Vulnerability Analysis of False Data Injection Attacks on Supervisory Control and Data Acquisition and Phasor Measurement Units

by

Jiazi Zhang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved September 2017 by the
Graduate Supervisory Committee:

Lalitha Sankar, Chair
Oliver Kosut
Kory Hedman
Vijay Vittal

ARIZONA STATE UNIVERSITY

December 2017

ABSTRACT

The electric power system is monitored via an extensive network of sensors in tandem with data processing algorithms, *i.e.*, an intelligent cyber layer, that enables continual observation and control of the physical system to ensure reliable operations. This data collection and processing system is vulnerable to cyber-attacks that impact the system operation status and lead to serious physical consequences, including systematic problems and failures.

This dissertation studies the physical consequences of unobservable false data injection (FDI) attacks wherein the attacker maliciously changes supervisory control and data acquisition (SCADA) or phasor measurement unit (PMU) measurements, on the electric power system. In this context, the dissertation is divided into three parts, in which the first two parts focus on FDI attacks on SCADA and the last part focuses on FDI attacks on PMUs.

The first part studies the physical consequences of FDI attacks on SCADA measurements designed with limited system information. The attacker is assumed to have perfect knowledge inside a sub-network of the entire system. Two classes of attacks with different assumptions on the attacker's knowledge outside of the sub-network are introduced. In particular, for the second class of attacks, the attacker is assumed to have no information outside of the attack sub-network, but can perform multiple linear regression to learn the relationship between the external network and the attack sub-network with historical data. To determine the worst possible consequences of both classes of attacks, a bi-level optimization problem wherein the first level models the attacker's goal and the second level models the system response is introduced.

The second part of the dissertation concentrates on analyzing the vulnerability of system to FDI attacks from the perspective of the system. To this end, an off-line vulnerability analysis framework is proposed to identify the subsets of the test system

that are more prone to FDI attacks.

The third part studies the vulnerability of PMUs to FDI attacks. Two classes of more sophisticated FDI attacks that capture the temporal correlation of PMU data are introduced. Such attacks are designed with a convex optimization problem and can always bypass both the bad data detector and the low-rank decomposition (LD) detector.

DEDICATION

This dissertation is dedicated to my father Guoqing Zhang and my mother Xihua Zhao, for their endless love and support through all these years.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation and thanks to my advisors, Dr. Lalitha Sankar and Dr. Oliver Kosut, for their guidance, encouragement, and invaluable support throughout my research work and my life as a graduate student.

I am also grateful to the members of my committee, Dr. Kory Hedman, and Dr. Vijay Vittal for their valuable suggestions and comments.

In addition, I would like to express my gratitude to the National Science Foundation (NSF), Department of Homeland Security (DHS), and Power Systems Engineering Research Center (PSERC) for the financial support provided.

Last but not least, I would like to thank all my friends in the areas of electric power and energy systems, industrial engineering, environmental engineering, and economics, for their compelling intellectual discussion as well as their friendship and support.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

LIST OF SYMBOLS

A_{GB}	Dependency matrix between the incremental injections and generations
A_{KN}	The line-to-bus connectivity matrix
C	Attack matrix
C'	Voltage angle attack matrix
$C_g(\cdot)$	The quadratic cost function for generator g.
$C_{S,i}$	The cost of the shedding load at bus i
D	Attacked measurement matrix
E	Noise matrix
G	The generator-to-bus connectivity matrix.
H	Jacobian matrix
K	The power transfer distribution factor (PTDF) matrix in $\mathcal{G}$.
L	Matrix of line outage distribution factor
M	A large enough constant
N_*	Nuclear norm constraint limit
N_0	The l_0-norm constraint integer
N_1	The l_1-norm constraint integer
P	Active power flow vector
P_S	Active power load shedding vector
P_D	Active power load vector
$P_{G,\max}$	The vector of maximum generation output limit.
$P_{G,\min}$	The vector of minimum generation output limit.
P_G	Active power generation vector
$P_{\max}$	The vector of transmission line rating.
R	Measurement error covariance matrix
V	Voltage magnitude state vector
Z	Measurement matrix

Γ	The dependency matrix between power flow and voltage angle state.
Ω_i	The set of lines that connected to bus i
Ω_{int}	The set of key transmission lines connecting two regions
$\alpha^{\pm}$	The dual variable vectors for the maximum and minimum generation limit constraints, respectively
$\delta_{\alpha}^{\pm}$	the vectors of binary variables associated with generation limits
$\delta_{\mu}^{\pm}$	the vectors of binary variables associated with line thermal limits
$\hat{x}$	Estimated state vector
λ	The dual variable vector for node balance constraints
$\mathcal{B}$	The set of boundary buses in $\mathcal{L}$
$\mathcal{E}$	The remaining network where the attacker has limited information or no information, $\mathcal{E} = \mathcal{G} \setminus \mathcal{L}$
$\mathcal{E}_{b,m}$	The set of buses in $\mathcal{E}$ on which marginal generators connecting to
$\mathcal{E}_{g,m}$	The set of marginal generators in $\mathcal{E}$
$\mathcal{G}$	The entire electric power network
$\mathcal{I}$	The set of internal buses in $\mathcal{L}$, $\mathcal{I} = \mathcal{L}_b \setminus \mathcal{B}$
$\mathcal{I}_a$	Set of attacked states
$\mathcal{L}$	The attack sub-network where the attacker has perfect system information
$\mathcal{L}_b, \mathcal{L}_{br}, \mathcal{L}_g$	the set of buses, lines, and generators of $\mathcal{L}$
$\mathcal{S}$	Attack subgraph
$\mu^{\pm}$	The dual variable vectors for the positive and negative directions of thermal limit constraints, respectively
σ_i^2	covariance of noise of measurement i
σ_{r_i}	The standard deviation of the i^{th} residual error r_i
τ	The load shift factor
$\text{diag}(\cdot)$	Diagonal matrix of a vector
θ	Voltage angle state vector
Θ	Voltage angle state matrix

ε	Converge threshold in Benders' decomposition
ζ	The weight of the norm of attack vector c
e	Noise vector
$h(x, \mathcal{G})$	A vector of nonlinear functions that describes the relationship between the system states and measurements for a topology $\mathcal{G}$
n_{br}	Number of branch
n_b	Number of bus
n_g	Number of generator
n_{load}	Number of load
n_z	Number of measurement
r	Residual vector
$supp\,(\cdot)$	The set of non-zero columns in a matrix
u	The slack vector used to linearize l_0-norm of c
x	State vector
z	Measurement vector
CA	Contingency analysis
CPF	Attacker-computed physical power flow
EMS	Energy management system
EPS	Electric power system
FDI	False data injection
KKT	Karush-Kuhn-Tucker
LMP	Locational marginal price
LNR	Largest normalized residual
MILP	Mixed-integer linear programming
MP	Master problem in Benders' Decomposition
OPF	Optimal power flow
PPF	Physical power flow
SCADA	Supervisory control and data acquisition

SE	State estimation
SP	Slave problem in Benders' Decomposition
WLS	Weighted least-squares
ZD	Lower bound in Benders' decomposition
ZU	Upper bound in Benders' decomposition

Chapter 1

INTRODUCTION

1.1 Overview

The electric power system (EPS) is a hierarchical network involving the transportation of power from the sources of power generation via an intermediate densely connected transmission network to a large distribution network of end-users. To ensure reliable operation of the entire power system, the system operators at each level should be aware of the real-time operation states. Therefore, secure and intelligent cyber data processing systems that monitor and control each level of the physical power system are crucial for reliable real-time operations. In traditional power systems, such a cyber layer includes (a) supervisory control and data acquisition (SCADA) system, and (b) energy management systems (EMSs) for data processing. In the past decade, phasor measurement units (PMUs) have been deployed in the electric power system to directly measure voltages and phase angles for key generation and transmission buses. Due to its high sampling rate and accuracy, PMUs have the potential to play a significant role in real-time power system state estimation (SE) [1], dynamic security assessment [2, 3], system protection [4], and system awareness [5]. In recent years, PMU measurements have gradually come to supplement the existing SCADA infrastructure. Moreover, in some specific regions, full observability of the power system can be achieved by PMU measurements alone. Figure 1.1 illustrates an example of the hierarchical electric power system. The figure demonstrates how at various levels from generation through transmission, substation, distribution, and even end-users, the network appropriate at each level is monitored using the collected SCADA and/or

PMU data via an EMS to enable control and actuation of the underlaying physical system.

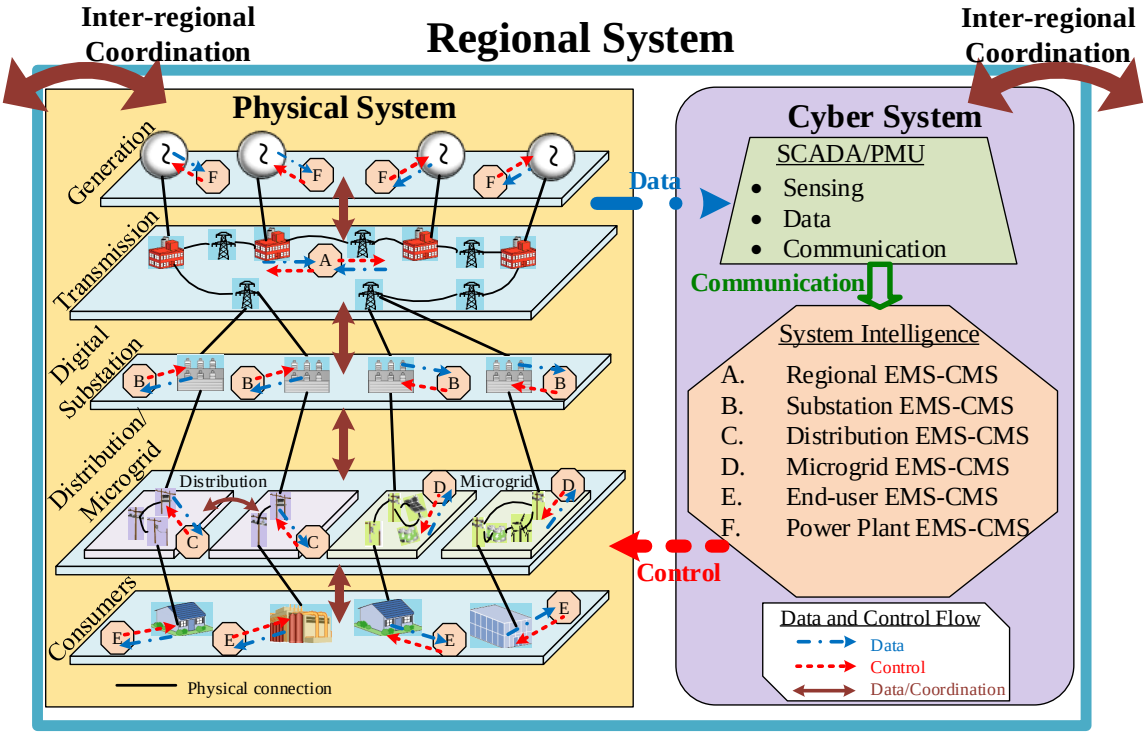


Figure 1.1: Hierarchical Cyber-Physical Power System.

However, at all levels of the hierarchy, SCADA, PMUs, and EMSs are vulnerable to cyber attacks. In fact, such cyber attacks can undermine the observability of physical systems, and hence, lead to mis-operation, violation, and inappropriate and/or untimely contingency response, which can potentially result in severe economic and social consequences. This has been verified by several relevant cyber incidents in recent years:

- In 2007, Idaho National Laboratory ran the Aurora Generator Test and demonstrated a cyber attack in which a diesel generator's circuit breaker was rapidly opened and closed by the attack. This attack led to the generator becoming out of phase from the rest of the power system and exploding [6].

- In 2009, the Stuxnet virus ravaged roughly 20% of Iran's nuclear centrifuges by causing them to spin out of control [7].

- In 2010, the Stuxnet malware attacked SCADA systems that use Siemens WinCC SCADA software, which in turn infected 14 power plants in Germany [8].

- In 2012, the Industrial Control Systems Cyber Emergency Response Team revealed that the number of reported cyber attacks is growing and the companies with access to the country's power grid have become the cyber attack targets [9]. The U.S. Department of Energy reported that from 2011 to 2014, 362 reports were received from electric utilities of physical or cyber attacks that interrupted power services [10]. In 2013, CNN reported that hacker hits on US power and nuclear targets spiked in 2012 [9]. A Department of Homeland Security branch recorded 161 cyber attacks on the energy sector in 2013, compared to just 31 in 2011 [10], which comprises 60% of all cyber attacks on cyber-physical systems.

