FFT) method [19]. The performance of W-FFT is affected by the choices of three parameters: *window size*, *sampling rate*, and *detection threshold.* The window size indicates the length of time series signal in each FFT window. The sampling rate indicates the time between two consecutive FFT window samples. The detection threshold indicates the smallest magnitude for the FFT or W-FFT coefficients around the natural frequency of a relocated system pole that triggers the detection of an attack frequency signature.

Suppose we set the sampling rate to 100 sec, window size to 200 sec, and detection threshold to 0.001. To assess the efficiency of W-FFT in detecting D-LAAs, we calculate the W-FFT coefficient at attack frequency 0.26 Hz for each sliding window. The results

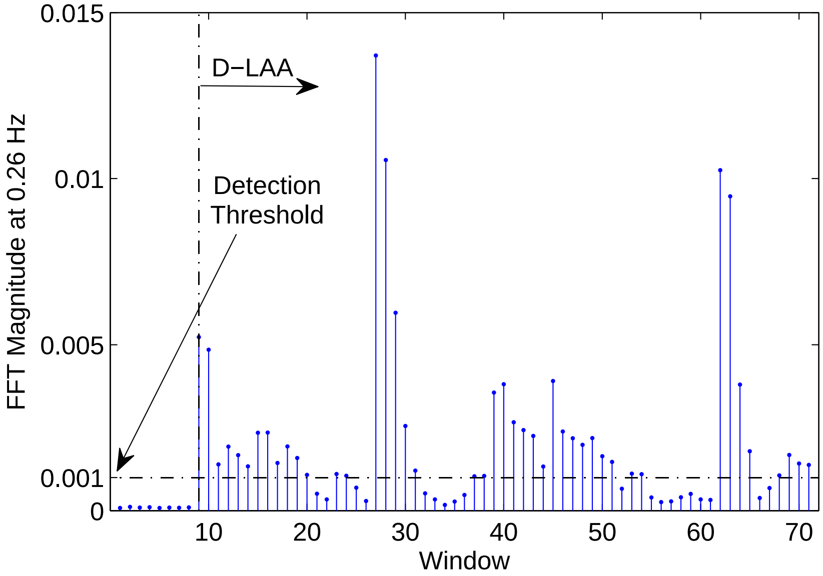


Figure 3.5: The W-FFT coefficient of a compromised load at attack frequency 0.26 Hz versus the W-FFT sliding windows for each home at bus 19.

are shown in Fig. 3.5. We can see that the W-FFT coefficient at attack frequency 0.26 Hz exceeds the detection threshold right after the attack is launched. This allows an immediate detection of the attack. However, there are also certain windows, e.g., windows number 21 and 22, where the W-FFT coefficient is below the detection threshold.

Finally, we must also point out a key limitation of detecting D-LAA solely based on load signals. Consider the W-FFT coefficients for a benign frequency-responsive load of a home at bus 23 in Fig. 3.6. We can see that there are still quite a few W-FFT coefficients that exceed the detection threshold, even though a frequency-responsive load is helping the grid.

Remark 5 *The frequency-domain analysis in this section is effective in detecting* load activities *around the natural frequencies of the system poles. However, it cannot distinguish between a compromised load (with adverse activity) and a frequency-responsive load (with benign activity), because such distinction is not possible by solely looking at the load signal and without considering frequency measurements.*

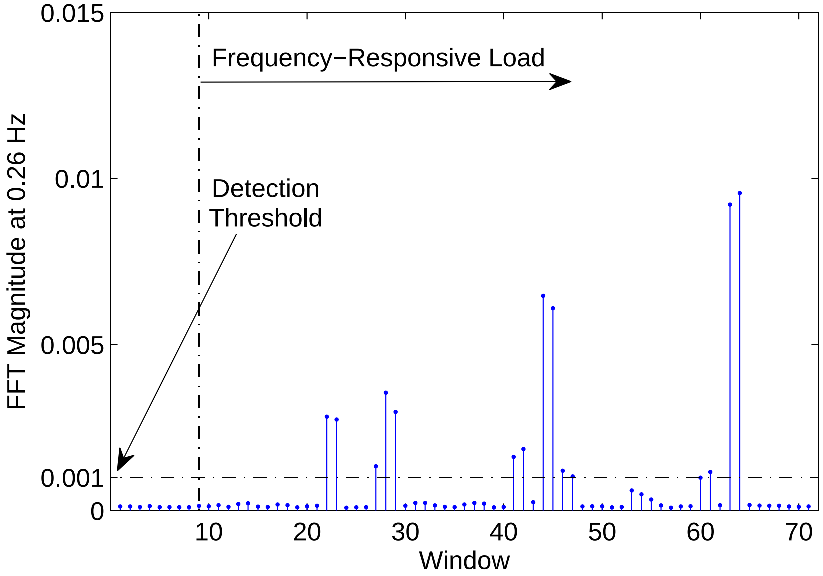


Figure 3.6: The W-FFT coefficient of a frequency-responsive load at attack frequency 0.26 Hz versus the W-FFT sliding windows for each home at bus 23.

3.3 Detection Based on Both Input and State Signals

In this section, we examine the possibility of detecting anomalies when there is access to both input and state signals. We show that the additional information that is provided by the state signal can particularly help in distinguishing between the adversarial action, e.g., D-LAA and benign action, e.g., frequency-responsive load.

The analysis in this section is in time-domain; and Cross-Correlation (CC) is the main mathematical tool [12]. Since we are interested in detecting anomalies in real-time, we use the Windowed Cross-Correlation (W-CC) method. Analogously to Section 3.2.2, three parameters of sampling rate, window size, and detection threshold can affect the analysis performance.

Suppose we set the sampling rate to 100 sec, window size to 200 sec, and detection threshold to 0.05. The results for the W-CC analysis of the load and frequency signals are shown in Figs. 3.7 and 3.8. For each W-CC sliding window, only the zero-lag cross-correlation coefficient is shown. Unlike in Figs. 3.5 and 3.6, where compromised loads

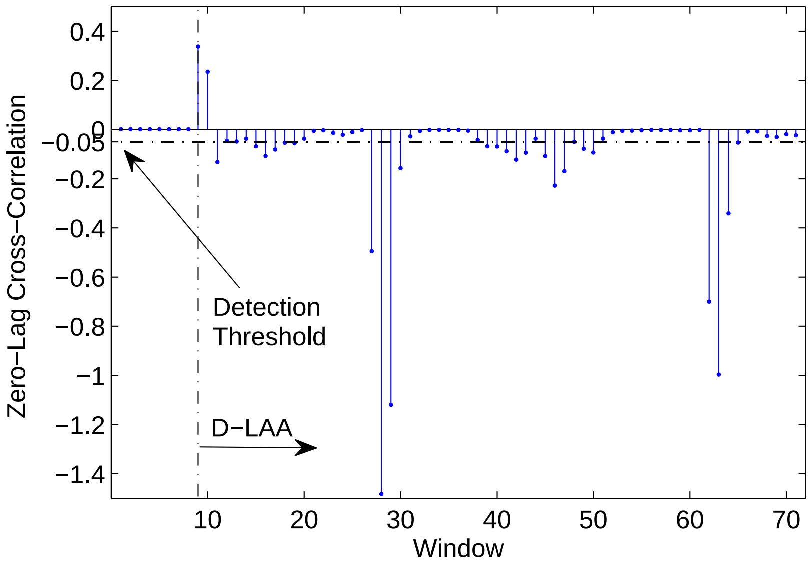


Figure 3.7: The zero-lag W-CC coefficient between the compromised load and frequency signals versus the W-CC sliding windows of a home at bus 19.

and frequency-responsive loads create similar coefficients, here, one can easily distinguish D-LAAs from frequency-responsive loads. Specifically, the zero-lag coefficients are *negative* for a compromised load under D-LAA and *positive* for a frequency-response load.

Recall from Remark 4 that the frequency-domain analysis in Section 3.2.1 requires the reading interval of the smart meter to be two seconds or less. Next, we examine the impact of smart meter time-resolution on detecting the correlations between the load and frequency signals. The results are shown in Fig. 3.9 for the zero-lag W-CC coefficient between a compromised load signal and the frequency signal of a home at bus 19. We can see that the magnitude of the correlation coefficients attenuate quickly as we lower the smart meter time-resolution.

Remark 6 *It appears that the need for high resolution smart meters does not depend on the method of detection, whether it is in time or frequency domain. This is an important observation because in practice most smart meters do not support high resolution readings. In fact, the need for such frequent meter readings has not been raised yet. In this regard, the*

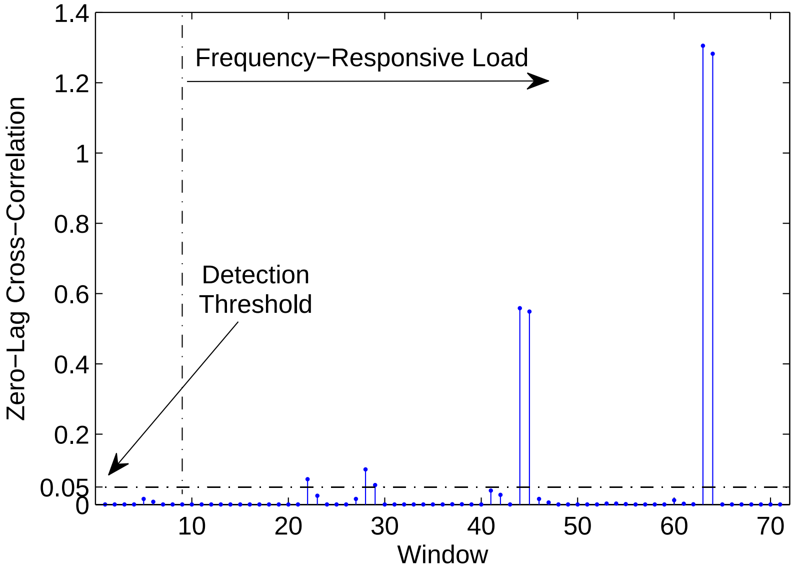


Figure 3.8: The zero-lag W-CC coefficient between the frequency-responsive load and frequency signals versus the W-CC sliding windows of a home at bus 23.

problem of detecting anomalies in this chapter appears to be one of the first smart meter data applications that can justify second-by-second or higher time-resolutions for smart meters.

Remark 7 *For the analysis in this chapter, it was implicitly assumed that the D-LAA sensor bus is the same as the D-LAA victim bus, i.e., $v = s$. In other words, the feedback on system frequency is measured at the same bus that the potential compromised load is located. However, in general, the sensor bus and the victim bus may not be the same in a D-LAA, see [4, 6]. Accordingly, if the location of the sensor bus is* unknown*, selecting the right frequency signal to be used as the base for cross-correlation analysis could be challenging. Another challenge in this case is to select the detection threshold for the CC and W-CC algorithms for each specific sensor bus location, even if such location is known.*

The challenges highlighted in Remark 7 suggest that, even if the frequency signals are available, one may not rely only on a time-domain cross-correlation analysis. It is more desirable if a frequency-domain analysis of the load signal is combined with a time-domain

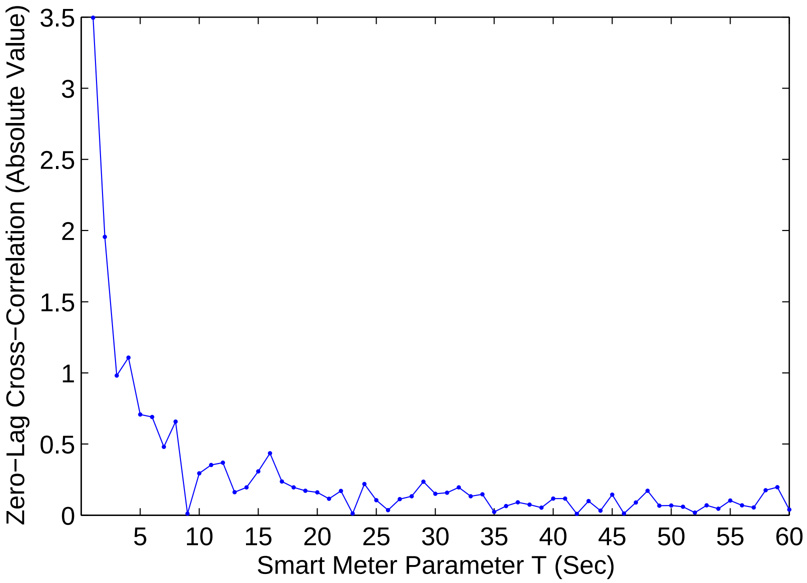


Figure 3.9: The zero-lag CC coefficient between the compromised load and frequency signals versus smart meter time-resolution for a home at bus 19.

analysis of the load and frequency signals. In fact, in addition to the concerns in Remark 7, since the performance of both the frequency-domain and time-domain detection methods are highly sensitive to the choices of parameters such as the detection threshold, a proper combination of the two methods may lead to a more accurate and robust detection.

Chapter 4

Location Identification Approaches

Consider a power system such as the one in Fig. 2.1. Suppose some fault(s) and/or attack(s) have affected a subset of inputs, such as the power generation level of certain generators and/or the power consumption level of certain loads, putting the system at the margin of instability. Suppose *the presence of the fault or attack has already been detected* using a frequency-domain data-driven analysis, such as the one in [5] or Section 3.2. That is, for a given threshold μ, the following expression holds:

$$\exists\, i : |Y_i(j\omega)| \geq \mu, \tag{4.1}$$

where Y_i is the Fourier Transform of the ith entry of the output signal y in (2.11). Thus, the fault/attack frequency ω^* is known:

$$\omega^* = \arg\max_{\omega}\, |Y_i(j\omega)|. \tag{4.2}$$

Parameter μ can be obtained from historical data, c.f. [5].

The next step is to answer the following two questions:

- *How many* power system inputs are affected?

- *Which* power system inputs are the ones that are affected?

We seek to answer both questions by using power system measurements. Only those measurements that capture frequency information around ω^* are of potential use. Such measurements are often provided only by advanced power system sensors such as PMUs that are used to monitor/estimate the *states* of the power system, where the reporting rate is a fraction of a second [56]. Note that, traditional SCADA systems do *not* support the reporting rate needed for this type of analysis, since their reporting rate is in the order of minutes.

4.1 Location Identification of Destabilizing Anomalies

In this section, we propose a novel optimization-based approach to identify which power system input(s), i.e., generators or loads, are affected by destabilizing faults and/or attacks.

4.1.1 Baseline Time-Domain Approach

Based on the existing literature, a somewhat standard approach to solve the destabilizing fault/attack location identification problem is to combine an *unknown input observer* (UIO) with any detection method, such as the one in [5] or Section 3.2. From [32, Definition 1], an UIO is defined for the system in (2.11). Its goal is to have estimation error vector approach zero asymptotically, despite the presence of the unknown input in the system. Different approaches can be used to design an UIO, e.g., see [72]. In principle, all approaches essentially seek to collect a time series of measurements $\hat{y}(t)$ from field sensors over a time interval D, and then reconstruct the unknown input signal $u^c(t)$ so as to minimize the residual observation error:

$$\|\hat{y}(t) - y(t)\|_2 \,, \tag{4.3}$$

subject to the power systems state space equations in (2.11) as constraints. Ideally, (4.3) approaches zero asymptotically. Note that, the constraints must hold at any time instance $t \in D$.

Once the UIO problem is solved and the unknown input signal $u^c(t)$ is reconstructed in time-domain, then one can identify the location(s) of the power system inputs that are affected by destabilizing fault/attack using, for example, the Fourier Transform of $u^c(t)$, see [5]. Accordingly, the set of affected power system inputs, denoted by $\mathcal{K}$, is obtained as:

$$\mathcal{K} = \{i \in \mathcal{B} \ | \ |U_i^c(j\omega^*)| \geq \mu\}. \tag{4.4}$$

4.1.2 Proposed Frequency-Domain Approach

The first fundamental step in our proposed approach is to transform the power system dynamics under destabilizing fault/attack in (2.11) from time-domain to frequency-domain. This can be done by applying the Fourier Transform to the model in (2.11) as follows:

$$E(j\omega\, X(j\omega) - x_0) = AX(j\omega) + BU^c(j\omega), \tag{4.5}$$

$$Y(j\omega) = CX(j\omega), \tag{4.6}$$

where x_0 denotes the power system's initial states in time domain. From (2.9), x_0 is related to y_0, i.e., the power system's initial outputs in time domain through $y_0 = Cx_0$.

We propose to identify destabilizing fault/attack location(s) by solving the following optimization problem:

$$\underset{X(j\omega^*),Y(j\omega^*),U^c(j\omega^*),x_0,I}{\textbf{minimize}} \quad \left\|\hat{Y}(j\omega^*) - Y(j\omega^*)\right\|_2 \tag{4.7a}$$

subject to

$$E(j\omega^*\, X(j\omega^*) - x_0) = AX(j\omega^*) + BU^c(j\omega^*), \tag{4.7b}$$

$$Y(j\omega^*) = CX(j\omega^*), \tag{4.7c}$$

$$\hat{y}_0 = Cx_0, \tag{4.7d}$$

$$\sum_{i\in\mathcal{B}} I_i = |\mathcal{K}|, \tag{4.7e}$$

$$|U^c(j\omega^*)| \leq \operatorname{diag}(I)\, U^{\max}, \tag{4.7f}$$

where the variables X, Y, and U^c are complex numbers, variable x_0 is scalar, and variable I is binary and defined as

$$I_i = \begin{cases} 1 & i \in \mathcal{K}, \\ 0 & i \notin \mathcal{K}. \end{cases} \tag{4.8}$$

The entry of I corresponding to location $i \in \mathcal{B}$ indicates whether or not the power system input i is a fault/attack location. The characteristics of problem (4.7) are as follows.

First, suppose we drop I as variable and also drop (4.7e) and (4.7f) as constraints. The remainder of the optimization problem in (4.7a)-(4.7d) is intended to reconstruct the frequency-spectrum of the unknown input signal $u^c(t)$, but *only* at frequency ω^*. Here, we make no effort in reconstructing the unknown input signal $u^c(t)$ at frequencies which are not ultimately of interest to the destabilizing fault/attack location identification problem. As we will see in Section 4.1.4, this will not only drastically reduce the computation complexity and thus the delay in identifying the fault/attack location(s), but it also will enhance the design performance, in the sense that one can now identify the destabilizing fault/attack location(s) with fewer number of sampled measurements.

Second, the binary decision making framework in optimization problem (4.7) eliminates the need to separately apply the data-driven method in [5] or Section 3.2, unlike in the case of the baseline time-domain approach in Section 4.1.1. Here, we assume that the number of affected power system input(s) is given, i.e., the cardinality of set $\mathcal{K}$, denoted by $|\mathcal{K}|$, is known. Accordingly, in (4.7e), we make sure that exactly $|\mathcal{K}|$ entries of vector I are non-zero. This assumption will be relaxed later in Section 4.1.3, where we develop an algorithm for the case where the number of affected location(s) is unknown.

Third, as for constraint (4.7f), it forces the frequency spectrum of the reconstructed unknown input signal $u^c(t)$ at each location i to have no signature at the fault/attack frequency ω^*, unless such input is indeed identified as a fault/attack location, i.e., $I_i = 1$. Notation $\operatorname{diag}(I)$ indicates a diagonal matrix with its diagonal entries being equal to the entries of vector I. The upper bound vector parameter $U^{\max}$ includes sufficiently large numbers in its entries. It can be obtained empirically.

One can explain the feasible set of optimization problem (4.7) by examining its constraints. First, consider constraint (4.7d). This constraint specifies the initial state of the power system based on the initial output measurements. Since a destabilizing fault or attack may affect only the system inputs but not the system outputs, from (2.10), $\hat{y}_0$ can directly be obtained from any given x_0. Therefore, (4.7d) always results in a solution for x_0. Next, consider constraints (4.7e) and (4.7f). Any arbitrary choice of I that satisfies constraint (4.7e) would result in a feasible solution for $U^c(j\omega^*)$ in constraint (4.7f). Finally, given the feasible solutions for both x_0 and $U^c(j\omega^*)$, constraints (4.7b) and (4.7c) simply provide the evolution of system states and outputs from the initial state and inputs according to the system model in (2.11). Hence, corresponding to the obtained feasible solutions of x_0 and $U^c(j\omega^*)$, there always exist solutions for $X(j\omega^*)$ and $Y(j\omega^*)$. Therefore, we can conclude that problem (4.7) always has a feasible solution. Of course, the extent of the accuracy of such feasible solutions depends on how small one can make the residual error $\hat{Y}(j\omega^*) - Y(j\omega^*)$ in the objective function of problem (4.7).

Although problem (4.7) is nonlinear and mixed-integer, it is tractable. In fact, once we slightly reformulate constraint (4.7f), we can present it as two separate linear inequality constraints on real and imaginary components. Therefore, the nonlinearity in (4.7) is solely due to the convex quadratic objective function. Accordingly, problem (4.7) is a standard mixed-integer least-square problem with linear constraints. Throughout this chapter, we solve optimization problem (4.7) using the MOSEK solver within the CVX software package [13]. CVX is installed in MATLAB to facilitate solving convex optimization problems.

Before we end this section, we shall point out that, an alternative option for the design in this section is to conduct a similar analysis as in the baseline design in Section 4.1.1, but this time in frequency-domain, and accordingly develop an UIO in frequency-domain. However, in principle, there is no advantage in doing so, as far as the reconstruction of the unknown input signal is concerned. Interestingly, we are *not* really concerned with the reconstruction of the unknown input signal. The UIO would be simply a middle step for us to ultimately identify the location(s) of power system inputs that are affected by destabilizing fault or attack. That explains why we took a rather different approach to tackle the problem, as it was described earlier in this section.

4.1.3 Proposed Algorithm

Problem (4.7) was formulated based on the assumption that the *number* of affected power system inputs, i.e., parameter $|\mathcal{K}|$, is *known* in advance. However, this is not always the case. In fact, the number of affected inputs is often unknown in practice. Accordingly, we propose to first conduct a sensitivity analysis of the objective function in (4.7a) with respect to parameter $|\mathcal{K}|$. We will then utilize the results to develop an algorithm to identify destabilizing fault/attack location(s), when the number of such location(s) is unknown.

Let $F(|\mathcal{K}|)$ denote the optimal objective value of problem (4.7) for a given $|\mathcal{K}|$. Next, we introduce a new definition.

Definition 8 (Sensitivity Function) *The difference between two consecutive optimal ob-*

jective values in (4.7)*, is referred to as the* sensitivity *with respect to* $|\mathcal{K}|$ *and defined as*

$$S(|\mathcal{K}|) = F(|\mathcal{K}|) - F(|\mathcal{K}| + 1), \quad |\mathcal{K}| = 1, ..., |\mathcal{B}| - 1. \tag{4.9}$$

The main properties of the above sensitivity function can be explained in a theorem, as it is presented next.

Theorem 9 (Properties of Sensitivity Function) *The sensitivity function,* $S(|\mathcal{K}|)$*, has the following two key properties:*

- *Non-negative Function:* $S(|\mathcal{K}|) \geq 0$
- *Non-increasing Function:* $S(|\mathcal{K}| + 1) \leq S(|\mathcal{K}|)$

proof: see Appendix A.2.

From the non-increasing property of the sensitivity function in Theorem 9, we can conclude that $S(1) \geq S(|\mathcal{K}|)$ for any $\mathcal{K}$. Accordingly, we can introduce a new definition for sensitivity.

Definition 10 (Normalized Sensitivity Function) *The normalized sensitivity function is defined as*

$$N(|\mathcal{K}|) = \begin{cases} 1 & |\mathcal{K}| = 0, \\ S(|\mathcal{K}|)/S(1) & |\mathcal{K}| \neq 0. \end{cases} \tag{4.10}$$

Corollary 11 (Identification Threshold) *For any arbitrary choice of parameter* ϵ*, there always exists a location set* $\mathcal{K}$ *for which the following conditions hold at the same time:*

$$\begin{cases} N(|\mathcal{K}| - 1) & > \epsilon, \\ N(|\mathcal{K}|) & \leq \epsilon, \end{cases} \tag{4.11}$$

where $0 < \epsilon < 1$ *is the identification threshold.*

```
Algorithm 1: Frequency-Domain Location Identification
Inputs: Measurements, Fault/Attack Frequency.
Parameters: System Model, Threshold $\epsilon$
Take Fourier Transform of $\hat{y}(t)$.
for $|\mathcal{K}| = 1$ to $|\mathcal{B}|$ do
    Solve optimization problem (4.7).
    if condition (4.11) holds then
        break
return $\mathcal{K}$
```

In Corollary 11, parameter ϵ specifies the residual error in state estimation. Set $\mathcal{K}$ is then selected through optimization to meet the limit on residual error that is set forth by parameter ϵ. The proposed frequency-domain location identification method, in presence of *uncertainty* about the number of affected power system inputs, is summarized in Algorithm 1.

According to Corollary 11, the number of affected inputs, i.e., $|\mathcal{K}|$, will increase by decreasing the value of ϵ. Decreasing ϵ does not change the fact that the inputs which are selected by Algorithm 1 are the ones that are most affected by the destabilizing fault or attack. For example, if decreasing ϵ results in selecting 3 instead of 2 inputs, then the third selected input is the third most affected input by the destabilizing fault or attack, e.g., due to the use of benign negative feedback but based on a state that is highly affected by the anomaly, see the illustrative example in Section 4.1.4. Nevertheless, one should be careful in selecting parameter ϵ, e.g., by using historical data of different fault and attack scenarios, so as to maintain a desirable sensitivity of the location identification system.

Note that, both detection and location identification would be implemented in real-time in practice in order to allow immediate and proper reaction in presence of an anomaly. Accordingly, we conduct our analysis in a window-based fashion, similar to the Windowed FFT (W-FFT), e.g., in [19], where the FFT is taken for a window of measurements. The size of the window in our case studies is 300 seconds.

4.1.4 Illustrative Example

Location Identification Performance

The performance of a fault/attack location identification algorithm can be evaluated in terms of two factors: the ability to find the location(s) that *are* affected; and the ability *not* to select the location(s) that are *not* affected. The latter is the ability to avoid false alarms. Therefore, we next introduce one metric, called *location identification accuracy* (LIA), that incorporates both factors:

$$\text{LIA (\%)} = \left[\frac{\text{\# of Correct} - \text{\# of Incorrect}}{\text{\# of Actual}}\right]^{+} \times 100. \tag{4.12}$$

The numerator in (4.12) is the total number of correctly identified affected input(s) minus the total number of benign input(s) that are incorrectly identified as affected. The denominator is the true total number of the affected input(s). This fraction is always less than one. Using operator $[x]^{+} = \max\{x, 0\}$, LIA is always between zero and one, or between 0% and 100%. As an example, suppose the power system is under a multi-point destabilizing attack where four power system inputs are affected. Suppose a location identification algorithm is applied, and it correctly identifies three of the four affected inputs. Suppose the algorithm also incorrectly identifies a benign input as affected. In that case, the numerator is 3 - 1 = 2 and the denominator is 4. Accordingly, LIA is obtained as 50%.