- In 2015, there was a cyber-attack on the computers and SCADA systems of four control centers in Ukraine, disconnecting the circuit breakers at nearly 60 substations. This attack resulted in a regional power outage for nearly 6 hours across various areas, affecting approximately 225 thousand customers [11].

The list above illustrates that the cyber layer of the power system is vulnerable to cyber attacks. Therefore, it is crucial to fully understand the consequences of realistic and credible cyber-attacks as a first step to thwart such attacks. This dissertation focuses on a specific class of cyber attacks, false data injection (FDI) attacks, on the transmission system, in which the attacker replaces a subset of measurements with counterfeits before they are processed by the cyber layer.

1.2 Literature Review

There has been much recent interest in understanding the cyber-security challenges facing the electric power system. Since it is not possible to review all these challenges, we will focus on the class of *unobservable attacks*. Unobservable attacks are a class of cyber attacks in which the attacker focuses on SE and within SE, the data change made by attacker appears exactly as if it originated from a normal state; thus, it cannot be detected by existing bad data detectors. We briefly review the existing literature on unobservable attacks.

In [12], the authors are the first to introduce a class of FDI attacks on DC SE. The authors show that an attacker with sufficient system knowledge can inject malicious data in SCADA measurements without being detected by existing bad data detection techniques. In [13], Sandberg *et al.* introduce two security indices which quantify the least effort to launch unobservable FDI attacks on DC SE. In [14], Teixeira *et al.* analyze how to completely protect SE by placing encrypted devices on a set of SCADA measurements into the power system. In [15], the authors further propose the minimum cost protection scheme to thwart unobservable FDI attacks. In [16], Kosut *et al.* discuss the trade-off between the attacker's efforts to maximize the attack strength and minimize the detection rate.

In contrast to the above mentioned references, in [17], Hug and Giampapa focus on FDI attacks on AC SE and introduce a class of unobservable attacks that are limited to a sub-graph of the networks. They demonstrate that although AC SE is vulnerable to unobservable FDI attacks, doing so requires the knowledge of both system topology and states to launch such attacks. In [18], Liang *et al.* introduce unobservable FDI attacks for a nonlinear measurement model of AC SE and demonstrate that such FDI attacks can lead to a re-dispatch of generators when none was actually needed; this

re-dispatch can in turn cause a line to overload in the physical system.

The impact of FDI attacks on electric power markets is studied in [19]. The authors demonstrate that FDI attacks can be utilized to manipulate locational marginal prices (LMPs) of ex-post real-time market, and thus, allow the attacker to profit. In general, there are many easier ways to make a profit from the market instead of attacking the electric power system. Therefore, the focus should be on whether the cyber attacks can have physical consequences on the power system.

Yet another class of FDI attacks is one that alters topology data in an unobservable manner. In [20], Kim and Tong formally introduce an undetectable topology attack as a specific class of FDI attacks on power systems and evaluate the attack's impact on the electric market. In [21], Rahman *et al.* study the impact of the undetectable topology attack introduced in [20] on DC OPF when DC SE is used; the attack is optimized to increase the total operation cost of the system. In [22], Ashok and Govindarasu analyze a different class of topology attacks wherein the attacker compromises the critical measurements of the system to result in incorrect contingency analysis results. In [23], the author and her collaborator study unobservable state-preserving line-maintaining attacks (*i.e.*, only topology data is changed) for which an algorithm using breadth-first search (BFS) is developed to find the smallest sub-network required to launch such an attack.

Recently, there are growing interests on understanding the consequences of FDI attacks. In [24] and [25], the authors propose a max-min attacker-defender model to study the most damaging FDI attacks with both a short term goal and a long term goal. They formulate a class of load redistribution attacks as a bi-level optimization problem. They find the attack which maximizes operation costs of the attack-induces re-dispatch. More recently, in [26], the authors introduce a bi-level optimization for the worst unobservable attacks on AC SE. The objective of the attack is to maximize

power flow on a specific line, and thereby, cause overflow violations in the physical system. This attack optimization problem can be solved using an equivalent mixed-integer linear programming (MILP) problem. In [27] and [28], four computationally efficient algorithms are proposed to scale the MILP in [26] to large systems so as to provide lower and upper bounds on the attack consequences. Besides pure cyber attacks on states, my prior work [29] demonstrates that a two-step attack strategy can be utilized to design unobservable cyber attacks to coordinate and mask physical attacks. Such a coordinated cyber-physical attacks can result in line overflow in the physical power system while it is unobservable in the cyber layer.

Besides FDI attacks designed with complete system information as introduced in the above literature, there are growing attempts to model FDI attacks with incomplete and localized system information. In [30], Rahman and Mohsenian-Rad introduce a class of FDI attacks modeled with incomplete system topology information which results from lack of real-time knowledge of circuit breaker status and transformer tap. The authors demonstrate that power systems are still vulnerable to FDI attacks designed with incomplete system information, however, the probability of the attacks being detected increases as the uncertainties of system information grow. In [31] and [32], Liu and Li propose a local load redistribution attack model with incomplete network information on DC SE and AC SE and demonstrate that an attacker can design unobservable DC and AC FDI attacks, respectively, based on localized information.

Thus far, the FDI attacks introduced in the above literature target on SCADA measurements. Such attacks can also be converted and injected in PMU measurements. In [33, 34], the authors study the cyber-security of PMU-based SE and classify potential cyber-attacks on PMUs as attacks againt communication links, denial of service attacks, and data spoofing attacks including GPS spoofing attacks and FDI

attacks.

To thwart unobservable FDI attacks, several protection mechanisms and attack detection approaches have been introduced. In [35], Kim and Tong introduce an approach to ensure system observability by placing secure PMUs so as to protect the system from FDI attacks. However, since PMU measurements may also be vulnerable to FDI attacks, this method cannot eliminate FDI attacks when PMUs are compromised by attackers. In [36], the authors propose a decentralized detection scheme for FDI attacks based on the Markov graph of bus phase angles. However, this method might not work well when the system experiences a disturbance. In [37], Lee and Kundur introduce a detector based on Expectation-Maximization to detect FDI attacks in PMU measurements. This method needs to be solved iteratively and the convergence rates are very slow for real-time detection (*e.g.,* 10^5 iterations) for even a small test system.

Recently in [38, 39, 40], low-rank decomposition (LD) has been proposed to detect FDI attacks on the electric power system using a block of consecutive measurement data. In [38], the authors propose an LD approach (introduced in [41] for arbitrary sparse datasets), for temporal SCADA data; specifically, they demonstrate that attacks designed without knowledge of the temporal correlations of the SCADA measurements can be detected by solving an LD problem. Furthermore, their model assumes that while the FDI attack matrix is sparse in each time instant, the attacker attacks a different set of measurements. While such a model is quite general, for attacks designed with a specific effect (financial or physical damage), sustaining attacks over time on the same meters can have more impact. Focusing on such sustained attacks, for PMU data, the authors of [39, 40] show that an LD-based detector can identify column sparse FDI attack matrix where the column sparsity is a result of the assumption that the attacker attacks the same set of PMU measurements over time.

1.3 Objective of the Dissertation

This dissertation analyzes the vulnerability and physical consequences of FDI attacks in which the attacker can change SCADA or PMU measurements. In particular, we focus on FDI attacks that can result in line overflow without being detected via measurements.

It has been shown in [26] that attacks targeting SCADA measurements can be designed via a bi-level optimization problem wherein the first level problem models the attacker's goal and the second level problem models the system response. However, the attack optimization problem in [26] requires the attacker to know system-wide information including topology, generation cost and capacity, as well as load data. In practice, obtaining all the required information can be difficult for the attacker. Therefore, one of the goals of this dissertation is to study the FDI attacks designed with limited attacker information. Although [30, 31, 32] have already studied several classes of FDI attacks inside a sub-network of the network graph, the analysis of attacks is limited to its feasibility and observability. The physical consequences of the worst-case limited information FDI attacks are yet to be studied. Therefore, we seek to understand the physical consequences from the perspectives of both the attacker and the system, so as to identify the subsets of the system that are more vulnerable to FDI attacks. This in turn can be used to develop attack detection resiliency mechanisms.

In addition, due to the lack of knowledge of the measurement temporal correlations, the FDI attacks introduced in [26] can be detected by the LD detector introduced in [38, 39, 40] when injecting to PMU measurements. Thus, the second goal of this dissertation is to study whether FDI attacks can be designed to capture the temporal correlation of PMU data and thereby bypass the LD detector.

1.4 Dissertation Organization

The remainder of this dissertation is organized as follows.

Chapter 2 presents the mathematical formulation for the various computational units of EMSs, including state estimation, bad data detection, and optimal power flow. The unobservable FDI attack model is also reviewed in this chapter.

In Chapter 3, prior work on the bi-level optimization problem for line overflow attacks with perfect information is reviewed.

In Chapter 4, two classes of unobservable FDI attacks designed with limited system information are studied. In the first class of attacks, the attacker is assumed to have perfect knowledge in an attack sub-network and limited estimated information outside of the sub-network. A modified bi-level attack optimization problem considering limited system information is presented. Discussion on the impact of incomplete and inaccurate knowledge outside the attack sub-network on attack consequences is given. The performance of the attack strategy and the long-term consequences of such attacks are illustrated via simulation. In the second class of attacks, the attacker is assumed to have no information outside of the attack sub-network. To overcome the limited information, a multiple linear regression model is developed to learn the relationship between the external network and the attack sub-network from historical data. The worst possible consequences of such FDI attacks are evaluated by solving a bi-level optimization problem. Justifications for the proposed attack strategy are provided, followed by numerical results of the attack consequences. A sensitivity analysis of the second class of attacks on multiple scenarios of imperfect historical datasets is provided. Discussions on the impact of the approximations made in the attack models are provided.

In Chapter 5, an off-line vulnerability analysis framework from the perspective

of the system control center is proposed. The performance of this framework is demonstrated with the IEEE 24-bus RTS system, IEEE 118-bus system, and IEEE 2383-bus (Polish) system.

Chapter 6 presents the vulnerability analysis to FDI attacks on PMU. In this chapter, the PMU data is modeled as a low-rank matrix and a low-rank decomposition detector is used to identify FDI attacks. Two classes of FDI attacks on PMU that can bypass the low-rank decomposition detector are introduced. The performance of the attack strategy is illustrated via simulation.

In Chapter 7, conclusions and future works are provided.

Chapter 2

SYSTEM AND ATTACK MODEL

In this chapter, we introduce the mathematical formulation for the state estimation (SE), and optimal power flow (OPF) units in EMS. The unobservable FDI attack model is also reviewed. Throughout, we assume there are n_b buses, n_{br} branches, n_g generators, n_{load} load buses, and n_z measurements in the system. The temporal nature processing of real-time power operation is illustrated in Figure 2.1.

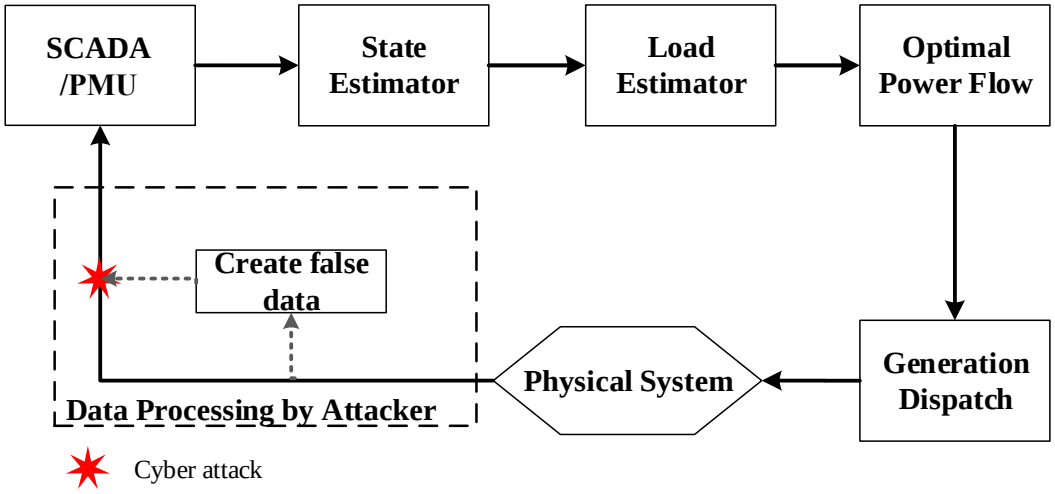


Figure 2.1: Temporal Nature of Real-Time Power System Operation.