Again consider the power system under destabilizing attack in the illustrative example in Section 2.3.1. Suppose all buses are equipped with measurement devices, such as PMUs. Also, suppose the number of affected inputs (two) is known in advance. The performance, in terms of LIA, of the time-domain versus frequency-domain approaches are compared in Fig. 4.1(a). The x-axis is the time sampling rate of sensors. We can see that the LIA for the proposed frequency-domain method reaches 100% at only 0.8 Hz. This is in fact the same sampling rate that is required to detect the fault/attack in this example,

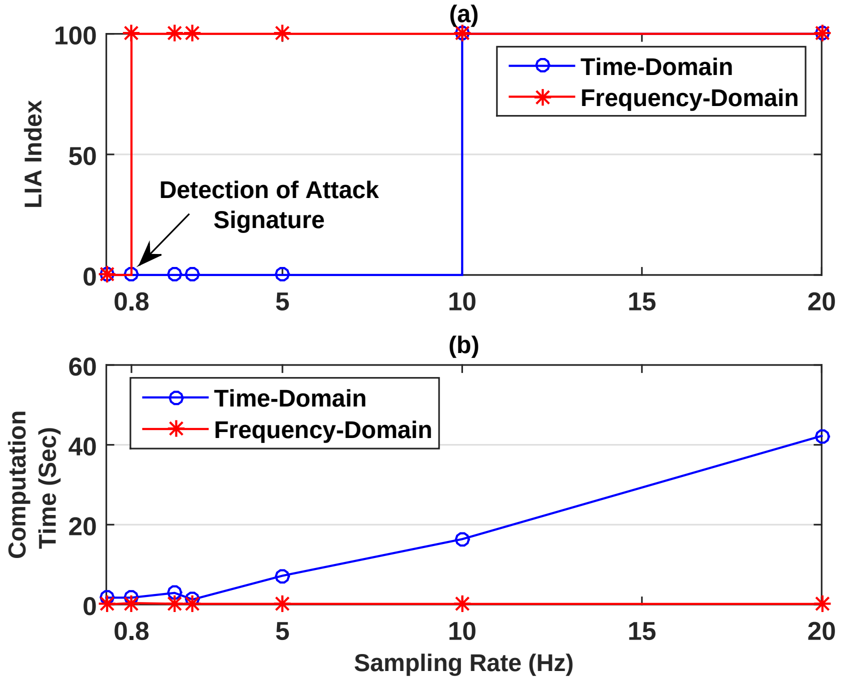


Figure 4.1: Comparing the performance of time-domain and frequency-domain location identification methods: a) LIA index; b) computation time.

see Section 2.3.1. In contrast, the time-domain method has a zero LIA all the way up to 10 Hz.

Another important performance metric is computation time, i.e., the time needed by the algorithm to identify the location(s) of faults/attacks. This is shown in Fig. 4.1(b). We can see that, the time-domain method needs at least 17 seconds before it can reach 100% accuracy. In contrast, it takes less than 1 second for the frequency-domain method to reach 100% accuracy. The exact computation platform is not a major factor; of importance is rather the *relative* computation time. That being said, the computation platform in this example was an Intel Core i7-2600 with 3.4 GHz CPU and 8 GB memory.

Ability to Identify the Number of Affected Inputs

Next, suppose all buses, except for bus 7, are equipped with sensors that take two samples per second. Suppose $\mu = 1.5$, which allows detecting the presence of the

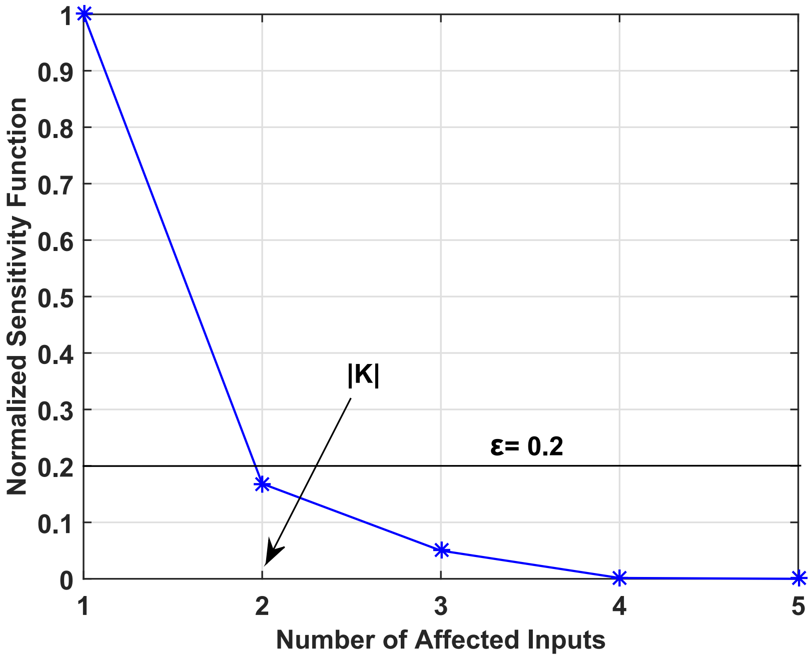


Figure 4.2: Identifying number of affected inputs, i.e., $|\mathcal{K}|$ using Algorithm 1.

fault/attack by examining the frequency-spectrum of the measurements at bus 6, as it was previously shown in Fig. 2.3. Now, suppose the identification threshold is $\epsilon = 0.2$, the unknown location(s) of affected power system inputs are identified using Algorithm 1. The results are shown in Fig. 4.2, where $N(|\mathcal{K}|)$ is plotted versus $|\mathcal{K}|$. The algorithm stops in this case at $|\mathcal{K}| = 2$, which is associated with solution $I = [0\,0\,1\,0\,0\,1]$. That is, $\mathcal{K} = \{6, 9\}$, which is exactly the correct locations of the attacks. Therefore, LIA $= 100\%$, despite not knowing the number of affected inputs.

Finally, the outcome of running Algorithm 1 for different choices of parameter ϵ is shown in Table 4.1. If $\epsilon = 1$, then only input 6 is identified, which is one of the two affected inputs. If $\epsilon = 0.2$, then inputs 6 and 9 are identified. This is the ideal result, because inputs 6 and 9 are the exact two affected inputs. As we keep decreasing the value of ϵ, inputs 6 and 9 will continue to be identified as the affected inputs; however, additional benign inputs will be added to set $\mathcal{K}$, which degrades the LIA.

Table 4.1: Impact of parameter ϵ on the performance of Algorithm 1.

ϵ	$\lvert\mathcal{K}\rvert$	$\mathcal{K}$	LIA
1	1	$\{\mathbf{6}\}$	50%
0.2	2	$\{\mathbf{6},\mathbf{9}\}$	100%
0.1	3	$\{\mathbf{6},\mathbf{9},7\}$	50%
0.01	4	$\{\mathbf{6},\mathbf{9},7,5\}$	0%
0.0001	5	$\{\mathbf{6},\mathbf{9},7,5,4\}$	0%
$\simeq$0	6	$\{\mathbf{6},\mathbf{9},7,5,4,8\}$	0%

Practical Implementation Challenges

Recall that in practice, W-FFT is used for detection and location identification instead of continuous Fourier transform. Sampling frequency is one of the important factors in accuracy of W-FFT. We seek to examine impact of sampling frequency on residual error of observation in (4.7) for the illustrated example in Section 2.3.1. Results are shown in Fig. 4.3 for different W-FFT's sampling frequencies. According to this figure, observation accuracy would be increased by increasing the sampling frequency of W-FFT. Therefore, not only PMU sampling time resolution, but also sampling frequency parameter of W-FFT affects on accuracy of frequency-domain location identification approach.

4.2 Hierarchical Approach

One possible application of the methodology developed in Section 4.1 is in WAMS to conduct fault and attack location identification in a *hierarchical* fashion. Consider a typical WAMS data collecting and data processing network, as in Fig. 4.4. In practice, it is divided into several *areas*. Multiple PMUs are often installed in each area, providing synchrophasor measurements at high resolutions, e.g., with 30 readings per second [56]. The PMUs in each area are connected to a Phasor Data Concentrator (PDC). PDCs are then connected to the control center. Applications of synchrophasors include state estimation, parameter identification, and model validation [56].

The focus in this section is on answering the following question: *How can we*

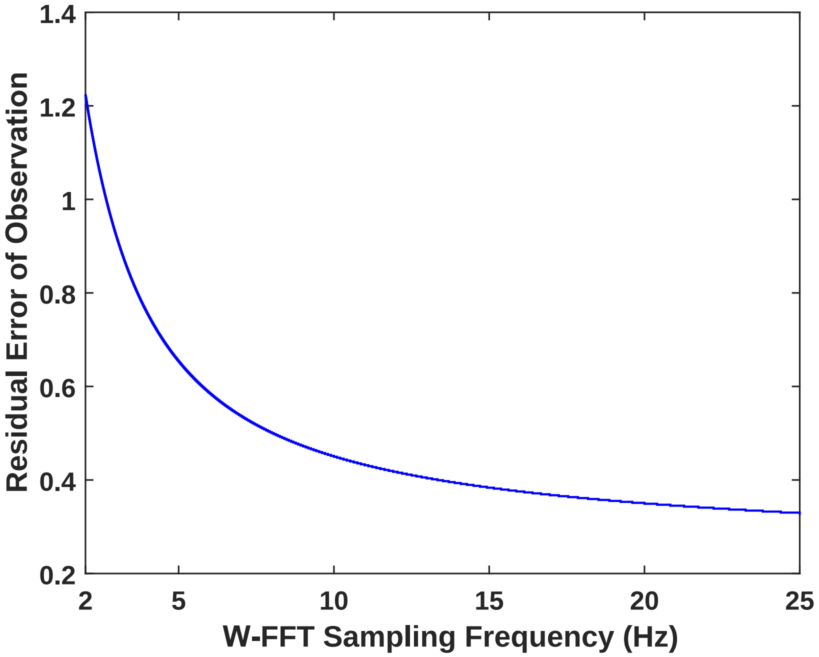


Figure 4.3: Impact of W-FFT's sampling frequency on location identification accuracy.

integrate a fault/attack location identification mechanism into a typical WAMS? There are at least two main challenges to address. First, any such mechanism is preferred to be *hierarchical* to fit into the multi-level structure of WAMS networks. Second, any such mechanism must be *light-weight* in its computational burden so as to have minimal overhead on PDCs and their existing data processing tasks.

4.2.1 System Configuration

Suppose the set of all buses in each area within an n-area system is denoted by $\mathcal{A}_a$, for $a = 1, ..., n$. The buses in each area are classified as *internal* versus *boundary*. An internal bus does not have any direct line to a bus outside its own area. A boundary bus has at least one direct line to a bus in another area. Two areas are *neighbors* if there is at least one direct line between their boundary buses. PDCs are configured to collect data from not only the PMUs in their own area; but also the PMUs on boundary buses in their neighboring areas. For example, the PDC corresponding to $\mathcal{A}_5 = \{25, 26, 27, 28, 29, 37, 38\}$

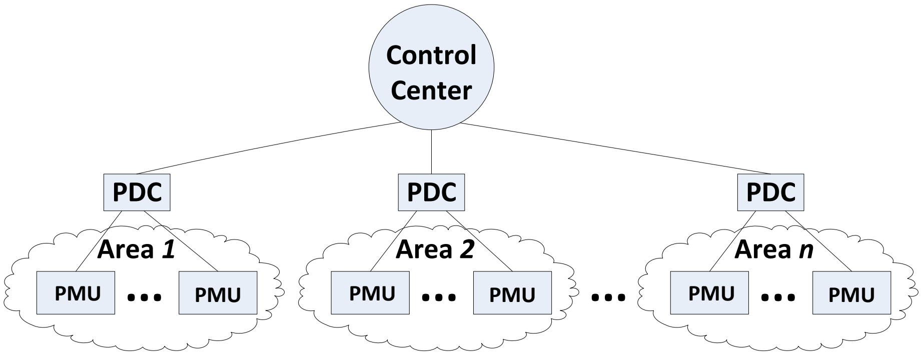


Figure 4.4: The hierarchical structure of a typical synchrophasor network.

in Fig. 4.5, collects PMU data from these buses: $\{2, 17, 25, 26, 27, 28, 29, 37, 38\}$. We refer to this latter set of buses as the *subsystem* of area $\mathcal{A}_5$.

4.2.2 Hierarchical Identification

The proposed hierarchical destabilizing fault/attack location identification algorithm is given in Algorithm 2. The central idea in this algorithm is to keep track of three sets, denoted by $\mathcal{P}$, $\mathcal{C}$, and $\mathcal{N}$. They specify the *previous*, *current*, and *next* areas to run Algorithm 1. In this regard, Algorithm 2 can be interpreted as an intelligent mechanism to hierarchically run Algorithm 1 across different areas in the system. In addition to breaking down the original *large system-wide* fault/attack identification problem into several *small area-level* identification tasks, Algorithm 2 is also capable of accurately identifying the fault/attack locations by examining only a small subset of the areas, see Section 4.2.3. Set $\mathcal{T}$ keeps track of the identified location(s) as Algorithm 2 examines different areas.

The *Initial area* to examine, i.e., the starting point for Algorithm 2, is area $\mathcal{A}_s$, which is obtained as

$$s = \arg\max_{a} \operatorname*{maximize}_{i \in \mathcal{A}_a} |\hat{Y}_i(j\omega^*)|. \tag{4.13}$$

Here, we start with the area that has detected the strongest fault/attack signature in the frequency spectrum.

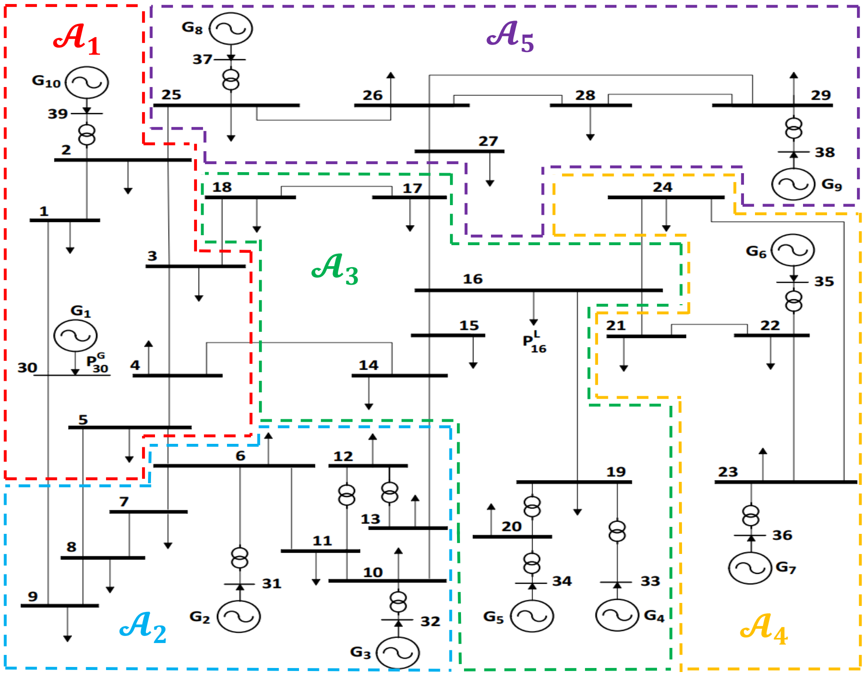


Figure 4.5: The IEEE 39 bus test system is partitioned into five areas.

The operation of Algorithm 2 is as follows. The outer loop in lines 4 to 13 is executed until the algorithm stops. The inner loop in lines 6 to 10 runs Algorithm 1 in all areas within set $\mathcal{C}$. The next areas to run Algorithm 1 are decided in line 9 based on the boundary buses that are identified as fault/attack locations. Only the internal buses that are identified as fault/attack locations are added to set $\mathcal{T}$ in line 10. From lines 11 and 12, set $\mathcal{C}$ is updated to identify a new set of areas in the next round of the algorithm. The algorithm ends if set $\mathcal{C}$ is empty, i.e., there is no need to examine any further area.

It is worth clarifying that the accuracy of the location identification approach can reach 100%, i.e., LIA=100%, when it is implemented in a centralized fashion, as long as parameter ϵ is selected properly. However, there is no similar guaranty for the hierarchical approach to achieve 100% accuracy. This is due to the fact that the hierarchical approach involves *model decomposition* and such model decomposition creates *additional residual error* in the input observation aspect of the proposed design. Of course, as we will show in

Algorithm 2: Coordination Algorithm

```
Inputs: Measurements Grouped into Subsystems.
Parameters: System Model.
Initialization: $\mathcal{P} = \{\}, \mathcal{C} = \{\mathcal{A}_s\}, \mathcal{T} = \{\}$.
repeat
    $\mathcal{N} \leftarrow \{\}$
    for any area $\mathcal{A}_i \in \mathcal{C}$ do
        Run Algorithm 1 on subsystem of $\mathcal{A}_i$ to obtain $\mathcal{K}$.
        for any boundary bus $j \in \mathcal{K} \setminus \mathcal{A}_i$ do
            $\mathcal{N} \leftarrow \mathcal{N} \cup \{\mathcal{A}_a | j \in \mathcal{A}_a\}$
        $\mathcal{T} \leftarrow \mathcal{T} \cup (\mathcal{K} \cap \mathcal{A}_i)$
    $\mathcal{P} \leftarrow \mathcal{P} \cup \mathcal{C}$
    $\mathcal{C} \leftarrow \mathcal{N} \setminus \mathcal{P}$
until $\mathcal{C} = \{\}$
return $\mathcal{T}$
```

our case studies in Section 4.2.3, the performance loss is not significant, such as at only 6%.

4.2.3 Case Study

Consider the IEEE 39-bus test system in Fig. 4.5. The parameters in (2.8), and the loads at load buses, are set as in [1]. All generator and load buses are equipped with PMUs. The grid is partitioned into five areas. Without loss of generality, the destabilizing anomaly is assumed to be due to D-LAAs [6]. We simulated 200 different D-LAA scenarios. In scenarios 1 to 80, 81 to 120, 121 to 160, 161 to 180, 181 to 200, the adversary compromised one, two, three, four, and five power system inputs, respectively. The number of affected power system inputs is assumed to be unknown to our algorithm.

We aim to compare the *centralized* location identification approach of running Algorithm 1 for the entire power system versus the *hierarchical* approach in Algorithm 2. The results are shown in Fig. 4.6(a). Note that, the choice of victim bus(es) in all ten scenarios is the same within each test group, but the choices of anomaly feedback gains are different in each scenario. In total $20 \times 10 = 200$ cases are examined. In order to save space in presenting the results, we grouped together the scenarios with the same choice of victim

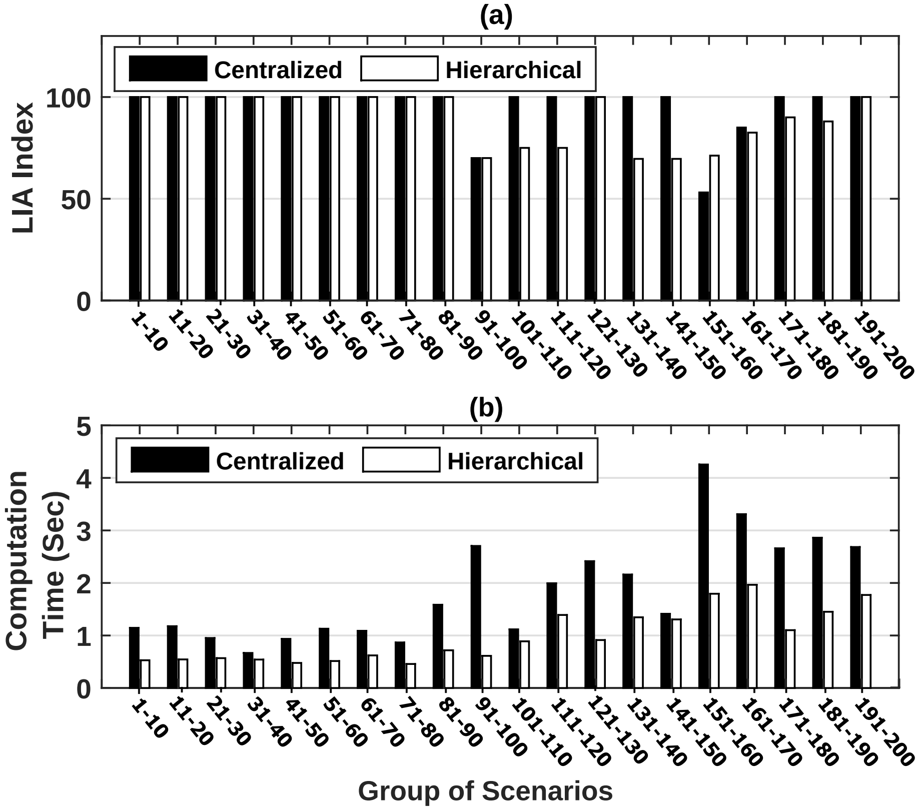


Figure 4.6: Comparing the centralized versus hierarchical location identification methods across 200 different test scenarios: a) Average LIA; b) Average computation time.

buses and showed their average results in one bar, thus showing a total of 20 bars for each design setup. We can see that the hierarchical approach can work almost as good as the centralized approach. While the average LIA across the 200 test scenarios is 95% for the centralized approach, it drops only by 6% to 89% in the hierarchical approach. In return, the hierarchical approach provides *a much better performance with respect to computation time*, as shown in Fig. 4.6(b). On average, the computation time for the hierarchical approach is almost half of that for the centralized approach, i.e., 0.98 second versus 1.86 second. Recall from Section 4.2 that the hierarchical approach does not guarantee 100% accuracy due to the presence of additional residual observation error caused by model decomposition.

Of interest is the perfect 100% LIA for both centralized and hierarchical designs on the *first 80* scenarios in Fig. 4.6(a), i.e., the first 8 bars. Recall from the setup of our

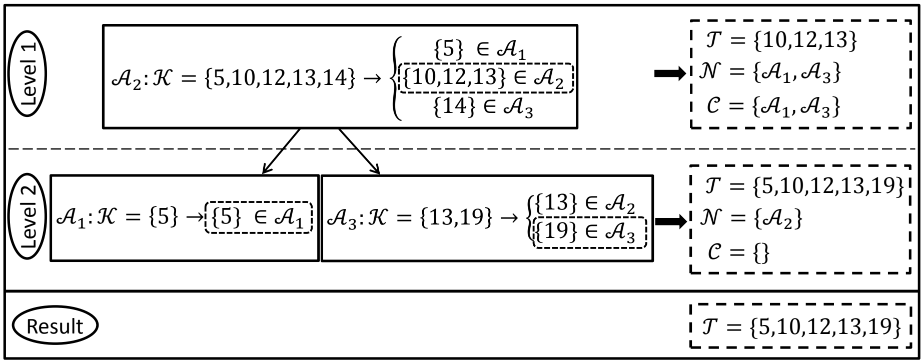

Level 1
$\mathcal{A}_2 : \mathcal{K} = \{5,10,12,13,14\} \rightarrow$
$\{5\} \in \mathcal{A}_1$
$\{10,12,13\} \in \mathcal{A}_2$
$\{14\} \in \mathcal{A}_3$
$\mathcal{T} = \{10,12,13\}$
$\mathcal{N} = \{\mathcal{A}_1, \mathcal{A}_3\}$
$\mathcal{C} = \{\mathcal{A}_1, \mathcal{A}_3\}$
Level 2
$\mathcal{A}_1 : \mathcal{K} = \{5\} \rightarrow \{5\} \in \mathcal{A}_1$
$\mathcal{A}_3 : \mathcal{K} = \{13,19\} \rightarrow$
$\{13\} \in \mathcal{A}_2$
$\{19\} \in \mathcal{A}_3$
$\mathcal{T} = \{5,10,12,13,19\}$
$\mathcal{N} = \{\mathcal{A}_2\}$
$\mathcal{C} = \{\}$
Result
$\mathcal{T} = \{5,10,12,13,19\}$

Figure 4.7: An example for step-by-step operation of the hierarchical approach.

case studies that, only one power system input is affected in each of these 80 scenarios. Accordingly, these are the cases that are more likely to occur in practice. The hierarchical approach improves the computation time significantly in all these 80 scenarios, without degrading the performance in location identification.

The step-by-step details of running Algorithm 2 for scenario number 200 is depicted in Fig. 4.7. Similar diagrams can be plotted for every other scenario. From Fig. 4.7, Algorithm 2 starts with $s = 2$, and by running Algorithm 1 on the subsystem of area $\mathcal{A}_2$. This results in identifying buses 5, 10, 12, 13, 14 as potential fault/attack locations. Buses 10, 12, and 13 are internal to area $\mathcal{A}_2$. Therefore, they are permanently added to set $\mathcal{T}$. However, buses 5 and 14 are boundary buses, as they belong to area $\mathcal{A}_1$ and area $\mathcal{A}_3$, respectively. Next, areas $\mathcal{A}_1$ and $\mathcal{A}_3$ are considered to run Algorithm 1. At the second level of the algorithm, running Algorithm 1 in area $\mathcal{A}_1$ results in identifying bus 5; and running Algorithm 1 in area $\mathcal{A}_3$ results in identifying buses 13 and 19. Bus 5 is internal to area $\mathcal{A}_1$ and bus 19 is internal to area $\mathcal{A}_3$. Therefore, they are added to set $\mathcal{T}$. Note that, bus 13 was already added to set $\mathcal{T}$ in the first level of the algorithm. We reach $\mathcal{C} = \{\}$ at this point. Therefore, the algorithm stops. The final set of identified fault/attack locations is $\mathcal{T} = \{5, 10, 12, 13, 19\}$.

4.3 Location Identification of Dynamic Load Altering Attacks

Consider the power system model under Dynamic Load Altering Attack in (2.16). This model can be rewritten as

$$\dot{\mathbf{x}} = (A + BA^p)\mathbf{x} + B(\mathbf{u}^n + \mathbf{u}^p), \tag{4.14}$$

where

$$A^p = \begin{bmatrix} 0 & 0 & -(D^L)^{-1}K^{LG} \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.15}$$

and the vector $\mathbf{u}^n$ denotes the control input under normal operation and $\mathbf{u}^p$ is the constant term for the proportional controller. The attack will be destabilizing if the matrix $(A+BA^p)$ has an eigenvalue whose absolute value is greater than 1 in discrete-time model of (4.14).