2.1 State Estimation

Consider an $n_z \times 1$ vector z of nonlinear measurements given as

$$z = h(x, \mathcal{G}) + e \tag{2.1}$$

where $x = [\theta, V]^T$ is the system state vector including voltage angles θ and voltage magnitudes V, and e is an $n_z \times 1$ noise vector which is independent of x and is modeled as Gaussian distributed with 0 mean and σ_i^2 covariance, such that the measurement error covariance matrix is given by $R = diag(\{\sigma_i^2\}_{i=1}^M)$. The function $h(x, \mathcal{G})$ is a vector of nonlinear functions that describes the relationship between the system states and measurements for a topology $\mathcal{G}$. Both the line status data s and the measurements z are collected by the SCADA system and/or PMUs. The commonly obtained measurements in the grid are the active and reactive line power flows and bus injections.

We use the weighted least-squares (WLS) AC SE to calculate the θ and V [42]. The objective of the estimation process is to minimize the sum of the squares of the weighted deviations of the estimated measurements from z. The states are solved as a least square problem with the following objective function

$$\min J(x) = (h(x) - z)^T R^{-1} (h(x) - z), \tag{2.2}$$

the solution to which satisfies

$$g(\hat{x}) = \frac{\partial J(\hat{x})}{\partial x} = H^T(\hat{x}) \cdot R^{-1} \cdot (h(\hat{x}) - z) = 0 \tag{2.3}$$

where $H = \frac{\partial h(x)}{\partial x} \mid_{x=\hat{x}}$ is the system Jacobian matrix, and $\hat{x}$ is the $2n_b \times 1$ estimated state vector. The WLS solution for this nonlinear optimization problem can be solved iteratively.

Specifically, the linearized measurement vector can be written as:

$$z = Hx + e \tag{2.4}$$

where x now is a $n_b \times 1$ vector with only voltage angle states.

We use the weighted least-squares (WLS) method to solve this problem [42] as

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} z \tag{2.5}$$

where $\hat{x}$ is the estimated system state vector.

Measurements collected by SCADA may contain errors that can undermine the accuracy of estimated states. Therefore, a bad data detector is equipped with SE to detect faulty measurements, and hence, protect SE from large errors. The measurement residual vector is used to detect bad data, as

$$r = z - h(\hat{x}, \mathcal{G}) \tag{2.6}$$

where r is the $n_z \times 1$ residual vector.

In this dissertation, the χ^2-detector is utilized to detect bad data. The threshold is determined by the χ^2-test. To bypass the bad data detection, the residual vector should satisfy the following relationship

$$r^T R^{-1} r \leqslant \chi^2_{(m-n),p} \tag{2.7}$$

where $\chi^2_{(m-n),p}$ is the value from the χ^2-distribution table corresponding to a detection confidence with probability p (*e.g.* 95%) and $m - n$ degrees of freedom.

If the threshold in (2.7) is violated, the largest normalized residual (LNR) method is further used for bad data identification as follows:

$$\underset{i}{\text{Max}} \frac{|r_i|}{\sigma_{r_i}} \leqslant \tau_r \tag{2.8}$$

where σ_{r_i} is the standard deviation of the i^{th} residual error r_i. If the LNR test is not passed, the measurement with maximum residual is identified as a bad measurement. The bad measurement is then removed from the measurement vector and SE is repeated until no bad data is detected.

The SE solution that passes the bad data detection is used to compute the power flow of the system, which hence yields the estimated loads of the system. The estimated loads then pass to the OPF module for an optimal power dispatch solution. If the state vector is maliciously altered by an attacker, it can result in a wrong dispatch solution, which can lead the system to uneconomic and/or insecure operation states.

2.2 Unobservable FDI Attack Model

In an unobservable FDI attack, the attacker aims to maliciously change the system states from x to $x+c$ without being detected by the bad data detector. In the absence of noise, the measurements for AC SE after such attacks, z^a, satisfy

$$z^a = h\left(x + c, \mathcal{G}\right) \tag{2.9}$$

where c is the $2n_b \times 1$ attack vector.

The measurements for DC SE after unobservable FDI attacks satisfy

$$z^a = z + Hc = H\left(x + c\right) \tag{2.10}$$

where c is the $n_b \times 1$ attack vector.

2.3 Optimal Power Flow

The optimal power flow (OPF) problem aims to solve the optimal power dispatch solution. The DC OPF problem can be written as:

$$\text{minimize} \quad \sum_{g=1}^{n_g} C_g\left(P_{Gg}\right) \tag{2.11}$$

$$\text{subject to} \quad GP_G - H\theta = P_D \tag{2.12}$$

$$-P_{\max} \leqslant \Gamma\theta \leqslant P_{\max} \tag{2.13}$$

$$P_{G,\min} \leqslant P_G \leqslant P_{G,\max} \tag{2.14}$$

where the optimization variables include the $n_b \times 1$ voltage angle vector θ and $n_g \times 1$ active power generation vector P_G. Furthermore, G is the $n_b \times n_g$ generator-to-bus connectivity matrix; $C_g(\cdot)$ is the quadratic cost function for generator g; H is the $n_b \times n_b$ dependency matrix between power injections and states θ; Γ is the $n_{br} \times n_b$ dependency matrix between power flows and states θ; P_D is the $n_b \times 1$ active power

load vector; $P_{G,\max}$ and $P_{G,\min}$ are $n_g \times 1$ maximum and minimum generation limit vectors, respectively; and $P_{\max}$ is the $n_{br} \times 1$ line rating vector.

The objective (2.11) is to minimize the total costs of all generators. Constraint (2.12) is the power balance constraint for each bus. Constraint (2.13) represents the power flow limit constraint for each transmission line. Note that power flow limit can either be a thermal limit, which prevents the transmission line from overheating, or a stability limit, which maintains synchronism and voltage stability among buses in the system. Constraint (2.14) represents the generation limit for each generator.

Note that (2.11)−(2.14) are DC OPF formulated with the $B-\theta$ method, in which the line power flow is calculated as the product of the dependency matrix of power flow and voltage angle B, and voltage angle vector θ. It can also be equivalently formulated using PTDF as follows:

$$\text{minimize} \quad \sum_{g=1}^{n_g} C_g\left(P_{Gg}\right) \tag{2.15}$$

$$\text{subject to} \quad \sum_{g=1}^{n_g} P_{Gg} = \sum_{i=1}^{n_b} P_{Di} \tag{2.16}$$

$$-P_{\max} \leqslant K(GP_G - P_D) \leqslant P_{\max} \tag{2.17}$$

$$P_{G,\min} \leqslant P_G \leqslant P_{G,\max} \tag{2.18}$$

where K is the $n_{br} \times n_b$ PTDF matrix, *i.e.,* the dependency matrix between power flows and power injections. In particular, constraint (2.16) is the power balance of the entire system.

Chapter 3

PRIOR WORK: PERFECT KNOWLEDGE FDI ATTACKS

In this chapter, we briefly review a closely related work [26] on unobservable FDI attacks assuming a perfect knowledge attacker. As in [26], we distinguish between two types of buses in the network: *load buses* that have load directly connected to that bus, and *non-load buses* with no load. The knowledge (denote as K1) and capabilities (denote as C1) of the attacker in [26] is described below:

K1. The attacker has knowledge of (i) the whole network topology; (ii) the cost, capacity, and operational status of all generators in the system; and (iii) the historical load data of the entire network.

C1. The attacker may choose a small area $\mathcal{S}$, which is a *subgraph* of the entire network $\mathcal{G}$, *i.e.*, a sub-network chosen in certain manner (see below for description) and bounded by load buses. The attacker may replace measurements inside $\mathcal{S}$ and has sufficient computational capability to perform SE.

Note that K1 and C1 model an omnipotent attacker who has perfect knowledge of the entire network. In practice, such knowledge and resources may be too strong to be achieved by attackers. However, the goal of this chapter is to evaluate the consequences of the worst-case FDI attacks. Thus, extra knowledge and capabilities are granted to attackers here.

3.1 Attack Design with Perfect Information

In [26], the authors formulate a bi-level optimization problem to determine the worst-case FDI attacks that can maximize the power flow on a target line. The bi-level

attack optimization problem is as follows:

$$\text{maximize} \quad P_l - \zeta \left\| c \right\|_0 \tag{3.1}$$

$$\text{subject to} \quad P = \Gamma\theta^* \tag{3.2}$$

$$\left\| c \right\|_0 \leqslant N_0 \tag{3.3}$$

$$-\tau P_D \leqslant Hc \leqslant \tau P_D \tag{3.4}$$

$$\{\theta^*, P_G^*\} = arg \left\{ \min_{\theta, P_G} \sum_{g=1}^{n_g} C_g \left(P_{Gg} \right) \right\} \tag{3.5}$$

$$\text{subject to} \quad GP_G - H\theta = P_D \qquad (\lambda) \tag{3.6}$$

$$-P_{\max} \leqslant \Gamma \left(\theta + c \right) \leqslant P_{\max} \quad \left(\mu^{\mp} \right) \tag{3.7}$$

$$P_{G,\min} \leqslant P_G \leqslant P_{G,\max} \qquad \left(\alpha^{\mp} \right) \tag{3.8}$$

where

P_G^* is the $n_g \times 1$ vector of optimal generation dispatch solved by DC OPF;

τ is the load shift factor;

N_0 is the l_0-norm constraint integer;

ζ is the weight of the norm of attack vector c.

λ is the $n_b \times 1$ dual variable vector for node balance constraints;

$\mu^{\mp}$ are the $n_{br} \times 1$ dual variable vectors for the negative and positive directions of thermal limit constraints, respectively.

$\alpha^{\mp}$ are the $n_g \times 1$ dual variable vectors for the minimum and maximum generation limit constraints, respectively.

The objective of the optimal attack problem is to maximize the power flow on the target line l while changing as few states as possible. In the first level, the attack

vector is chosen subject to the l_0-norm constraint of the attack vector in (3.3) and the load shift limitation in (3.4). In the second level, the system response to the attack determined in the first level is modeled via DC OPF in (3.6)−(3.8).

The bi-level optimization problem introduced above is non-linear and non-convex. For tractability, several constraints have been modified to convert the original formulation into an equivalent mixed-integer linear program (MILP). The modifications include:

1. Relax the l_0-norm constraint in (3.3) to an l_1-norm constraint with limit N_1 and linearize it by introducing a slack vector u as

$$c \leqslant u, \quad -c \leqslant u, \quad \sum_{i=1}^{n_b} u_i \leqslant N_1. \tag{3.9}$$

   The objective function (3.1) can be rewritten as

$$\underset{c,u}{\text{maximize}} \;\; P_l - \zeta \sum_{i=1}^{n_b} u_i \tag{3.10}$$

2. Replace the second level DC OPF problem by its Karush-Kuhn-Tucker (KKT) optimality conditions introduced in [43] as

$$\begin{aligned}
&(3.6) - (3.8) \\
\mathbf{0} &= \nabla \left[C_G(P_G) \right] + \nabla (G P_G - H\theta - P_D) \cdot \lambda \\
&\quad + \nabla \left[\Gamma(\theta + c) \mp P_{\max} \right] \cdot \mu^{\pm} \\
&\quad + \nabla \left(P_G - P_{G,\max} \right) \cdot \alpha^{+} + \nabla \left(P_{G,\min} - P_G \right) \cdot \alpha^{-} && (3.11) \\
\mathbf{0} &\leqslant \mu^{\pm} && (3.12) \\
\mathbf{0} &\leqslant \alpha^{\pm} && (3.13) \\
\mathbf{0} &= \operatorname{diag}(\mu^{\pm}) \left[\Gamma(\theta + c) \mp P_{\max} \right] && (3.14) \\
\mathbf{0} &= \operatorname{diag}(\alpha^{+}) \left(P_G - P_{G,\max} \right) && (3.15)
\end{aligned}$$

$$\mathbf{0} = \mathrm{diag}(\alpha^-)\,(P_{G,\min} - P_G) \tag{3.16}$$

where constraint (3.11) is the partial gradient optimal condition, (3.12) and (3.13) are the dual feasibility constraints, (3.14)–(3.16) represent the complementary slackness constraints.

3. Linearize the complementary slackness conditions in KKT by introducing a new vector δ of binary variables for dual variables and a large constant M as

$$[\delta_\mu^\pm; \delta_\alpha^\pm] \in \{0, 1\} \tag{3.17}$$

$$\begin{cases} \mu^\pm \leqslant M\delta_\mu^\pm \\ P_{\max} \mp \Gamma\,(\theta + c) \leqslant M(\mathbf{1} - \delta_\mu^\pm) \end{cases} \tag{3.18}$$

$$\begin{cases} \alpha^\pm \leqslant M\delta_\alpha^\pm \\ P_G^{\max} - P_G \leqslant M(\mathbf{1} - \delta_\alpha^+) \\ P_G - P_G^{\min} \leqslant M(\delta_\alpha^- - \mathbf{1}) \end{cases} \tag{3.19}$$

The whole problem then becomes a single level MILP with objective (3.10), subject to (3.21), (3.4), (3.6)–(3.9), (3.11)–(3.13), and (3.17)–(3.19).

In (3.5)−(3.8), B-θ method is used to formulate the DC OPF problem, in which the line power flow is calculated as the product of the dependency matrix of power flow and voltage angle B and the voltage angle vector θ. In contrast, PTDF can also be utilized to formulate DC OPF problem, where the line power flow is calculated as the product of PTDF matrix and power injection. Note that in this formulation, the variable vector θ is eliminated, and hence, the thermal limit constraints become independent of each other. The bi-level attack optimization problem is as follows:

$$\begin{aligned} &\text{maximize} \quad P_l - \zeta \left\| c \right\|_0 \\ &\text{subject to} \end{aligned} \tag{3.20}$$

$$P = K(GP_G^* - P_D) \tag{3.21}$$

$$\|c\|_0 \leqslant N_0 \tag{3.22}$$

$$-\tau P_D \leqslant Hc \leqslant \tau P_D \tag{3.23}$$

$$\{P_G^*\} = \arg\left\{\underset{P_G}{\text{minimize}} \sum_{g=1}^{n_g} C_g\left(P_{Gg}\right)\right\} \tag{3.24}$$

subject to

$$\sum_{g=1}^{n_g} P_{Gg} = \sum_{i=1}^{n_b} P_{Di} \tag{3.25}$$

$$-P_{\max} \leqslant K(GP_G - P_D + Hc) \leqslant P_{\max} \tag{3.26}$$

$$P_{G,\min} \leqslant P_G \leqslant P_{G,\max} \tag{3.27}$$

This problem can also be converted into MILP with modifications in (3.9)−(3.19).

3.2 Attack Implementation

Once the attack vector c is determined, the attacker can identify an attack subgraph $\mathcal{S}$ in which an unobservable attack can be implemented. The procedure for attack subgraph identification is as follows:

1. Identify a set of buses that correspond to non-zero entries of c (denoted as *center buses*).
2. Let $\mathcal{S}$ be the set of all center buses.
3. Extend $\mathcal{S}$ by including all branches and buses adjacent to center buses.
4. If any bus on the boundary of $\mathcal{S}$ is a *non-load bus* (*i.e.*, no load is present), extend $\mathcal{S}$ by including all branches and buses adjacent to this bus.
5. Repeat step 3 until all boundary buses are load buses.

This method is first proposed in [17]. Constructing $\mathcal{S}$ with this method ensures that nothing is changed outside of $\mathcal{S}$ while the changes needed at the non-center buses inside $\mathcal{S}$ are presented as load changes.

Given attacker's knowledge K1 and capabilities C1, the authors in [26] introduce the FDI attacks on DC SE as follows:

$$z_i^a = \begin{cases} z_i \,, & i \notin \mathcal{J} \\ z_i + H_i c \,, & i \in \mathcal{J} \end{cases} \tag{3.28}$$

where $\mathcal{J}$ denotes the set of measurements in $\mathcal{S}$.

Note that, this attack may not be unobservable to AC SE [44], but can be converted to an unobservable AC attacks as

$$z_i^a = \begin{cases} z_i \,, & i \notin \mathcal{J} \\ h_i(\hat{x} + c) \,, & i \in \mathcal{J} \end{cases} \tag{3.29}$$

where $\hat{x}$ is the state vector that the attacker estimated with measurements in $\mathcal{S}$. The method is first introduced in [17] and [26]. Furthermore, in our prior work [26, 29, 45], we have demonstrated that the consequences of the AC attacks track those of the original DC attacks.

Chapter 4

FALSE DATA INJECTION ATTACKS DESIGNED WITH LIMITED SYSTEM INFORMATION

In this chapter, two classes of unobservable FDI attacks designed with limited system information that can result in line overflow are studied. We first define the *attack sub-network* (denoted with $\mathcal{L}$) as the subset of the entire system where the attacker has perfect system knowledge. Such system information will be specified in sequel. The remaining network is defined as the *external network* denoted with $\mathcal{E}$. For each class of attacks, the theoretical attack model is presented, the implementation of attacks is provided, and the worst-case attack and its consequences on the physical system are exhaustively studied.

Throughout this chapter, we define an attack to be *successful* if the physical power flow on the target line is greater than the line rating post-attack.

4.1 False Data Injection Attacks with Limited External Network Information

In this section, a class of unobservable FDI attacks designed with limited external network information is studied. We introduce a bi-level optimization problem to find unobservable line-overflow FDI attacks on DC SE when the attacker's knowledge is mostly limited to a given sub-network. In particular, the attacker has access to the following information: (i) inside the sub-network, perfect system information including topology, load data, and generator data, (ii) outside the sub-network, estimated (*i.e.* possibly incorrect) generator data for the marginal generators, and (iii) estimated power transfer distribution factor (PTDF) data for the entire network. we demonstrate that using this bi-level optimization problem for limited information at-

tacks, an attacker can cause a line overflow in the IEEE 24-bus RTS system even given incomplete or inaccurate information.

4.1.1 Attack Strategy

In this subsection, we build upon the attack in [26] by replacing assumptions K1 and C1 with the following more limited assumptions on the attacker's knowledge (K2(a),K2(b)) and capability (C2):

K2(a) Attacker has perfect knowledge of the topology, the historical load, and the generator data including operational status, capacity and cost, inside a sub-network $\mathcal{L}$ of the entire network $\mathcal{G}$. This sub-network $\mathcal{L}$ is bounded by load buses.

K2(b) Attacker has estimated knowledge of (i) the power transfer distribution factor (PTDF) of $\mathcal{G}$; (ii) operational status, capacity and cost of only marginal generators (*i.e.*, generators that are re-dispatched after attack) outside $\mathcal{L}$.

C2 The attacker is restricted to a small sub-graph $\mathcal{S}$ within $\mathcal{L}$, *i.e.,* $\mathcal{S} \subseteq \mathcal{L}$, to access and modify measurements.

In practice, to achieve K2(a) and C2, an attacker can access to the databases of several adjacent substations. The K2(b) knowledge, however, is difficult to achieve even when the attacker has already hacked into the databases. Thus, one should note that extra knowledge is also granted here to evaluate worst-case limited information FDI attacks.

Notation: Let $\mathcal{N}$ denote an electric power network. The subset of buses, lines, and generators in $\mathcal{N}$ are denoted as $\mathcal{N}_b$, $\mathcal{N}_{br}$, and $\mathcal{N}_g$, respectively. For the limited information attack we study, the attacker is assumed to have perfect information of $\mathcal{L}$, a sub-network of the entire network $\mathcal{G}$. Thus, $\mathcal{L}_b$, $\mathcal{L}_{br}$, and $\mathcal{L}_g$ are the set of buses,

lines, and generators of $\mathcal{L}$, respectively. We define the remaining network that the attacker has limited information about as $\mathcal{E} = \mathcal{G} \setminus \mathcal{L}$. The set of buses in $\mathcal{E}$ is denoted as $\mathcal{E}_b$. We define the set of boundary buses in $\mathcal{L}$ as $\mathcal{B}$, such that each bus in $\mathcal{B}$ is connected to at least one bus in $\mathcal{E}$. The set of remaining buses in $\mathcal{L}$ is defined as the internal bus set $\mathcal{I} = \mathcal{L}_b \setminus \mathcal{B}$. We define the set of marginal generators in $\mathcal{E}$ as $\mathcal{E}_{g,m}$. The set of buses that each element of $\mathcal{E}_{g,m}$ connects to is denoted as $\mathcal{E}_{b,m}$. In addition, let $\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}$ denote the set of lines that connect a bus in $\mathcal{L}$ to a bus in $\mathcal{E}$. In the following, we distinguish different sets of elements (buses, lines, and generators) by using corresponding superscripts. For matrices, the superscript (E, F) represents a sub-matrix, such that E and F represent the sets of rows and columns of the original matrix, respectively. Finally, we use the subscript 0 to denote the variable values before the attack.

Rewriting OPF with Limited Information

To form an optimization problem that only uses the limited information as detailed in K2(a) and K2(b), we rewrite each line of the OPF as follows:

1. For the objective (3.5), the attacker is limited to only generators in $\mathcal{L}_g$ and $\mathcal{E}_{g,m}$ (knowledge K2(a) and K2(b)). Therefore, the objective can be rewritten as:

$$\text{minimize} \sum_{g \in \mathcal{L}_g} C_g\left(P_{Gg}\right) + \sum_{g \in \mathcal{E}_{g,m}} C_g\left(P_{Gg}\right); \tag{4.1}$$

2. For the thermal limit constraint (3.7), the attacker is limited to only the subset for lines in $\mathcal{L}_{br}$ as:

$$-P_{\max}^{\mathcal{L}_{br}} \leqslant \Gamma^{(\mathcal{L}_{br},\mathcal{L}_b)}\left(\theta^{\mathcal{L}_b} + c^{\mathcal{L}_b}\right) \leqslant P_{\max}^{\mathcal{L}_{br}}; \tag{4.2}$$

3. For the generation limit constraint (3.8), only those for generators in $\mathcal{L}_g$ and

$\mathcal{E}_{g,m}$ can be formulated as:

$$P_{G,\min}^{\mathcal{L}_g} \leqslant P_G^{\mathcal{L}_g} \leqslant P_{G,\max}^{\mathcal{L}_g} \tag{4.3}$$

$$P_{G,\min}^{\mathcal{E}_{g,m}} \leqslant P_G^{\mathcal{E}_{g,m}} \leqslant P_{G,\max}^{\mathcal{E}_{g,m}}. \tag{4.4}$$

4. For the power balance constraint (3.6), the attacker can only formulate those for the buses in $\mathcal{I}$ and $\mathcal{B}$ as:

$$G^{(\mathcal{I},\mathcal{L}_g)} P_G^{\mathcal{L}_g} - H^{(\mathcal{I},\mathcal{L}_b)} \theta^{\mathcal{L}_b} = P_D^{\mathcal{I}} \tag{4.5}$$

$$G^{(\mathcal{B},\mathcal{L}_g)} P_G^{\mathcal{L}_g} - [H^{(\mathcal{B},\mathcal{L}_b)} \; H^{(\mathcal{B},\mathcal{E}_b)}] \begin{bmatrix} \theta^{\mathcal{L}_b} \\ \theta^{\mathcal{E}_b} \end{bmatrix} = P_D^{\mathcal{B}}; \tag{4.6}$$