In this section, we propose a location identification method that directly estimates the K^{LG} matrix. Our method automatically determines which load buses are compromised and can distinguish between destabilizing and benign loads. Our method requires access only to synchronized measurements of the state vector $\mathbf{x}$, and does not require access to the control input $\mathbf{u}$. These state measurements are widely available in existing modern power systems through Phasor Measurement Units (PMUs).

4.3.1 Attack Identification Method

Our attack identification procedure has two steps. First we estimate the K^{LG} matrix using *dual state estimation.* This is a standard technique that applies the unscented Kalman filter (UKF) [71] to simultaneously estimate the entries of matrix K^{LG} and the system states $\mathbf{x}$. Unfortunately, the standard application of this technique does not work well for our problem. It is too slow computationally and has poor accuracy. So we introduce a novel rank-1 approximation which lets us effectively apply dual state estimation to our problem. Finally, once K_{LG} is estimated, we apply a thresholding procedure to identify the attacked buses.

Standard Dual State Estimation and Its Limitations

Dual state estimation is traditionally described using the system's discrete state equations, so we begin our presentation by discretizing (4.14) as

$$\mathbf{x}_{t+1} = (sA + sBA^p_t + I)\mathbf{x}_t + sB(\mathbf{u}^n_t + \mathbf{u}^p_t) + \epsilon. \tag{4.16}$$

The subscripts indicate the timestep, s is a scalar that represents the length of a time step, and $\epsilon \sim \mathcal{N}(0, Q^\epsilon)$ is an error term capturing both modeling and observation errors.

Now we describe how to estimate the A^p_t matrix. Recall that in the definition of A^p_t, the K^{LG}_t matrix is unknown and determined by the attacker; all other elements are statically known. In dual state estimation, we augment the original dynamical system's state variables to also include the elements of K^{LG}_t. The resulting augmented system is

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \operatorname{vec} K^{LG}_{t+1} \end{bmatrix} = \begin{bmatrix} sA + sBA^p_t + I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \operatorname{vec} K^{LG}_t \end{bmatrix} + \begin{bmatrix} sB & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{u}^n_t + \mathbf{u}^p_t \\ \mathbf{u}^m_t \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon^m \end{bmatrix}, \tag{4.17}$$

where

$$\epsilon \sim \mathcal{N}(0, Q^\epsilon), \quad \epsilon^m \sim \mathcal{N}\left(0, Q^{\epsilon^m}\right). \tag{4.18}$$

Here we have also introduced a new control input $\mathbf{u}^m_t$ with error ϵ_m. It is unobserved and controlled by the attacker. Specifically, the attacker uses $\mathbf{u}^m_t$ to manipulate the entries of K^{LG}_t, and hence A^p_t. The notation $\operatorname{vec} K^{LG}_t$ refers to the column vector constructed by stacking the columns of K^{LG}_t on top of each other.

Next we note that the control inputs $\mathbf{u}^n_t$, $\mathbf{u}^p_t$, and $\mathbf{u}^m_t$ are unobserved. A standard technique for modeling unobserved inputs is to replace them with random error terms. The true distribution of these random errors is unknown, but for computational convenience we assume they are normally distributed. In particular, we assume the control inputs are zero-mean Gaussians with covariance Q^n, Q^p, and Q^m respectively. Under these assumptions, we can rewrite the dualized system dynamics described in (4.18) as

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \operatorname{vec} K^{LG}_{t+1} \end{bmatrix} = \begin{bmatrix} sA + sBA^p_t + I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \operatorname{vec} K^{LG}_t \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon^{KL} \end{bmatrix}, \tag{4.19}$$

where

$$\begin{aligned} &\epsilon \sim \mathcal{N}\left(0, B(Q^n + Q^p) + Q^\epsilon\right), \\ &\epsilon^{KL} \sim \mathcal{N}\left(0, Q^m\right). \end{aligned}$$

Observe that the dynamical system described by (4.19) is nonlinear because the K^{LG} term appears in the definition of A^p_t. It is standard to solve systems of this form using the UKF [71]. We defer to the cited paper for details.

The UKF encounters two problems when run on (4.19). The first is that the problem is *underspecified*. The number of parameters we are trying to estimate (i.e. the number of entries in K^{LG}) grows as $O(|\mathcal{G}||\mathcal{L}|)$, but the size of the observed data (i.e. the size of $\mathbf{x}$) grows at the slower rate of $O(|\mathcal{G}| + |\mathcal{L}|)$. In general, underspecified problems are difficult to solve without introducing additional statistical assumptions. As the size of the power grid increases, the degree of underspecification increases, so we would expect this method to have low accuracy on large grids. The second problem is computational. At each time step, the UKF takes the inverse of a matrix whose dimension depends on the number of state variables. There are $O(|\mathcal{G}||\mathcal{L}|)$ states in (4.19), and so the runtime of this inversion is $O((|\mathcal{G}||\mathcal{L}|)^3)$. This poor scaling makes the standard method impractical to run on power systems with more than about 50 buses. These limitations of the standard method motivate our proposed rank-1 method, which we describe next.

The rank-1 method

In this method, we assume that the K^{LG} matrix has rank 1. We justify this assumption as follows. In a typical destabilizing attack, only a small number of buses are compromised and subject to positive feedback. For each of these compromised buses, there is a corresponding nonzero entry in the K_t^{LG} matrix. A basic fact of linear algebra is that the rank of a matrix is less than or equal to the number of nonzero entries in the matrix. Specifically, we have

$$\text{Rank}\{K_t^{LG}\} \leq \text{Non-zero entries in } K_t^{LG}. \tag{4.20}$$

Therefore, assuming that there are a small number of compromised buses is equivalent to assuming that K^{LG} has low rank.

Specifically, we assume that

$$K_t^{LG} = \mathbf{k}_t^L {\mathbf{k}^G_t}^{\mathsf{T}}, \tag{4.21}$$

where $\mathbf{k}_t^L$ and $\mathbf{k}_t^G$ are column vectors. Under this assumption, we can rewrite the standard method's dynamics from (4.19) as

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{k}_{t+1}^L \\ \mathbf{k}_{t+1}^G \end{bmatrix} = \begin{bmatrix} sA + sBA_t^p + I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{k}_t^L \\ \mathbf{k}_t^G \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \tag{4.22}$$

where

$$\begin{aligned}
\epsilon &\sim \mathcal{N}\left(0, B(Q^n + Q^p) + Q^\epsilon\right), \\
\epsilon_1 &\sim \mathcal{N}\left(0, Q^m + Q^{\epsilon_1}\right), \\
\epsilon_2 &\sim \mathcal{N}\left(0, Q^m + Q^{\epsilon_2}\right),
\end{aligned}$$

and the new attack matrix is

$$A_t^p = \begin{bmatrix} 0 & 0 & -(D^L)^{-1}{\mathbf{k}^L_t}^{\mathsf{T}}\mathbf{k}_t^G \\ 0 & 0 & 0 \end{bmatrix}.$$

This system remains nonlinear and is solved using the UKF.

The rank-1 method has improved statistical and computational performance. Statistically, there are only $O(|\mathcal{G}| + |\mathcal{L}|)$ parameters to estimate in the rank-1 method. This matches the size of the state vector $\mathbf{x}$, so the problem is no longer underspecified. We no longer expect statistical performance to degrade as the problem size increases. Computationally, the run time of each iteration of the UKF is only $O((|\mathcal{G}| + |\mathcal{L}|)^3)$. This is much faster than the $O((|\mathcal{G}||\mathcal{L}|)^3)$ required for the standard method.

Thresholding

Once the matrix K^{LG} is estimated, we apply a thresholding procedure to identify the attack. Define the function

$$f_t(i) = \sum_{j=1}^{n} K_t^{LG}(i,j), \tag{4.23}$$

to be the sum of the entries in the ith row of the K_t^{LG} matrix. This value is the total predicted attack on the ith bus in the power grid. Also define

$$\alpha_t = \arg\max_i |f_t(i)|, \tag{4.24}$$

to be the bus we predict has the most compromised load and so is under the heaviest attack. If $f_t(\alpha_t)$ is greater than some threshold τ, then we declare that the system is under attack at bus α_t. At this point, the system operator can take defensive measures such as isolating the bus from the system.

4.3.2 Simulation Results

In this section we compare the performance of the proposed method in Section 4.3.1 with a baseline approach that does not apply the rank-1 assumption from Section 4.3.1. We refer to the latter method as the *standard* method. In this section, we show that compared to the standard method, our proposed method can:

1. significantly lower the computation time;

2. significantly lower the identification error; and

3. better distinguish positive and negative feedback.

We begin with a qualitative demonstration of these facts, and then conclude with a quantitative demonstration.

Test Setup and Qualitative Results

All experiments in this section use a single randomly generated power grid with 20 generator and 20 load buses. We test on this relatively small grid size because the standard method that estimates a full rank K^{LG} matrix cannot scale to larger problems. On this size problem, a single iteration of our rank-1 method takes about 1 second, and a single iteration of the standard method takes about 1 minute. On a problem with 100 generators and 100 loads, a single iteration of our rank-1 method takes about 5 seconds, and a single iteration of the standard method takes over an hour. The computation advantage of our proposed method is evident.

We follow the *clusterSmallWorld* procedure for generating the power grid [73]. Note that, standard methods for generating random graphs do not exhibit the topological and electrical properties of real world power grids [29], but *clusterSmallWorld* was designed specifically for modeling real world power grids. An outline of the procedure is: First generate a random number of ring shaped grids with fewer than 10 buses each; Then randomly add connections between the buses until the average degree of each node is 4. To ensure the stability of the resulting system, scale matrix A so that its maximum eigenvalue is no greater than 0.999. This model generates realistically shaped power grids up to about 300 buses. Once the power grid has been generated, a load input vector, i.e., $\mathbf{u}_t$, is sampled from a Gaussian process truncated so that values are always non-negative.

The first experiment has 6 separate scenarios that test how the proposed method and the standard method perform in identifying 1, 3, and 5 compromised buses. In each case, the attack begins at time 0.1 seconds. Matrix A_t^p is selected such that $(A + BA_t^p)$ has maximum eigenvalue 1.05, ensuring that the attack destabilizes the system. Figure 4.8 shows the results. The proposed method clearly has better qualitative performance on this particular problem. Specifically, it identifies the compromised buses faster and more accurately.

To look carefully into how our proposed method can differentiate between benign and malicious loads, next we randomly selected a load i and generator j, then set the ith row and jth column of K^{LG} to -10. Recall that negative values of the K^{LG} matrix correspond to benign loads. The results are shown in Figure 4.9. Negative feedback does not destabilize the system, yet we are able to detect the feedback. The standard method (not shown) has

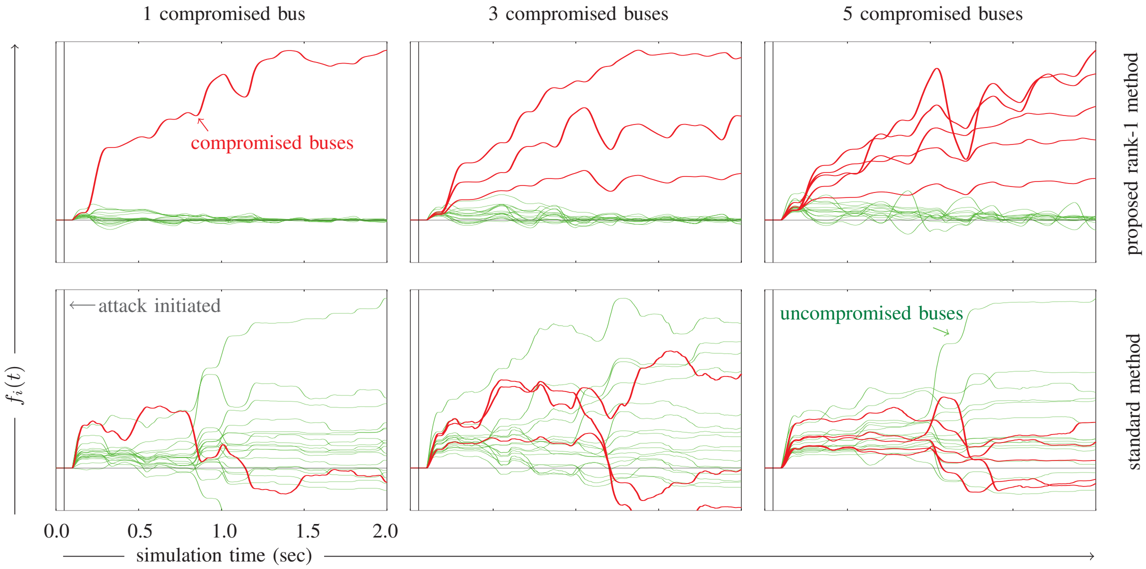


Figure 4.8: Each line in the figures above represents the predicted positive feedback of a particular load bus. Compromised buses are drawn in bold red, and uncompromised buses are drawn in thin green. For all times t before the attack begins, each bus i has $f_i(t)$ near zero. After the attack begins, the $f_i(t)$ deviate from zero. Our method is correctly identifying the attacked buses whenever the red lines are above the green lines. In the top row, we see that our rank-1 approximation of K^{LG} provides relatively accurate predictions even when the number of attacks increases and the rank-1 approximation is no longer true. In the bottom row, we see that the standard method has poor accuracy.

difficulty with this problem as it takes much longer for the standard method to converge.

Quantitative Results

We now explore the quantitative performance of our methods by measuring its performance on several power systems. We generated two sets of power grids, one with 20 generators and 20 loads (as in the previous section), and the other one with 100 generators and 100 loads. The standard method was run only on the smaller grid, again because it is computationally infeasible to run it on the larger one, and the proposed method was run only on both grids.