In constraint (4.6), the term $H^{(\mathcal{B},\mathcal{E}_b)}\theta^{\mathcal{E}_b}$ is a vector of the power injections that flow from $\mathcal{E}$ to the buses in $\mathcal{B}$; however, this information is not available to the attacker. For each bus in $\mathcal{B}$, such a term equals to the sum of power flows on lines in $\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}$ that connects to that bus, *i.e.,* $H^{(\mathcal{B},\mathcal{E}_b)}\theta^{\mathcal{E}_b} = A_{KN}^{(\mathcal{B},\mathbb{W}_{\mathcal{B}}^{\mathcal{E}})} P^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}}$, where A_{KN} is the $n_b \times n_{br}$ line-to-bus connectivity matrix. The vector of power flow in $\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}$, $P^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}}$, can be computed as $P^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}} = K^{(\mathbb{W}_{\mathcal{B}}^{\mathcal{E}},\mathcal{G}_b)}(GP_G - P_D)$. The attacker has no knowledge of load in $\mathcal{E}_b$ and generation of non-marginal generators in $\mathcal{E}$. However, if the attacker's knowledge K2(b) is correct and complete, the loads and the output of non-marginal generators in $\mathcal{E}$ will remain unchanged after attack since only marginal generators will re-dispatch. We later study the errors that can result from incorrect attack information. Then $P^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}}$ can be formulated with the sum of $P_0^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}}$, *i.e.,* the power flow on the lines before attack, and the incremental power flow resulting from re-dispatch of generators in $\mathcal{L}_g$ and $\mathcal{E}_{g,m}$ as:

$$\begin{aligned} P^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}} = P_0^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}} &+ K^{(\mathbb{W}_{\mathcal{B}}^{\mathcal{E}},\mathcal{L}_b)} G^{(\mathcal{L}_b,\mathcal{L}_g)} \Delta P_G^{\mathcal{L}_g} \\ &+ K^{(\mathbb{W}_{\mathcal{B}}^{\mathcal{E}},\mathcal{E}_{b,m})} G^{(\mathcal{E}_{b,m},\mathcal{E}_{g,m})} \Delta P_G^{\mathcal{E}_{g,m}}, \end{aligned} \tag{4.7}$$

where $\Delta P_G^{\mathcal{L}_g}$ and $\Delta P_G^{\mathcal{E}_{g,m}}$ represent the output changes of generators in $\mathcal{L}_g$ and $\mathcal{E}_{g,m}$, respectively, with $\Delta P_G^{\mathcal{L}_g} = P_G^{\mathcal{L}_g} - P_{G,0}^{\mathcal{L}_g}$ and $\Delta P_G^{\mathcal{E}_{g,m}} = P_G^{\mathcal{E}_{g,m}} - P_{G,0}^{\mathcal{E}_{g,m}}$. For ease of expression, we define the dependency matrix between the incremental injections and generations as A_{GB}, such that the sub-matrices of A_{GB} for buses in $\mathcal{B}$ and generations in $\mathcal{L}_g$ (denoted as $A_{GB}^{(\mathcal{B},\mathcal{L}_g)}$) and in $\mathcal{E}_{g,m}$ (denoted as $A_{GB}^{(\mathcal{B},\mathcal{E}_{g,m})}$), respectively, are:

$$A_{GB}^{(\mathcal{B},\mathcal{L}_g)} = A_{KN}^{(\mathcal{B},\mathbb{W}_{\mathcal{B}}^{\mathcal{E}})} K^{(\mathbb{W}_{\mathcal{B}}^{\mathcal{E}},\mathcal{L}_b)} G^{(\mathcal{L}_b,\mathcal{L}_g)} \tag{4.8}$$

$$A_{GB}^{(\mathcal{B},\mathcal{E}_{g,m})} = A_{KN}^{(\mathcal{B},\mathbb{W}_{\mathcal{B}}^{\mathcal{E}})} K^{(\mathbb{W}_{\mathcal{B}}^{\mathcal{E}},\mathcal{E}_{b,m})} G^{(\mathcal{E}_{b,m},\mathcal{E}_{g,m})}. \tag{4.9}$$

We define $P_{\text{inj},0}^{\mathcal{B}}$ to represent a vector of constant values which corresponds to the portion of power injections in $\mathcal{B}$ that are from the entire network loads and the output of the non-marginal generators in $\mathcal{E}$, *i.e.,* $P_{\text{inj},0}^{\mathcal{B}} = A_{KN}^{(\mathcal{B},\mathbb{W}_{\mathcal{B}}^{\mathcal{E}})} P_0^{\mathbb{W}_{\mathcal{B}}^{\mathcal{E}}} - A_{GB}^{(\mathcal{B},\mathcal{L}_g)} P_{G,0}^{\mathcal{L}_g} - A_{GB}^{(\mathcal{B},\mathcal{E}_{g,m})} P_{G,0}^{\mathcal{E}_{g,m}}$. Therefore, the constraint (4.6) can be rewritten as

$$\begin{aligned} &G^{(\mathcal{B},\mathcal{L}_g)} P_G^{\mathcal{L}_g} - H^{(\mathcal{B},\mathcal{L}_b)} \theta^{\mathcal{L}_b} - A_{GB}^{(\mathcal{B},\mathcal{L}_g)} P_G^{\mathcal{L}_g} - A_{GB}^{(\mathcal{B},\mathcal{E}_{g,m})} P_G^{\mathcal{E}_{g,m}} \\ &\quad = P_D^{\mathcal{B}} + P_{\text{inj},0}^{\mathcal{B}}. \end{aligned} \tag{4.10}$$

Attack Optimization Problem under Limited Information

The limited information bi-level attack optimization problem can now be rewritten as

$$\text{maximize} \quad P_l - \zeta \left\| c^{\mathcal{L}_b} \right\|_0 \tag{4.11}$$

$$\text{subject to} \quad P^{\mathcal{L}_{br}} = \Gamma^{(\mathcal{L}_{br},\mathcal{L}_b)} \theta^{\mathcal{L}_b *} \tag{4.12}$$

$$\left\| c^{\mathcal{L}_b} \right\|_0 \leqslant N_0, \quad c^{\mathcal{B}} = \mathbf{0} \tag{4.13}$$

$$-\tau P_D^{\mathcal{L}_b} \leqslant H^{(\mathcal{L}_b,\mathcal{L}_b)} c^{\mathcal{L}_b} \leqslant \tau P_D^{\mathcal{L}_b} \tag{4.14}$$

$$\begin{aligned} &\left\{ \theta^{\mathcal{L}_b *}, P_G^{\mathcal{L}_g *}, P_G^{\mathcal{E}_{g,m} *} \right\} \\ &= \arg \left\{ \underset{\theta^{\mathcal{L}_b}, P_G^{\mathcal{L}_b}, P_G^{\mathcal{E}_{g,m}}}{\text{minimize}} \sum_{g \in \mathcal{L}_g} C_g(P_{Gg}) + \sum_{g \in \mathcal{E}_{g,m}} C_g(P_{Gg}) \right\} \end{aligned} \tag{4.15}$$

$$\text{subject to } (4.5), (4.2), (4.3) - (4.4), (4.10)$$

Note that, in addition to the changes in the OPF sub-problem described in Sec. 4.1.1, we have also changed the constraint on the attack vector in (4.13) to require the attack to be within the sub-network $\mathcal{L}$.

The bi-level optimization problem introduced above is non-linear and non-convex. For tractability, we modify several constraints to convert the original formulation into an equivalent mixed-integer linear problem as in [26]. The modifications include: (a) relaxing the l_0-norm constraint in (4.13) to an l_1-norm constraint with limit N_1 and linearizing it by introducing a slack vector u; (b) replacing the second level DC OPF problem by its Karush-Kuhn-Tucker (KKT) optimality conditions; and (c) linearizing the complementary slackness conditions in the KKT conditions by introducing new binary variables δ for dual variables and a large constant M.

4.1.2 Discussion

As stated in Sec. 4.1.1, each of the attacker's knowledge and capabilities in K2(a), K2(b) and C2 may cause the attack solved by the limited information optimization problem to be suboptimal compared to the perfect information optimization problem. Recall that in assumption K2(b), the attacker has an estimated knowledge of certain system parameters, which may be incomplete or inaccurate. As a result, the attacker may obtain a wrong evaluation of attack consequences. We focus on the following three scenarios to show that limited knowledge can possibly provide incorrect generation re-dispatch in the optimization problem, which in turn can finally lead to wrong maximal power flow solution.

1. Congested lines in $\mathcal{E}$: the attacker cannot capture the changes of generation dispatch resulting from the thermal limit constraints of congested lines in $\mathcal{E}$.

2. Wrong external marginal generators: the attacker may choose a wrong $\mathcal{E}_{g,m}$, and hence, the approximation of the power injection in (4.10) will be incorrect.

3. Wrong PTDF: the attacker may obtain wrong PTDF, K, information due to real-time topology changes in $\mathcal{E}$.

Despite the limitations listed above, there is still a large chance that the system operation will be worsened by such attacks, as illustrated in the following section, for both perfect limited information and inaccurate knowledge scenarios.

4.1.3 Numerical Results

In this subsection, we illustrate the effect of attacks designed with the optimization problem in Sec. 4.1.1. We first solve the optimization problem to find the optimal attack vector $c^{\mathcal{L}_b}$ inside $\mathcal{L}$. Subsequently, we use the attack vector $c^{\mathcal{L}_b}$ to simulate an AC attack. The test system is the IEEE 24-bus reliability test system (RTS). We assume: (i) the system is operating under optimal power flow; and (ii) the loads of the system are constant and are equivalent to the historic load data. We use MATPOWER to run AC power flow and AC OPF. The optimization problem is solved with CPLEX.

Note that, to model realistic power systems, we assume that there are congested lines prior to the attack and the attacker chooses one line in $\mathcal{L}$ as the target to maximize power flow. This is achieved in simulation by uniformly reducing all line ratings by 50% except for that of line 11 (to ensure no congestion in $\mathcal{E}$). The test system is shown in Fig. 4.1. When the attacker has perfect knowledge of $\mathcal{L}$, we refer to this case as *perfect case*.

We illustrate our results for the following choice of parameters: the weight of the l_1-norm of attack vector in (4.11), ζ, is set to 1% of the original power flow value

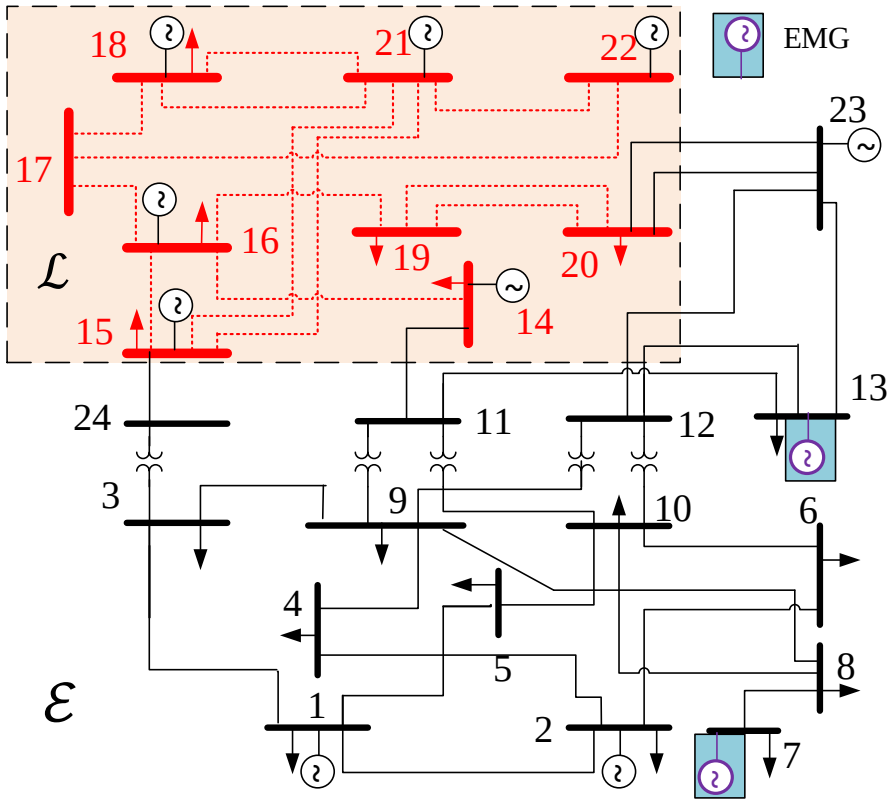


Figure 4.1: IEEE 24-Bus RTS System Decomposed into The Attack Sub-network and External Network.

of the target line; and the load shift factor in (4.14), τ, is set to 10%. We focus on 2 scenarios: (1) attacker's estimate of information in $\mathcal{E}$ is perfect (perfect case); (2) attacker's estimate is inaccurate (three different imperfect cases).

Scenario 1: *Perfect Estimated Knowledge of* $\mathcal{E}$

We first compare the attack consequences determined by the optimization problems for two cases: (i) limited but perfect knowledge (henceforth identified as *local case*); (ii) complete system knowledge as in [26] ((identified as *global case*)). This comparison for target line 28 is illustrated in Fig. 4.2 with sub-plots (a)−(c) illustrating the maximal power flow, the l_1-norm, and the l_0-norm of the attack vector, respectively, as a function of the l_1-norm constraint N_1. In each subplot, we plot

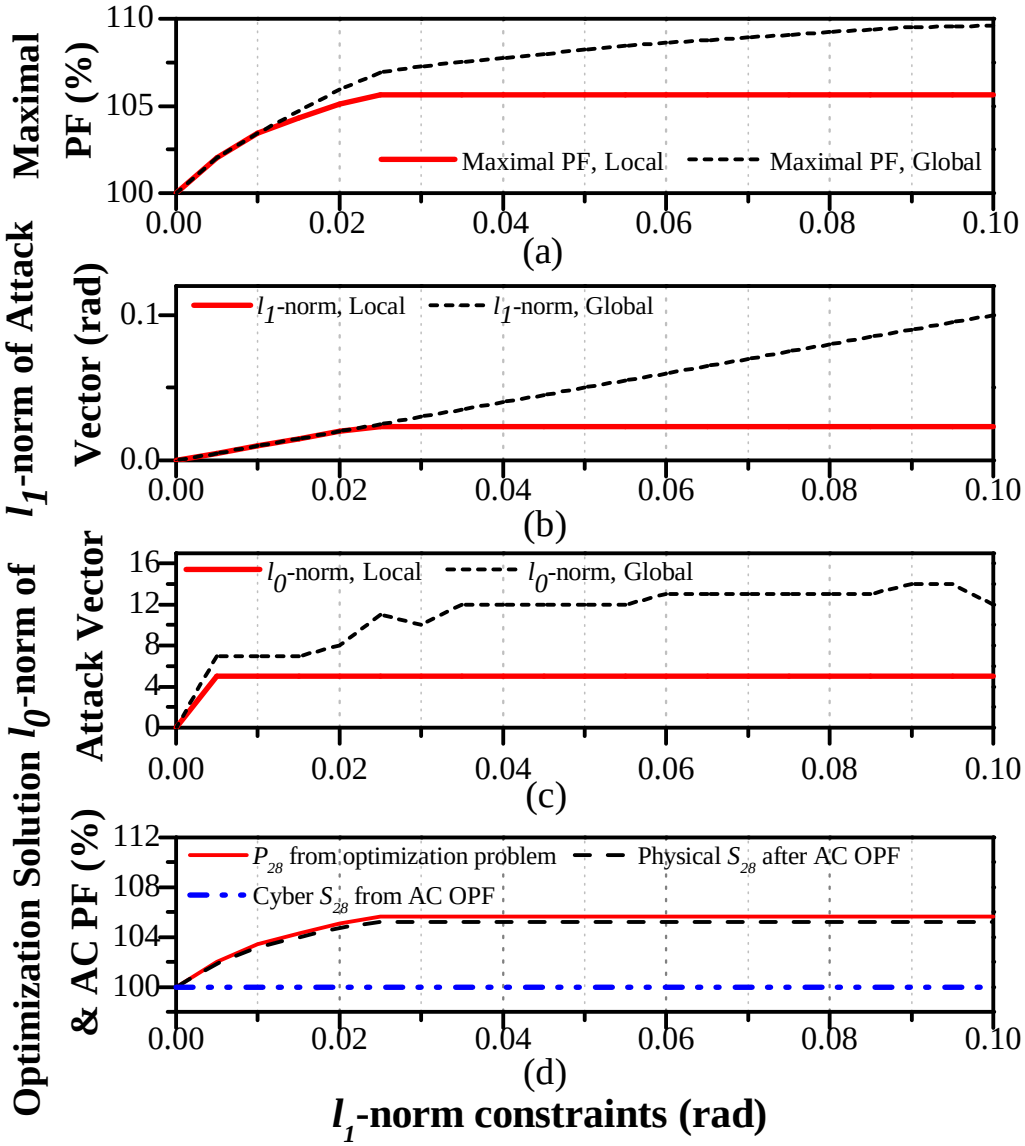


Figure 4.2: The Maximum Power Flow on The Target Line, The l_1-Norm and l_0-Norm of Solved Attack Vector c V.S. The l_1-Norm Constraint (N_1) When Load Shift (τ) Is Limited by 10%; Target Line 28 (Bus 16 – Bus 17).

two curves, one for local and one for global. As N_1 increases, the maximal PF on the target line, the l_1-norm and l_0-norm of attack vector in both limited and perfect information attacks increase before N_1 reaches 0.025. However, when $N_1 > 0.025$, getting a larger overflow on the target line requires measurements in both $\mathcal{L}$ and $\mathcal{E}$ to be modified. Therefore, target line power flow resulting from global case can in-

crease after $N_1 > 0.025$, while that resulting from local case remains unchanged due to limited attack resources.

Fig. 4.2 (d) illustrates the result of an AC attack using the attack vector $c^{\mathcal{L}_b}$. The system re-dispatch in response to this attack is via AC OPF in this case. It is shown that although the attack vector is solved by a linear optimization problem, it can still cause overflows in the AC system.

Finally, we exhaustively test all 13 lines inside $\mathcal{L}$ as targets in the perfect test case with $N_1 = 0.05$ and $\tau = 10\%$. We observe that attacks that target lines 23 and 28, which are congested prior to the attack, can successfully cause overflows on target lines. On the other hand, attacks that target lines 25, 26, 30, and 31 do not result in target line overflows, however, they lead to overflows on lines 23 and 28. This observation indicates that congested lines are more vulnerable to this class of attacks, and therefore should be better protected.

Scenario 2: *Inaccurate Estimated Knowledge of* $\mathcal{E}$

We use three specific test cases to illustrate the effects of incomplete or inaccurate information conditions highlighted in Sec. 4.1.2. The modifications of the three test cases on the perfect test case are:

- **Case 1**: decrease the rating of line 11 by 50% and that of line 7 by 60% to create two congested lines in $\mathcal{E}$;
- **Case 2**: assume the $\mathcal{E}_{g,m}$ are generators that on buses 1 and 23 in contrast with the correct ones on buses 7 and 13;
- **Case 3**: use wrong PTDF matrix which corresponds to a topology with an outage on line 6.

For each test case, we first solve the limited information attack optimization problem to obtain the attack vector $c^{\mathcal{L}_b}$ as well as the dispatch results solved in second level problem DC OPF. Then we create the false load patterns determined by $c^{\mathcal{L}_b}$ and use such patterns to run a DC OPF as in (2.11)−(2.14). Throughout, N_1 is set to be 0.05, and the load shift is $\tau = 10\%$. We observe that the generation re-dispatch given by the attack optimization problems in all three cases described above are different from the actual system re-dispatch. Such differences may lead to incorrect post attack power flow estimation. We compare the power flow on target line solved with our limited information optimization problems (denoted as *Computed PF*), and that solved with the system response DC OPF (denoted as *Actual PF*) in Table 4.1 for target line 28. From the Table, we can see that, all these three cases can distort the attacker's evaluations of target line power flow. However, even such limited inaccurate attacks can cause damage to a congested system.

Table 4.1: Comparison of Computed PF and Actual PF within Attacks for Target Line 28.

Case	Actual PF	Computed PF
Perfect	105.64%	105.64%
1	104.60%	105.64%
2	104.82%	105.95%
3	104.95%	105.90%

4.2 False Data Injection Attack with No External Network Information

In this section, a class of unobservable line overflow FDI attacks designed with no external network information is studied. The attacker is assumed to have per-

fect information only inside the attack sub-network, but can collect historical data of the loads and generations inside the sub-network. To this end, we introduce a new vector, *pseudo-boundary injections*, to represent the power flow delivered from external network to the boundary buses and compute the power flow inside the attack sub-network with power injections at buses inside the attack sub-network and the pseudo-boundary injections. The multiple linear regression method is then utilized to learn the relationship between pseudo-boundary injections and power injections in attack sub-network with historical data. We introduce a bi-level attack optimization problem similar as [26] in which the second level DC OPF is modeled with knowledge inside attack sub-network and the pseudo-boundary injections. The existence of linear relationship between pseudo-boundary injections and power injections inside the attack sub-network under certain circumstance is proved. The upper bound on the target line flow within the designed attack is provided. We demonstrate that the attacks designed with the proposed attack strategy can result in line overflows in IEEE 24-bus RTS system, IEEE 118-bus system, and IEEE 2383-bus (Polish) system.

The limited information attack problem studied in this chapter is similar to the "seamless" market problem [46] which aims to achieve maximum social welfare across several adjacent markets, while allowing each market to model its own system, exchanging boundary information with its neighbors. In order to predict the behavior of the adjacent market, in [47, 46], linear regression model is used. However, in contrast to the market problem that requires perfect prediction of the external network, the attacker only needs partial prediction to overload a target line. In this section, we demonstrate that even if the prediction of the external network re-dispatch is inaccurate, the attacker can still cause overflow on the target line.

4.2.1 Attack Strategy

In this subsection, we build upon the attack in Chapter 3 by replacing assumptions K1 and C1 with the following limited assumptions on the attacker's knowledge (K3) and capability (C3):

K3 Within a sub-network $\mathcal{L}$, the attacker has perfect knowledge of the topology, historical load data, generator data including operational status, capacity, cost, and historical dispatch information, and locational marginal price (LMP). In particular, we assume that the attacker has enough historical data to perform the multiple linear regression described in the sequel. This sub-network $\mathcal{L}$ is bounded by load buses.

C3 The attacker may modify measurements within an attack sub-graph $\mathcal{S}$ within $\mathcal{L}$, *i.e.,* $\mathcal{S} \subseteq \mathcal{L}$.

As stated in Sec. 4.1, to achieve K3 and C3, the attacker can access the databases of several adjacent substations. In addition, historical data can also be collected from the historical market data released by ISOs, *e.g.,* ERCOT releases all data after six months.

An example of the attack sub-network in the IEEE 24-bus RTS system is shown in Fig. 4.3.

Notation: The area outside $\mathcal{L}$, where the attacker has no knowledge, is denoted as the external network, *i.e.*, $\mathcal{E} = \mathcal{G} \setminus \mathcal{L}$. We define the set of boundary buses in $\mathcal{L}$ as $\mathcal{B}$, such that each bus in $\mathcal{B}$ is connected to at least one bus in $\mathcal{E}$. The set of remaining buses in $\mathcal{L}$ is defined as the internal bus set $\mathcal{I} = \mathcal{L} \setminus \mathcal{B}$. For any vector or matrix associated with the entire network $\mathcal{G}$ such as c, G, H, K, P, P_G, and P_D, we write the equivalent parameters corresponding to the sub-network $\mathcal{L}$ with $\bar{(\cdot)}$. For

example, $\bar{H}$ refers to the dependency matrix between power injection measurements and state variables only inside $\mathcal{L}$. Sub-vectors are denoted by subscripts with the corresponding set of elements (buses, lines, or generators). Sub-matrices are denoted with a subscript giving the set of rows, and a superscript giving the set of columns.

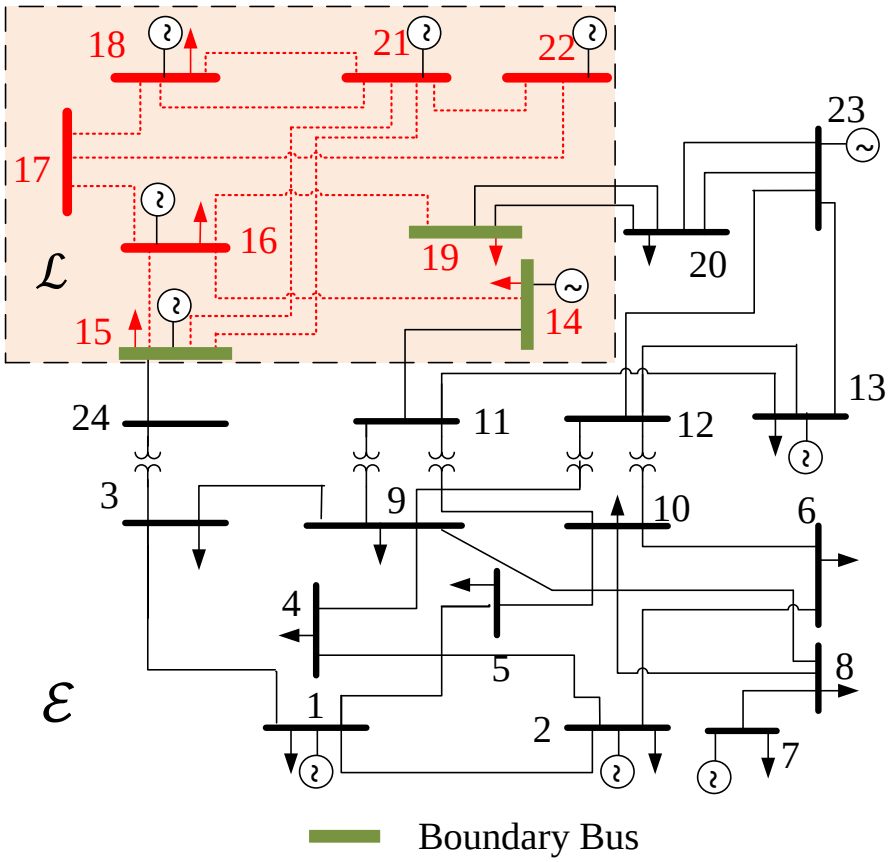


Figure 4.3: IEEE 24-Bus RTS System Decomposed into Attack Sub-Network and Attack External Network.