A major strength of both methods is that they experienced no false positives. We define a false positive to be the detection of an attack when no attack occurred. It does

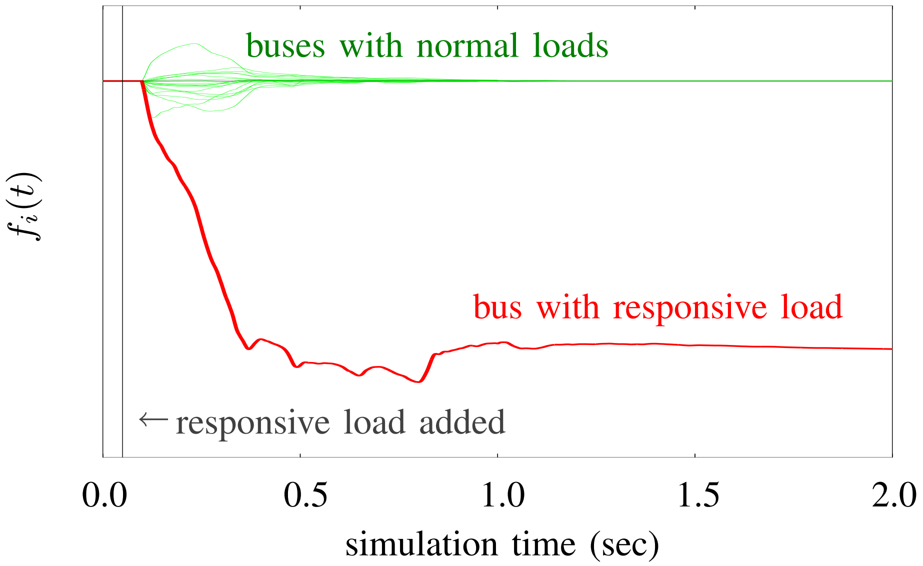


Figure 4.9: In this simulation, we added a benign frequency responsive load. Our rank-1 method is able to quickly identify the load with the responsive feedback. The corresponding value of $f_i(t)$ is negative because the feedback is negative.

not matter if the value of α_t is correct. When no attack is underway, the largest entries of the estimated K_t^{LG} are typically less than 10^{-6}. When an attack is underway, the largest values of the estimated K^{LG} skyrocket to well above 10^{-1}. Therefore, it is easy to set the threshold τ to avoid false positives.

Finally, we evaluate the method's *accuracy of identification*. We define the accuracy at time point t to be the fraction of α_t values that correctly predict the attacked bus. Figure 4.10 shows that the longer we wait to declare an attack occurs (i.e. the larger we set τ), the higher our accuracy is. In the case of the rank-1 method detecting a single attack, we observed 99% accuracy in under one second. The rank-1 method operating on the 200 bus system has much higher accuracy than the standard method operating on the significantly easier 40 bus system. The standard method's accuracy is little better than random guessing after two seconds.

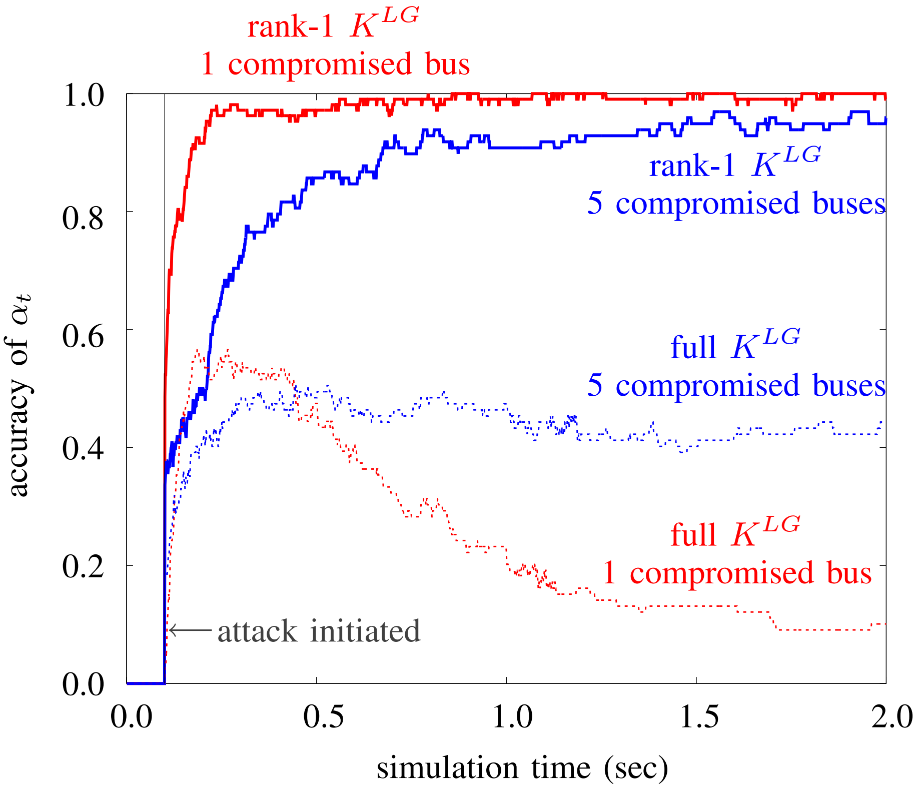


Figure 4.10: It takes only about a quarter of a second for our rank-1 method to detect a single attack with 99% accuracy. As the number of attacks increases, our rank-1 method takes longer to achieve high accuracy. With 5 compromised buses, the attack is detected with 95% accuracy by two seconds. This is fast enough to implement corrective actions. The full K^{LG} method has much worse accuracy no matter how many buses are compromised.

Chapter 5

Protection and Mitigation Schemes

We assume that each vulnerable load can be protected, e.g., by implementing reinforced security measures, but at some cost. The cost is due to adding hardware and software security components, whether at device level [41, Section 6.2] or at communication level [40, 76]. Such cost is incurred directly to utility companies and indirectly to end consumers. Accordingly, we propose an algorithm to determine the *minimum* amount of load that must be protected at each load bus in order to assure power system stability under destabilizing anomalies against the remaining unprotected vulnerable loads. Specifically, we address designing protection scheme against D-LAAs in this chapter.

Note that, besides protecting the load, there might also exist some compensators to counter-attack D-LAAs to keep the power system stable. This may include frequency-responsive loads or load protection mechanisms such as UFLS protection relays (see Section 2.3.2), as well as ancillary generation mechanisms that respond to under- or over- frequencies. All such compensators can be integrated into our analysis by adding their corresponding system dynamics to the state-space system model in (2.14). Once such state-space model is updated, the rest of the attack analysis in Section 2.2.1 as well as the protection scheme design approaches in this Section can still be applied similarly to the new system model.

5.1 Protection Problem Formulation

The foundation of the proposed protection mechanism is to protect enough vulnerable loads such that we can maintain the system in (2.14) stable. Specifically, we want to keep the poles of the system on the left-half complex plane even if all unprotected vulnerable loads are compromised. This requires formulating and solving a *non-convex pole placement optimization problem*, as we will explain in details next.

The stability of the closed-loop system (2.16) can be analyzed using the Linear Quadratic Lyapunov Theory that is overviewed in Appendix A.1. Specifically, the closed-loop system in (2.16) is stable if there exists a *symmetric positive semi-definite* matrix X such that

$$\left(A - B \begin{bmatrix} 0 \\ 0 \\ K^{LG} \end{bmatrix}^T \right)^T X + X \left(A - B \begin{bmatrix} 0 \\ 0 \\ K^{LG} \end{bmatrix}^T \right) < 0. \tag{5.1}$$

For each victim load bus v, let P_v^{LP} denote the potentially vulnerable but *protected* load. Note that, we have $0 \leq P_v^{LP} \leq P_v^{LV}$. Accordingly, the amount of unprotected vulnerable load at bus v is calculated as $P_v^{LV} - P_v^{LP}$. This puts an upper bound on the attack controller gain K_{vs}^{LG}. Specifically, we have

$$K_{vs}^{LG} \omega_s^{\max} \leq \left(P_v^{LV} - P_v^{LP} \right) / 2, \tag{5.2}$$

where $\omega_s^{\max}$ denotes the maximum admissible frequency deviation for generator s before its over or under frequency relays trip. The division by two on the right hand side is due to the fact that the compromised load $P_v^{LV} - P_v^{LP}$ must provide enough room to allow both over or under frequency fluctuations, e.g., see Fig. 2.6(c) and (f), before the attack can trip the frequency relays at generator s, e.g., see Fig. 2.6(f).

To design an efficient load protection plan against D-LAAs, we need to solve the following optimization problem:

$$\begin{aligned}
\textbf{minimize} \quad & \sum_{v \in \mathcal{V}} P_v^{LP} \\
\textbf{subject to} \quad & 0 \leq P^{LP} \leq P^{LV}, \\
& X \succeq 0, \\
& X = X^T, \\
& \text{Eqs. (5.1) and (5.2)}, \quad \forall v \in \mathcal{V},
\end{aligned} \tag{5.3}$$

where the variables are P^{LP}, K^{LG}, and X. Notation $\succeq$ indicates matrix positive semi-definiteness. Here, we seek to deploy the minimum total load protection that guarantees power system stability under D-LAA attacks against any unprotected vulnerable load when the frequency sensor is located at generator bus s. Problem (5.3) is a non-convex optimization problem due to the non-convex quadratic constraint in (5.1).

5.2 Solution Method

First, we note that the inequality constraint in (5.2) must hold as equality for any optimal solution of problem (5.3). This can be proved by contradiction. Note that, if at optimality, the constraint in (5.2) holds as strict inequality at a victim load bus v, then one can reduce P_v^{LP} and lower the objective function, thus, contradicting the optimality status. Therefore, K_{vs}^{LG} acts as a slack variable as far solving optimization problem (5.3) is concerned. Once P_v^{LP} is known, we have

$$K_{vs}^{LG} = \left(P_v^{LV} - P_v^{LP}\right) / \left(2\omega_s^{\max}\right). \tag{5.4}$$

Therefore, there are only two sets of variables in the optimization problem in (5.3), P^{LP} and X. They are *coupled* through the non-convex inequality constraint in (5.2). To tackle this non-convexity, we propose to solve problem (5.3) using the *coordinate descent* *method* [9, pp. 207]. The idea is to first take P^{LP} as a constant and solve problem (5.3) over X only:

$$\begin{aligned} \textbf{Minimize} \quad & \sum_{v\in\mathcal{V}} P_v^{LP} \\ \textbf{Subject to} \quad & X \succeq 0, \\ & X = X^T, \\ & \text{Eqs. (5.1) and (5.4)}, \quad \forall v \in \mathcal{V}, \end{aligned} \tag{5.5}$$

where the variables are the entries of matrix X. Here, the objective function could be *anything* because problem (5.5) is essentially a *feasibility* problem, c.f. [10, pp. 129]. Problem (5.5) can also be classified as a *semi-definite program* [10, pp. 168]. Next, we take X as a constant based on the solution of problem (5.5) and solve problem (5.3) over P^{LP} only:

$$\begin{aligned} \textbf{Minimize} \quad & \sum_{v\in\mathcal{V}} P_v^{LP} \\ \textbf{Subject to} \quad & 0 \le P^{LP} \le P^{LV}, \\ & \text{Eqs. (5.1) and (5.4)}, \quad \forall v \in \mathcal{V}, \end{aligned} \tag{5.6}$$

where the variables are the entries of vector P^{LP}. This procedure is repeated, leading to an iterative algorithm. As for the initial condition, we start with full protection, i.e., we initially set $P_v^{LP} = P_v^{LV}$ for all potential victim load buses v. Next, we continue improving the protection system by lowering the amount of protected load while maintaining the stability of the system using the Lyapunov criteria in (5.1). The convergence of the coordinated descent algorithm is guaranteed, c.f. [9, Proposition 2.5]. Note that, at each iteration, the total protected load either reduces or remains unchanged. Therefore, the iterations continue until either we find the exact optimal solution for (5.3) or we reach a stationary point that is sub-optimal. As we will see in Section 5.3, the optimality gap for the above algorithm is typically very small.

5.2.1 Protection System Design Under Uncertainty

For the analysis in Sections 5.1 and 5.2, it was implicitly assumed that the power system operator knows where the frequency sensor is deployed. That is, it knows the location of sensor bus s. However, this assumption may not always hold in practice. This creates uncertainty when designing the protection system. The key to tackle uncertainty is to design the protection system in a way that it is robust to any scenario for the location of the sensor bus. This can be done by solving the following optimization problem which is an extension of problem (5.3) across various sensor bus location scenarios:

$$\begin{aligned} \textbf{minimize} \quad & \sum_{v \in \mathcal{V}} P_v^{LP} \\ \textbf{subject to} \quad & 0 \leq P^{LP} \leq P^{LV}, \\ & X_s \succeq 0, \qquad \forall s \in \mathcal{S}, \\ & X_s = X_s^T, \quad \forall s \in \mathcal{S}, \\ & \text{Eqs. (5.1) and (5.4)}, \quad \forall v \in \mathcal{V}, \ \forall s \in \mathcal{S}, \end{aligned} \tag{5.7}$$

where the variables are P^{LP}, K^{LG}, and X_s for any $s \in \mathcal{S}$. Here, $\mathcal{S} \subseteq \mathcal{G}$ denotes the set of all potential locations for the sensor bus. Problem (5.7) can be solved similar to problem (5.3) using the coordinated descent method, see Section 5.2.

5.3 Case Studies

Consider the power system in Section 2.3.2. We would like to protect this system against closed-loop D-LAAs.

5.3.1 Known Sensor Bus Location

Suppose the sensor bus is located at bus $s = 33$ and this is known to the grid operator. The results for solving the protection system optimization problem in (5.3) in

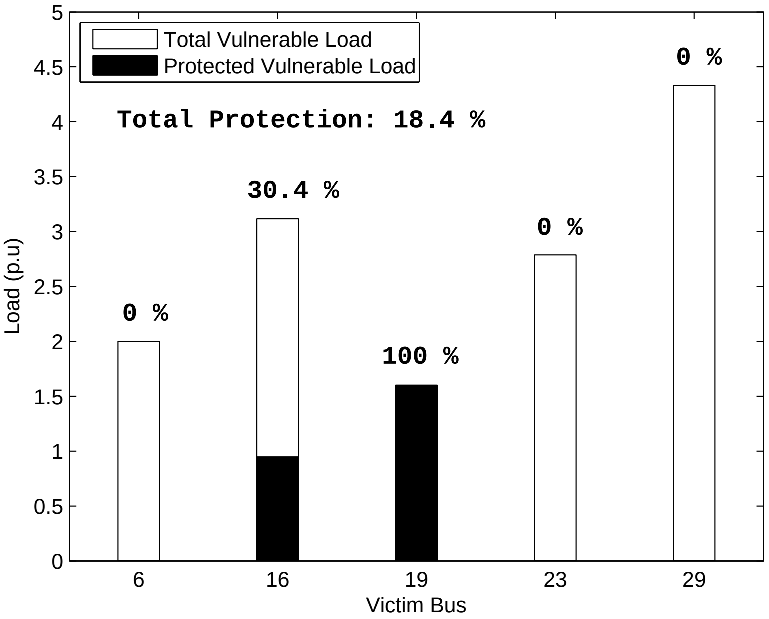


Figure 5.1: The optimal load protection scheme when the sensor location is known.

this case are shown in Fig. 5.1. We can see that as long as we fully protect the vulnerable load at bus 19 and protect 30.4% of the vulnerable load at bus 16, then no D-LAA with $s = 33$ can make the power system unstable. Note that, the total optimal load protection in this case is only 18.4% of the total vulnerable load in the system.

The operation of our proposed iterative algorithm to solve problem (5.3) is illustrated in Fig. 5.2. Recall from Section 5.2 that the algorithm starts from full protection and iterates until it reaches a stationary point at a much lower protection level. We can see that, the algorithm has indeed converged to the global optimal solution in this case after less than 45 iterations. Here, the global optimal solution is verified by conducting an exhaustive search based on an extensive root locus analysis.

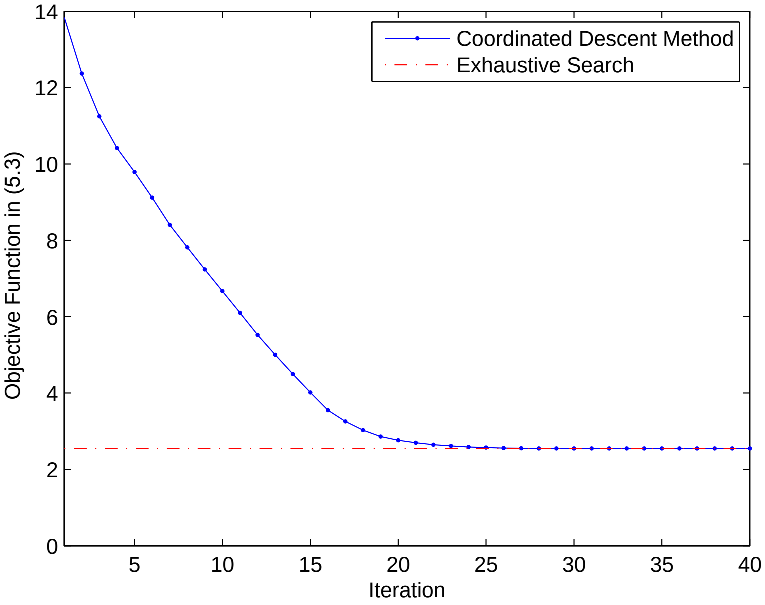


Figure 5.2: The iterative approach to solve the optimization problem in (5.3).

5.3.2 Unknown Sensor Bus Location

Next, consider the more practical scenario where the operator does *not* know where the attack frequency sensor is located. Accordingly, it needs to solve the extended optimization problem in (5.7). The results are shown in Fig. 5.3. As expected, the amount of vulnerable loads that need to be protected is higher in this case. However, such amount is still not too high and only at 26.8% of the total vulnerable load in the system. We can see that the uncertainty about the attack sensor location can be tackled by slightly adjusting the load protection plan, where we also protect 29.3% of the vulnerable load at bus 29. Interestingly, such protection allows some decrease, from 30.4% to 26.9%, in the level of vulnerable load that must be protected at bus 16.

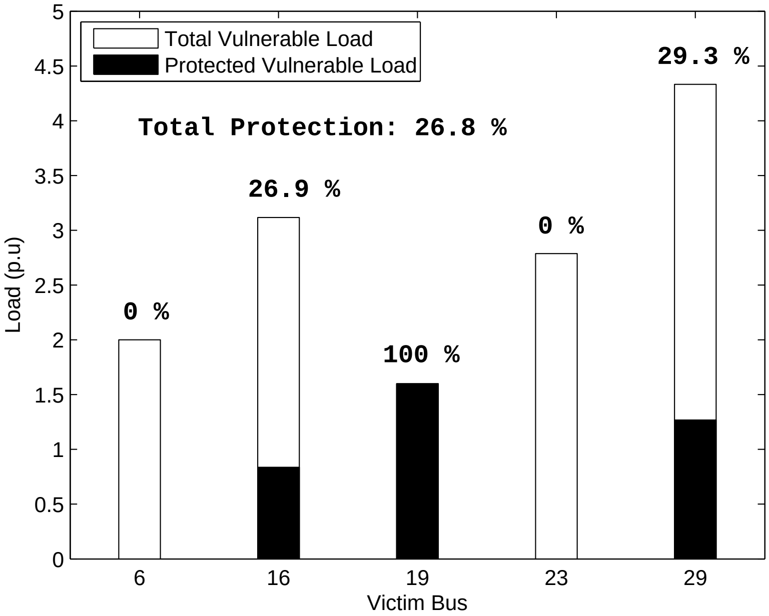


Figure 5.3: The optimal load protection scheme when the sensor location is unknown.

Chapter 6

Conclusions and Future Work

Conclusions

Destabilizing anomalies in the form of natural faults and malicious attacks were introduced. Also, model of power system dynamics in presence of such anomalies were investigated. Specifically, Dynamic Load Altering Attacks were introduced, characterized, and classified. Of particular interest was a closed-loop D-LAA against power system stability with feedback from power system frequency. Both single-point and coordinated multi-point attacks were investigated.

The problem of detecting presence of destabilizing anomalies in power systems from measurement data provided by Phasor Measurement Units and without knowledge of the power system dynamics was studied. Specifically, two scenarios were addressed: detecting anomalies solely based on load signal using a frequency-domain analysis and detecting anomalies based on both load and frequency signals using a time-domain analysis. Several detailed remarks are made in each case to gain analytical and practical insights. It was shown that depending on the type of anomaly and available data, both time-domain and frequency-domain detection analysis could be needed in order to ensure accurate anomaly detection.

The crucial problem of identifying location of destabilizing faults/attacks in power systems was also studied in this thesis. A novel optimization-based approach was proposed to identify the location(s) of destabilizing faults and attacks in power systems using synchronized measurements. The proposed method works in frequency-domain. It makes direct use of the information that is obtained during the detection phase. Compared to its time-domain counterpart, it needs much lower time-resolution in power system measurements. It does not require knowing the number of affected input location(s). It is also more computationally efficient. Importantly, it is well-suited to be deployed in wide area monitoring systems to do fault/attack location identification in a hierarchical fashion. It was also observed in this thesis that destabilizing anomalies can be modeled as a reparameterization of the power system's dynamical model. Therefore, an identification method that uses the unscented Kalman filter to jointly estimate both the system states and parameters of the anomaly was developed. A low-rank modification to the Kalman filter was also proposed that improves computational efficiency while maintaining the identification accuracy. The proposed method does not require prior knowledge on the number of buses that are compromised. It also does not require conducting a separate analysis at each bus. Instead, it naturally identifies anomaly on the entire system considered as a whole. Therefore, it has low computational complexity. Furthermore, it is capable of distinguishing destabilizing anomalies, i.e., load or generation control loops that are malicious and based on positive feedback, from the many load and generation control loops that exist in a power system that are benign and based on negative feedback. Numerical results show that this method successfully identifies complex destabilizing anomalies involving many buses.

Finally, a protection scheme was designed against destabilizing anomalies by formulating and solving a non-convex pole placement optimization problem. The non-convexity was tackled by using an iterative algorithm which solves a sequence of semi-definite optimization and convex feasibility optimization problems.

Various case studies were presented in this thesis to assess system vulnerabilities and the impacts of destabilizing anomalies on power system dynamics. Also, accuracy and efficiency of the proposed detection, identification, and protection approaches are verified in standard IEEE 9 and 39 bus test systems.

Future Work

This thesis can be extended in several following directions:

- Developing a more comprehensive model for capturing behavior of anomalies that also target power system measurements in addition to the inputs. Accordingly, developing new approaches for detection, identification, and protection.

- As we explained in this thesis, there are some challenges regarding to the practical implementation of frequency-domain detection and location identification approaches. Of interest for future work is optimal and adaptive selection of the sampling rate, window size, and detection threshold in windowed FFT for detection and location identification purposes.

- In this thesis, we investigated the location identification problem at transmission level of power system. Once the affected substation is identified, further examination of of the underneath distribution system is necessary to exactly pinpoint the location of anomalies. Therefore, similar frequency-domain techniques can be exploited to develop location identification approaches for distribution system as well.

Bibliography

[1] http://sys.elec.kitami-it.ac.jp/ueda/demo/WebPF/39-New-England.pdf.

[2] http://www.mainsfrequency.com/meter.htm.

[3] http://www.kios.ucy.ac.cy/testsystems/index.php/dynamic-ieee-test-systems.

[4] S. Amini, H. Mohsenian-Rad, and F. Pasqualetti. Dynamic load altering attacks in smart grid. In *IEEE PES Conference on Innovative Smat Grid Technologies (ISGT)*, Washington, D.C, February 2015.

[5] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad. Detecting dynamic load altering attacks: A data-driven time-frequency analysis. In *Proc. of IEEE Smart Grid Communications*, Miami, FL, November 2015.

[6] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad. Dynamic load altering attacks against power system stability: Attack models and protection schemes. *IEEE Trans. Smart Grid*, PP(99), October 2016.

[7] P. J. Antsaklis and Anthony N Michel. *Linear systems*. Springer Science & Business Media, 2006.

[8] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart: An open data set and tools for enabling research in sustainable homes. In *Proc. of the SustKDD*, Beijing, China, August 2012.

[9] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, Belmont, MA, 1997.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, 2004.

[11] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. SIAM, 1994.

[12] J. R. Buck, M. M. Daniel, and A. C. Singer. *Computer explorations in signals and systems using MATLAB*. Prentice Hall, 2002.

[13] CVX: Matlab Software for Disciplined Convex Programming. M. c. grant and s p. boyd. http://cvxr.com/cvx/, July 2014.

[14] J. W. Van der Woude. A graph-theoretic characterization for the rank of the transfer matrix of a structured system. *Signals and Systems Mathematics of Control*, 4(1):33–40, 1991.

[15] R. C. Dorf. *Modern control systems.* Addison-Wesley Longman Publishing Co., Inc., 1995.

[16] P. Du and N. Lu. Appliance Commitment for Household Load Scheduling. *IEEE Trans. on Smart Grid*, 2(2):411–419, June 2011.

[17] P. Duhamel and M. Vetterli. Fast fourier transforms: a tutorial review and a state of the art. *Signal processing*, 19(4), April 1990.

[18] C. Efthymiou and G. Kalogridis. Smart grid privacy via anonymization of smart metering data. In *Proc. of the IEEE SmartGridComm*, Gaithersburg, MD, October 2010.

[19] F. Zhang and Z. Geng and W. Yuan. The algorithm of interpolating windowed FFT for harmonic analysis of electric power system. *IEEE Trans. Power Delivery*, 16(2):160–164, April 2001.

[20] J. D. Fernandez and Andres E Fernandez. Scada systems: vulnerabilities and remediation. *Journal of Computing Sciences in Colleges*, 20(4):160–168, 2005.

[21] E. Ghahremani and I. Kamwa. Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units. *IEEE Transactions on Energy Conversion*, 26(4):1099–1108, 2011.

[22] E. Ghahremani and I. Kamwa. Local and wide-area pmu-based decentralized dynamic state estimation in multi-machine power systems. *IEEE Transactions on Power Systems*, 31(1):547–562, 2016.

[23] M. Ghamkhari and H. Mohsenian-Rad. Energy and performance management of green data centers: A profit maximization approach. *Smart Grid, IEEE Transactions on*, 4(2):1017–1025, 2013.

[24] J. D. Glover, M. S. Sarma, and T. J. Overbye. *Power System Analysis and Design.* Cengage Learning, 5 edition, 2009.

[25] U. Grenander. *Probability and Statistics: The Harald Cram.* Alqvist & Wiksell, 1959.

[26] Y. Guo, D. J. Hill, and Y. Wangi. Global transient stability and voltage regulation for power systems. *IEEE Trans. Power Systems*, 16(4):678–688, August 2002.