System Power Flow with Localized Information

According to assumption K3, the attacker only has knowledge inside $\mathcal{L}$. Therefore, the attacker cannot calculate the line power flow inside $\mathcal{L}$ with (3.21) since both the PTDF matrix K of the network $\mathcal{G}$ and the subset of power injections in external

network $\mathcal{E}$ are unavailable to attacker. To form the line power flow with K3, we introduce a vector of *pseudo-boundary injection* $\bar{P}_{I,\mathcal{B}}$ as illustrated in Fig. 4.4. The ith entry of $\bar{P}_{I,\mathcal{B}}$, namely $\bar{P}_{I,i}$, corresponding to boundary bus i, represents the sum of power flows delivered from $\mathcal{L}$ to $\mathcal{E}$ at boundary bus i, $i \in \mathcal{B}$, as

$$\bar{P}_{I,i} = \sum_{k \in \mathbb{W}_i^{\mathcal{E}}} P_k \tag{4.16}$$

where $\mathbb{W}_i^{\mathcal{E}}$ represents the lines located in $\mathcal{E}$ that are connected to boundary bus i.

Using (4.16), the vector of line power flows in $\mathcal{L}$ can be written as

$$\bar{P} = \bar{K}^{\mathcal{I}}(\bar{G}_{\mathcal{I}}\bar{P}_G - \bar{P}_{D,\mathcal{I}}) + \bar{K}^{\mathcal{B}}(\bar{G}_{\mathcal{B}}\bar{P}_G - \bar{P}_{D,\mathcal{B}} - \bar{P}_{I,\mathcal{B}}) \tag{4.17}$$

where $\bar{K}$ is split into column-wise sub-matrices $\bar{K}^{\mathcal{I}}$ and $\bar{K}^{\mathcal{B}}$, and $\bar{G}$ is split into row-wise sub-matrices $\bar{G}_{\mathcal{I}}$ and $\bar{G}_{\mathcal{B}}$, both corresponding to buses in $\mathcal{I}$ and $\mathcal{B}$, respectively. This equation can be further simplified as

$$\bar{P} = \bar{K}(\bar{G}\bar{P}_G - \bar{P}_D) - \bar{K}^{\mathcal{B}}\bar{P}_{I,\mathcal{B}}. \tag{4.18}$$

Multiple Linear Regression

The optimal line overflow attack introduced in Sec. 3.1 involves determining the attack vector in the first level and estimating the system response to the attack via the whole system DC OPF in the second level. However, due to limited knowledge, the attacker must predict the response of the OPF using only local knowledge. The OPF may be reformulated to include power balance, thermal limit, and generation limit constraints only in $\mathcal{L}$, and apply (4.18) to capture all effects in the external network through the pseudo-boundary injections $\bar{P}_{I,\mathcal{B}}$. However, with this formulation, the attacker still cannot predict how the attack affects $\bar{P}_{I,\mathcal{B}}$ since it depends on both power injections in $\mathcal{L}$ and $\mathcal{E}$. Therefore, before the attack is executed, the attacker cannot estimate the system re-dispatch after the attack accurately.

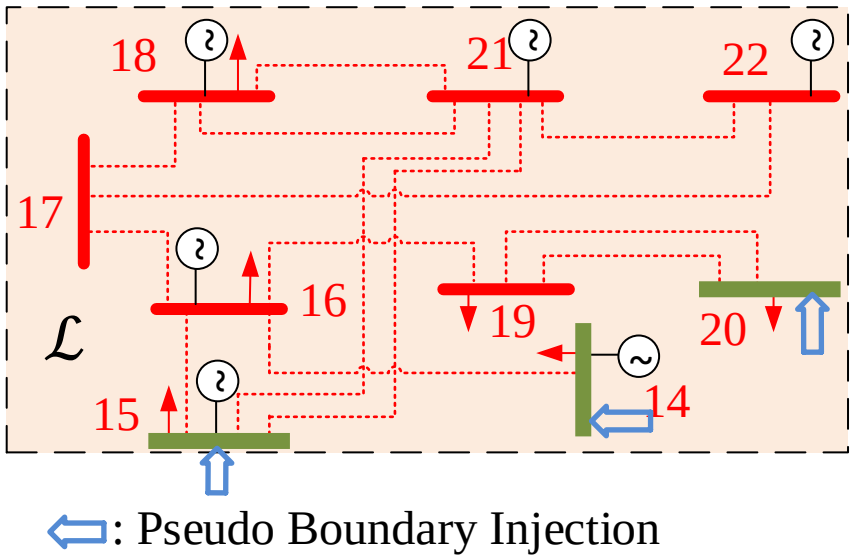


Figure 4.4: Equaivalent Attack Sub-Network in IEEE 24-Bus RTS System.

If the attacker can obtain a large amount of historical power injections and pseudo-boundary injections data in $\mathcal{L}$ (for example, by observing the system over a long time), it can learn a functional relationship between pseudo-boundary injection, $\bar{P}_{I,\mathcal{B}}$, and power injections inside $\mathcal{L}$. The attacker can then predict the pseudo-boundary injections with the power injection in $\mathcal{L}$ as

$$\hat{\bar{P}}_{I,\mathcal{B}} = \hat{F}\left(\bar{G}\bar{P}_G - \bar{P}_D\right) + \hat{f}_0 \tag{4.19}$$

where $[\hat{f}_0 \ \hat{F}]$ represent an affine relationship, and $\hat{\bar{P}}_{I,\mathcal{B}}$ is the attacker's prediction of pseudo-boundary injection by capturing the functional relationship via a linear model. Note that the historical pseudo-boundary injections can be computed with data in $\mathcal{L}$ as

$$\bar{P}_{I,i} = \sum_{g\in\mathbb{G}_i} \bar{P}_{G,g} - \bar{P}_{D,i} - \sum_{k\in\mathbb{W}_i^{\mathcal{L}}} \bar{P}_k \tag{4.20}$$

where $\mathbb{G}_i$ is the set of generators connected to bus i, and $\mathbb{W}_i^{\mathcal{L}}$ is the set of lines in $\mathcal{L}$ that connected to bus i. We suppose the attacker uses multiple linear regression to

learn $[\hat{f}_0\ \hat{F}]$.

Multiple linear regression is a statistical method to find a linear relationship between multiple inputs and single output [48]. Take boundary bus i for an example. Let the output $y_i = \bar{P}_{I,i}$ and inputs $\mathbf{x}^T = \bar{G}\bar{P}_G - \bar{P}_D$. At one instance of time t, $y_{i,t}$ satisfies

$$y_{i,t} = \begin{bmatrix} x_{1,t} & x_{2,t} & \dots & x_{k,t} \end{bmatrix} \begin{bmatrix} f_{i,1} \\ f_{i,2} \\ \vdots \\ f_{i,k} \end{bmatrix} + f_{i,0} + \varepsilon_{i,t} \tag{4.21}$$

where $f_{i,j},\ j = 0, ..., k$, are regression coefficients for boundary bus i, and $\varepsilon_{i,t}$ is random error. In the following, we let $\hat{F}_i = [f_{i,1}\ \ f_{i,2}\ ...\ f_{i,k}]$ be the coefficient vector for boundary bus i.

Consider a problem with an $m \times 1$ observed output vector $\mathbf{y}_i$, and an $m \times k+1$ input matrix $\mathbf{X} = [\mathbf{1}\ \mathbf{x}_1\ ...\ \mathbf{x}_k]$. The relationship in (4.21) can be written in matrix notation as

$$\mathbf{y}_i = \mathbf{X}\left[\hat{f}_{i,0}\ \ \hat{F}_i\right]^T + \varepsilon_i. \tag{4.22}$$

Least squares estimation (LSE) can be used to estimate the regression coefficients $\hat{F}_i$ in (4.22) as

$$[\hat{f}_{i,0}\ \ \hat{F}_i]^T = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}_i. \tag{4.23}$$

Note that as we stated in K3, we assume the attacker has enough historical data. Therefore, $\left(\mathbf{X}^T\mathbf{X}\right)$ is full-rank. We repeatedly use this process to obtain $[\hat{f}_{i,0}\ \hat{F}_i]$ for each i, $i \in \mathcal{B}$. Thus, the attacker can use historical data to obtain the estimate $\hat{F}$, such that $[\hat{f}_0\ \hat{F}] = [\hat{f}_{i,0}\ \hat{F}_i]$, $\forall i \in \mathcal{B}$. The dimension of $[\hat{f}_0\ \hat{F}]$ is $n_{\mathcal{B}} \times (k+1)$, where $n_{\mathcal{B}}$ is the number of boundary buses. The attacker now can approximate (4.18) as

$$\bar{P} = \bar{K}\left(\bar{G}\bar{P}_G - \bar{P}_D\right) - \bar{K}^{\mathcal{B}}\left(\hat{F}\left(\bar{G}\bar{P}_G - \bar{P}_D\right) + \hat{f}_0\right). \tag{4.24}$$

In the following subsection, (4.24) is used to evaluate the vulnerability to attacker with limited information.

Attack Optimization Problem under Localized Information

In this part, we introduce a bi-level attack optimization problem to formulate the limited information attack. The first level determines the attack vector in $\mathcal{L}$ that maximize target line flow and the second level represents system re-dispatch after attack via DC OPF formulated with only information in $\mathcal{L}$. However, since the attacker does not have knowledge of either the topology or the generator information in $\mathcal{E}$, we assume that the attacker only minimizes the total cost of generation in $\mathcal{L}$ and approximates the effect of the total generation cost in $\mathcal{E}$ as the total cost of the pseudo-boundary injections in the second level modified OPF. For boundary bus i, this cost is estimated as the product of the LMP, λ_i, and the pseudo-boundary injection at bus i. The limited information bi-level attack optimization problem is as follows:

$$\underset{\bar{c},\bar{P}}{\text{maximize}} \;\; \bar{P}_l - \zeta \left\|\bar{c}\right\|_0 \tag{4.25}$$

subject to

$$\bar{P} = \bar{K}\left(\bar{G}\bar{P}_G^* - \bar{P}_D\right) - \bar{K}^{\mathcal{B}}\bar{P}_{I,\mathcal{B}}^* \tag{4.26}$$

$$\left\|\bar{c}\right\|_0 \leq N_0, \quad \bar{c}_{\mathcal{B}} = \mathbf{0} \tag{4.27}$$

$$-\tau\bar{P}_D \leq \bar{H}\bar{c} \leq \tau\bar{P}_D \tag{4.28}$$

$$\left\{\bar{P}_G^*, \bar{P}_{I,\mathcal{B}}^*\right\} = \arg\left\{\underset{\bar{P}_G,\bar{P}_{I,\mathcal{B}}}{\text{minimize}} \sum_{g\in\mathcal{L}} C_g\left(\bar{P}_{Gg}\right) + \sum_{i\in\mathcal{B}} \lambda_i \bar{P}_{I,i}\right\} \tag{4.29}$$

subject to

$$\bar{P}_{I,\mathcal{B}} = \hat{F}\left(\bar{G}\bar{P}_G - \bar{P}_D + \bar{H}\bar{c}\right) + \hat{f}_0 \tag{4.30}$$

$$\sum_{g\in\mathcal{L}} \bar{P}_{G,g} - \sum_{i\in\mathcal{B}} \bar{P}_{I,i} = \sum_{i\in\mathcal{L}} \bar{P}_{D,i} \tag{4.31}$$

$$-\bar{P}_{\max} \leq \bar{K}\left(\bar{G}\bar{P}_G - \bar{P}_D + \bar{H}\bar{c}\right) - \bar{K}^{\mathcal{B}}\bar{P}_{I,\mathcal{B}} \leq \bar{P}_{\max} \tag{4.32}$$

$$\bar{P}_{G,\min} \leq \bar{P}_G \leq \bar{P}_{G,\max} \tag{4.33}$$

where (4.29) captures the modified OPF objective as the first term represents the total cost of generation in $\mathcal{L}$ and the second term is the total cost of pseudo-boundary injections. Constraint (4.30) represents the attacker's prediction of the pseudo-boundary injection after attack resulting from the counterfeit loads. Note that, in the cyber system (OPF with attack vector), the power injections in $\mathcal{L}$ is $\bar{G}\bar{P}_G - \bar{P}_D + \bar{H}\bar{c}$, thus, the corresponding pseudo-boundary injection should respond to these injections with attack. In (4.32), we directly write the second term with $K^{\mathcal{B}}\bar{P}_{I,\mathcal{B}}$ instead of $K^{\mathcal{B}}\left(\hat{F}\left(\bar{G}\bar{P}_G - \bar{P}_D + \bar{H}c\right) + \hat{f}_0\right)$. In addition, we have changed the constraint on the attack vector in (4.27) to limit the attack to be within the sub-network $\mathcal{L}$.