[27] E. Hammad, A. M. Khalil, A. Farraj, D. Kundur, and R. Iravani. A class of switching exploits based on inter-area oscillations. *IEEE Trans. Smart Grid*, PP(99), February 2017.

[28] R. Hassan, M. Abdallah, G. Radman, F. Marco, S. Hammer, J. Wigington, J. Givens, D. Hislop, J. Short, and S. Carroll. Under-frequency load shedding: towards a smarter smart house with a consumer level controller. In *Proc. of the IEEE SoutheastCon*, Nashville, TN, March 2011.

[29] P. Hines, S. Blumsack, C. E. Sanchez, and C. Barrows. The topological and electrical structure of power grids. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.

[30] L. Husheng and Z. Han. Manipulating the electricity power market via jamming the price signaling in smart grid. In *Proc. of the IEEE GLOBECOM Workshops*, Houston, TX, December 2011.

[31] M. Izbicki, S. Amini, C. R. Shelton, and H. Mohsenian-Rad. Identification of destabilizing attacks in power systems. In *Proc. of IEEE American Control Conference*, Seattle, WA, May 2017.

[32] R. J. Patton J. Chen and H. Zhang. Design of unknown input observers and robust fault detection filters. *International Journal of control*, 63(1):85–105, January 1996.

[33] P. Kadurek, C. Ioakimidis, and P. Ferrao. Electric vehicles and their impact to the electric grid in isolated systems. In *Proc. of International Conference on Power Engineering, Energy and Electrical Drives*, Lisbon, Portugal, March 2009.

[34] S. Kiliccote, S. Lanzisera, A. Liao, O. Schetrit, and M. A. Piette. Fast dr: Controlling small loads over the internet. *Forthcoming Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, 2014.

[35] P. Kundur. *Power System Stability and Control.* McGraw-Hill, 1994.

[36] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu. Securing smart grid: cyber attacks, countermeasures, and challenges. *IEEE Communications Magazine*, 50(8):38–45, 2012.

[37] J. Liang, S. K. Ng, G. Kendall, and J. W. Cheng. Load signature study—part i: Basic concept, structure, and methodology. *IEEE Trans. on Power Delivery*, 25(2):551–560, 2010.

[38] C. Liu, Z. Chen, C. L. Bak, and Z. Liu. Adaptive voltage stability protection based on load identification using phasor measurement units. In *Proc. of IEEE APAP*, Beijing, China, October 2011.

[39] X. Liu and Z. Li. Local load redistribution attacks in power systems with incomplete network information. *IEEE Trans. on Smart Grid*, 5(4):1665–1676, July 2014.

[40] R. Ma, H. Chen, Y. Huang, and W. Meng. Smart grid communication: Its challenges and opportunities. *Smart Grid, IEEE Transactions on*, 4(1):36–46, 2013.

[41] M. Mahmoud, J. Misic, X. Shen, et al. Investigating public-key certificate revocation in smart grid. *IEEE Internet of Things Journal*, March 2015.

[42] G. Marks. Opportunities for demand response in california agricultural irrigation: A scoping study. 2014.

[43] A. K. Marnerides, P. Smith, A. Schaeffer-Filho, and A. Mauthe. Power consumption profiling using energy time-frequency distributions in smart grids. *IEEE Communications Letters*, 19(1):46–49, 2015.

[44] A. Medina, J. Segundo, P. Ribeiro, W. Xu, K. Lian, G. Chang, V. Dinavahi, and N. Watson. Harmonic analysis in frequency and time domain. *IEEE Trans. on Power Delivery*, 28(3), July 2013.

[45] L. Meegahapola and D. Flynn. Impact on transient and frequency stability for a power system at very high wind penetration. In *Proc. of IEEE PES General Meeting*, Minneapolis, MN, July 2010.

[46] S. Mishra, X. Li, T. Pan, A. Kuhnle, M. T. Thai, and J. Seo. Price modification attack and protection scheme in smart grid. *IEEE Trans. on Smart Grid*, 2016.

[47] Y. Mo and B. Sinopoli. False data injection attacks in control systems. In *Proc. of Workshop on Secure Control Sys.*, Stockholm, Sweden, 2010.

[48] H. Mohsenian-Rad and A. Leon-Garcia. Optimal Residential Load Control with Price Prediction in Real-Time Electricity Pricing Environments. *IEEE Trans. on Smart Grid*, 1:120–133, September 2010.

[49] H. Mohsenian-Rad and A. Leon-Garcia. Distributed internet-based load altering attacks against smart power grids. *IEEE Trans. on Smart Grid*, 2(4):667–674, December 2011.

[50] H. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia. Autonomous Demand Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid. *IEEE Trans. on Smart Grid*, 1(3):320–331, December 2010.

[51] A. Molina-Garcia, F. Bouffard, and D. S. Kirschen. Decentralized demand-side contribution to primary frequency control. *IEEE Trans. on Power Systems*, 26(1):411–419, February 2011.

[52] K. Moslehi and R. Kumar. A reliability perspective of the smart grid. *Smart Grid, IEEE Transactions on*, 1(1):57–64, 2010.

[53] F. Pasqualetti, A. Bicchi, and F. Bullo. A graph-theoretical characterization of power network vulnerabilities. In *Proc. of IEEE American Control Conference*, San Francisco, CA, June 2011.

[54] F. Pasqualetti, F Dorfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Trans. Automatic Control*, 58(11):2715–2729, June 2013.

[55] A. Pertsinidis, I. Grossmann, and G. McRae. Parametric optimization of MILP programs and a framework for the parametric optimization of MINLPs. *Computers & Chemical Eng.*, 22:205–212, 1998.

[56] A. G. Phadke and J. S. Thorp. *Synchronized Phasor Measurements and Their Applications.* Springer International Publishing, 2017.

[57] X. Qing, H. Karimi, Y. Niu, and X. Wang. Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems. *International Journal of Electrical Power & Energy Systems*, 65:26–33, 2015.

[58] C. Rehtanz and J. Bertsch. Wide area measurement and protection system for emergency voltage stability control. In *Proc. of IEEE Power Engineering Society Winter Meeting*, New York, NY, January 2002.

[59] S., W. Gao, and A.P. Sakis Meliopoulos. An alternative method for power system dynamic state estimation based on unscented transform. *IEEE Transactions on Power Systems*, 27(2):942–950, 2012.

[60] S. V. Buldyrev and R. Parshani and G. Paul and H. E. Stanley and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, 2010.

[61] S. Shao, M. Pipattanasomporn, and S. Rahman. Demand response as a load shaping tool in an intelligent grid with electric vehicles. *IEEE Trans. on Smart Grid*, 2(4):624–631, December 2011.

[62] S. Shao, T. Zhang, M. Pipattanasomporn, and S. Rahman. Impact of tou rates on distribution load shapes in a smart grid with phev penetration. In *Proc. of IEEE PES Transmission and Distribution Conference and Exposition*, 2010.

[63] D. Soudbakhsh, A. Chakrabortty, F. Alvarez, and A. Annaswamy. A delay-aware cyber-physical architecture for wide-area control of power systems. *Control Engineering Practice*, 60(1), March 2017.

[64] P. Srikantha and D. Kundur. A DER attack-mitigation differential game for smart grid security analysis. *IEEE Trans. Smart Grid*, 7(3):1476 – 1485, August 2015.

[65] G. Strbac. Demand side management: Benefits and challenges. *Energy Policy*, 36(12):4419–4426, December 2008.

[66] A. Teixeira, H. Sandberg, and K. H. Johansson. Networked control systems under cyber attacks with applications to power networks. In *Proc. of IEEE American Control Conference*, Baltimore, MD, July 2010.

[67] C. W. Ten, L. Chen-Ching, and M. Govindarasu. Vulnerability assessment of cybersecurity for scada systems. *Power Systems, IEEE Transactions on*, 23(4):1836–1846, 2008.

[68] Y. Tomita, C. Fukui, H. Kudo, J. Koda, and K. Yabe. A cooperative protection system with an agent model. *IEEE Transactions on Power Delivery*, 13(4):1060–1066, 1998.

[69] K. Vanthournout, R. D'hulst, D. Geysen, and G. Jacobs. A smart domestic hot water buffer. *Smart Grid, IEEE Transactions on*, 3(4):2121–2127, 2012.

[70] J. C. Vieira, W. Freitas, W. Xu, and A. Morelato. Performance of frequency relays for distributed generation protection. *Power Delivery, IEEE Transactions on*, 21(3):1120–1127, 2006.

[71] E. A. Wan and R. V. Merwe. The unscented Kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. Ieee, 2000.

[72] S. Wang, E Wang, and P Dorato. Observing the states of systems with unmeasurable disturbances. *IEEE Trans. Automatic Control*, 20(5):716–717, October 1975.

[73] Z. Wang, A. Scaglione, and R. J. Thomas. Generating statistically correct random topologies for testing smart grid communication and control networks. *IEEE transactions on Smart Grid*, 1(1):28–39, 2010.

[74] D. Wei, Y. Lu, M. Jafari, P. Skare, and K. Rohde. An integrated security system of protecting smart grid against cyber attacks. In *Innovative Smart Grid Technologies (ISGT), 2010*, pages 1–7. IEEE, 2010.

[75] J. Yan, C. C. Liu, and M. Govindarasu. Cyber intrusion of wind farm scada system and its impact analysis. In *Proc. of IEEE/PES Power Systems Conference and Exposition (PSCE)*, Phoenix, AZ, March 2011.

[76] Y. Yan, Y. Qian, H. Sharif, and D. Tipper. A survey on cyber security for smart grid communications. *Communications Surveys & Tutorials, IEEE*, 14(4):998–1010, 2012.

[77] L. Yang, S. Hu, and T. Ho. Vulnerability assessment and defense technology for smart home cybersecurity considering pricing cyberattacks. In *Proc. of the IEEE/ACM International Conference on Computer-Aided Design*, San Jose, CA, November 2014.

[78] L. Yao and L. Hau-Ren. A Two-Way Direct Control of Central Air-Conditioning Load Via the Internet. *IEEE Trans. on Power Delivery*, 24(1):240–248, January 2009.

[79] Y. Zhang, Z. Wang, J. Zhang, and J. Ma. Fault localization in electrical power systems: A pattern recognition approach. *Int. Journal of Electrical Power & Energy Systems*, 33:791–798, March 2011.

[80] C. Zhao, U. Topcu, N. Li, and S. Low. Design and stability of load-side primary frequency control in power systems. *Automatic Control, IEEE Transactions on*, 59(5):1177–1189, 2014.

[81] C. Zhao, U. Topcu, and S. H. Low. Optimal load control via frequency measurement and neighborhood area communication. *IEEE Trans. on Power Systems*, 28(4):3576–3587, November 2013.

Appendix A

Appendix

A.1 Linear Quadratic Lyapunov Theory

Consider the linear time-invariant system $\dot{x} = Ax$. Using Lyapunov function $V(x) = x^T Xx$, one can show that:

1. The system is stable if there exists a real, symmetric, and positive definite matrix X such that $C = A^T X + XA$ and C is negative definite [7, Theorem 7.3]; and
2. The system is unstable if $C = A^T X + XA$ is negative definite, and X is real, symmetric, and either negative definite or indefinite [7, Theorem 7.3].

A.2 Proof of Theorem 9

To prove the first property, recall from Section 4.1.2 that constraint (4.7e) determines the number of non-zero entries in vector I. Accordingly, constraint (4.7f) is equivalent to

$$|U_i^c(j\omega^*)| \leq 0 \qquad\qquad i \notin \mathcal{K}, \tag{A.1a}$$

$$|U_i^c(j\omega^*)| \leq U_i^{\max} \qquad\qquad i \in \mathcal{K}. \tag{A.1b}$$

Constraint (A.1b) is *less restrictive* than constraint (A.1a). Therefore, as we increase $|\mathcal{K}|$, we expand the feasible set, i.e., we make the optimization problem more relaxed. As a result, the optimal objective value in problem (4.7) either decreases or remains the same. Therefore, we can conclude that function $F(|\mathcal{K}|)$ is non-increasing. That is, we have:

$$F(|\mathcal{K}| + 1) \leq F(|\mathcal{K}|). \tag{A.2}$$

From (4.9) and (A.2), and after reordering the terms, we have:

$$S(|\mathcal{K}) = F(|\mathcal{K}|) - F(|\mathcal{K}| + 1) \geq 0. \tag{A.3}$$

Next, we prove the second property. According to the *mixed integer problem sensitivity analysis* in [55], the optimal objective value of problem (4.7) is a convex function of parameter $|\mathcal{K}|$. In other words, $F(|\mathcal{K}|)$ is a convex function. From the definition of convexity, for any $0 \leq \theta \leq 1$, we have:

$$F(\theta x + (1 - \theta)y) \leq \theta F(x) + (1 - \theta)F(y), \; \forall x, y. \tag{A.4}$$

Suppose $\theta = 0.5, x = |\mathcal{K}|$, and $y = |\mathcal{K}| + 2$. We can derive:

$$F(|K| + 1) = F(0.5|K| + 0.5(|K| + 2)) \leq 0.5F(|K|) + 0.5F(|K| + 2), \tag{A.5}$$

where the inequality is due to (A.4). Once we multiply both sides by two, and after reordering the terms, we have:

$$F(|\mathcal{K}|+1) - F(|\mathcal{K}|+2) \leq F(|\mathcal{K}|) - F(|\mathcal{K}|+1). \tag{A.6}$$

From (4.9) and (A.6), we can conclude the second property.