As with the bi-level optimization problem for perfect information, (4.25)−(4.33) is non-linear and non-convex. We employ the same modifications as detailed in Sec. 3.1 to convert it into a MILP.

Note that attacker can only overload lines in $\mathcal{L}$. The attack optimization problem ensures that only measurements inside $\mathcal{L}$ can be changed by attacker. The post-attack system re-dispatch (OPF), on the other side, forces all the cyber line power flows within the thermal limits. Therefore, the attacker can only hide the physical overflow inside $\mathcal{L}$ with FDI attack.

4.2.2 Justification of the Localized Information FDI Attacks

In this subsection, we make a distinction between the *physical* system, as it actually exists, and the *cyber* system, as seen by the control center, which may differ from the physical system due to the FDI attack. We use the superscripts p and c to

denote the physical and cyber power flows, respectively. Due to limited information, the attacker can only use data in $\mathcal{L}$ to compute the physical and cyber power flows which may be different from the actual values. Therefore, we refer the physical and cyber power flows computed by the attacker as *attacker-computed* physical and cyber power flows, respectively.

We prove that: (i) there exists a linear relationship F between pseudo-boundary injection and power injections in $\mathcal{L}$ under certain circumstances; and (ii) even if $\hat{F}$ does not accurately predict the system response after attack, the attacker can still compute an upper bound on the physical power flow with limited information.

The following assumptions are made about the historical data available to the attacker: (i) the topology for all the historical data remains the same, (ii) each instance of historical data satisfies OPF, and (iii) there exists a subset of buses $\mathcal{Z}$ in $\mathcal{E}$, for which power injections remain constant in the historical data. The subset of remaining buses in $\mathcal{E}$ is denoted as $\mathcal{Y} = \mathcal{E} \setminus \mathcal{Z}$. In our prior work [26, 29, 45, 27], we have shown that congested lines are more vulnerable to line overflow FDI attacks. Analogously, in this section, we assume the target line is congested.

Validation of Multiple Linear Regression Method

In this part, we prove the existence of linear relationship between pseudo-boundary injections and power injections in $\mathcal{L}$ under certain circumstances.

For simplicity, we define the set of lines in the network $\mathcal{G}$ that are the congested for each instance of historical data as $\mathbb{C}$, where $\mathbb{C}^+$ and $\mathbb{C}^-$ are the subsets in $\mathbb{C}$ for which the power flow directions are positive and negative, respectively. We assume there are n_c congested lines in $\mathbb{C}$, $n_{\mathcal{L}}$, $n_{\mathcal{E}}$, and $n_{\mathcal{Y}}$ buses in $\mathcal{L}$, $\mathcal{E}$, and $\mathcal{Y}$, respectively.

In order to evaluate the performance of the coefficient matrix $\hat{F}$, we define a matrix $B = [K_{\mathbb{C}}^{\mathcal{Y}};\ \mathbf{1}^T]$, where $K_{\mathbb{C}}^{\mathcal{Y}}$ is the sub-matrix of K whose rows correspond to

the congested lines in $\mathbb{C}$ and columns correspond to the buses in $\mathcal{Y}$.

Theorem 1. *The coefficient matrix $\hat{F}$ perfectly predicts the pseudo-boundary injections with power injections in $\mathcal{L}$ linearly if and only if B is full column rank.*

Proof. We denote the vector of power injections in $\mathcal{G}$ as v; that is $v = GP_G - P_D$; the vectors $v_{\mathcal{L}}$, $v_{\mathcal{E}}$, $v_{\mathcal{Y}}$, and $v_{\mathcal{Z}}$ represent the subsets of v corresponding to buses in $\mathcal{L}$, $\mathcal{E}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. We define $\mathbb{W}_i$ as the set of lines connecting to boundary bus i, $i \in \mathcal{B}$, where $\mathbb{W}_i^{\mathcal{L}}$ and $\mathbb{W}_i^{\mathcal{E}}$ are the subsets of lines located in $\mathcal{L}$ and $\mathcal{E}$, respectively. We define a vector J_i as the sum of row vectors in K corresponding to lines in $\mathbb{W}_i^{\mathcal{E}}$; that is $J_i = \sum_{k \in \mathbb{W}_i^{\mathcal{E}}} K_k$. The matrices $J_i^{\mathcal{L}}$, $J_i^{\mathcal{Y}}$, and $J_i^{\mathcal{Z}}$ are the submatrices of J in which the columns of the matrices corresponding to buses in $\mathcal{L}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. As introduced in Sec. 4.2.1, the pseudo-boundary injection at bus i is a linear combination of power injections at each bus in $\mathcal{L}$, $\mathcal{Y}$, and $\mathcal{Z}$ is given by

$$\bar{P}_{I,i} = J_i^{\mathcal{L}} v_{\mathcal{L}} + J_i^{\mathcal{Y}} v_{\mathcal{Y}} + J_i^{\mathcal{Z}} v_{\mathcal{Z}}. \tag{4.34}$$

Note that $v_{\mathcal{Z}}$ is a constant across all instances of historical data. Since each instance of historical data resulted from an converged OPF, $v_{\mathcal{L}}$ and $v_{\mathcal{Y}}$ satisfy the following:

$$K_k^{\mathcal{L}} v_{\mathcal{L}} + K_k^{\mathcal{Y}} v_{\mathcal{Y}} = P_{k,\max} - K_k^{\mathcal{Z}} v_{\mathcal{Z}} \qquad \forall k \in \mathbb{C}^+ \tag{4.35}$$

$$K_r^{\mathcal{L}} v_{\mathcal{L}} + K_r^{\mathcal{Y}} v_{\mathcal{Y}} = -P_{r,\max} - K_r^{\mathcal{Z}} v_{\mathcal{Z}} \quad \forall r \in \mathbb{C}^- \tag{4.36}$$

$$\mathbf{1}^T v_{\mathcal{L}} + \mathbf{1}^T v_{\mathcal{Y}} = -\mathbf{1}^T v_{\mathcal{Z}} \tag{4.37}$$

where (4.35) and (4.36) are the thermal limit constraints for congested lines in $\mathbb{C}^+$ and $\mathbb{C}^-$, respectively, and (4.37) is the power balance constraint.

Equations (4.35)−(4.37) can be collected as

$$A v_{\mathcal{L}} + B v_{\mathcal{Y}} = d. \tag{4.38}$$

where $A = [K_{\mathbb{C}}^{\mathcal{L}};\ \mathbf{1}^T]$ and $d = [SP_{\mathbb{C},\max} - K_{\mathbb{C}}^{\mathcal{Z}} v_{\mathcal{Z}};\ -\mathbf{1}^T v_{\mathcal{Z}}]$. The matrix S is a $n_c \times n_c$ diagonal matrix with $S_{kk} = 1$, $\forall k \in \mathbb{C}^+$, and $S_{kk} = -1$, $\forall k \in \mathbb{C}^-$.

The dimensions of A and B are $(n_c + 1) \times n_{\mathcal{L}}$ and $(n_c + 1) \times n_{\mathcal{Y}}$, respectively. Note that the number of columns in B represents the total number of buses in $\mathcal{Y}$.

Suppose that B is full column rank. Thus, $B^T B$ is non-singular; that is, there exists a pseudoinverse $B^+ = \left(B^T B\right)^{-1} B^T$, such that $B^+ B = I$. Therefore, applying B^+ to (4.38), the vector $v_{\mathcal{Y}}$ can be rewritten as

$$v_{\mathcal{Y}} = -B^+ A v_{\mathcal{L}} + B^+ d. \tag{4.39}$$

The pseudo-boundary injection $\bar{P}_{I,i}$ in (4.34) can be written as

$$\bar{P}_{I,i} = \left(J_i^{\mathcal{L}} - J_i^{\mathcal{Y}} B^+ A\right) v_{\mathcal{L}} + J_i^{\mathcal{Y}} B^+ d + J_i^{\mathcal{Z}} v_{\mathcal{Z}}. \tag{4.40}$$

Therefore, the linear coefficient F_i between $\bar{P}_{I,i}$ and $v_{\mathcal{L}}$ is

$$\begin{aligned} F_i &= J_i^{\mathcal{L}} - J_i^{\mathcal{Y}} B^+ A \\ f_{i,0} &= J_i^{\mathcal{Y}} B^+ d + J_i^{\mathcal{Z}} v_{\mathcal{Z}}. \end{aligned} \tag{4.41}$$

From (4.41), we see that F_i is unique and is the perfect linear predictor. The linear coefficient matrix between $\bar{P}_{I,\mathcal{B}}$ and $v_{\mathcal{L}}$ is $F = [F_i]$, $\forall i \in \mathcal{B}$.

Suppose that B is not full column rank. Thus there exist infinitely many of $v_{\mathcal{Y}}$ satisfying (4.38), *i.e.*, $v_{\mathcal{Y}}$ cannot be uniquely determined by $v_{\mathcal{L}}$. Therefore, the multiple linear regression will not perfectly predict the pseudo-boundary injections.

□

In Sec. 4.2.3, we provide two test cases in both the IEEE 24-bus system and Polish system for which the B matrices are full column rank. We demonstrate that for each test case, the $\hat{F}$ obtained with multiple linear regression method does indeed lead to perfect prediction of $\bar{P}_{I,\mathcal{B}}$. We also provide four counter examples (one in the

IEEE 24-bus system, two in the IEEE 118-bus system, and the other in the Polish system). For these illustrated counter-examples, B satisfies $(n_c + 1) < n_{\mathcal{Y}}$, which indicates that B is not full column rank. However, even for a case with B satisfying $(n_c + 1) \geq n_{\mathcal{Y}}$, B cannot be assumed to be full column rank. An example is a system with 3 buses in $\mathcal{Y}$ and 2 parallel congested lines. For this system, $\text{rank}\left(K_{\mathbb{C}}^{\mathcal{Y}}\right) = 1$ since the row vectors in K for the parallel lines are the same. The matrix B, hence, is not a full rank matrix since $\text{rank}\,(B) \leq 2$ and by Theorem 1, $\hat{F}$ cannot result in an accurate prediction.

Note that B does not determine the feasibility of the limited information FDI attacks. In fact, B only determines whether $\bar{P}_{I,\mathcal{B}}$ can be perfectly predicted by $v_{\mathcal{L}}$ or not. However, that does not mean that when B is not full column rank, such attacks are infeasible. The matrix B which is not full column rank may undermine the attacker's evaluation of the attack consequences via the bi-level attack optimization problem. But the attacker can still find attack vector c and design the attack.

Upper Bound on Physical Consequences of Attack

Although $\hat{F}$ in general cannot accurately predict $\bar{P}_{I,\mathcal{B}}$ when B is not full column rank, the attacker can still utilize $\hat{F}$ in the bi-level attack optimization problem (4.25)−(4.33) to predict the physical power flow on target line. However, the attacker-computed physical power flow may not match the physical power flow. The following theorem shows that even so, the attacker can compute an upper bound P_l^{ub} on the physical power flow on the target line subsequent to an attack.

Theorem 2. *The physical power flow on the target line l resulting from attack vector $\bar{c}^*$ is upper bounded by*

$$P_l^{ub} = P_{l,max} - \bar{K}_l \bar{H} \bar{c}^*. \tag{4.42}$$

Proof. Solving the attack optimization problem (4.25)−(4.33), the attacker can obtain the optimal attack vector $\bar{c}^*$. The resulting attack vector for the whole system is c^*, where $c_i^* = \bar{c}_i^*$ for $i \in \mathcal{L}$ and $c_i^* = 0$ for $i \in \mathcal{E}$. Injecting c^* in the system will result in a system re-dispatch determined by (3.5)−(3.8). The difference between the physical and cyber power flows (P_l^p and P_l^c, respectively) on target line l after the post-attack system re-dispatch is

$$P_l^p - P_l^c = -K_l H c^*. \tag{4.43}$$

Thus, the physical power flow on target line l satisfies

$$P_l^p = P_l^c - K_l H c^* \leq P_{l,\max} - K_l H c^*. \tag{4.44}$$

where the upper bound follows from the thermal limit constraint on P_l^c in (3.7). Note that K_l and H are unknown to the attacker with limited information. However, the attacker has the knowledge of $\bar{K}_l$ and $\bar{H}$. We now show that the upper bound in (4.44) is equivalent to P_l^{ub} defined in (4.42).

The PTDF matrices $\bar{K}$ and K satisfy the following

$$\bar{K} = \bar{\Gamma}\bar{H}^+ \tag{4.45}$$

$$K = \Gamma H^+ \tag{4.46}$$

where Γ and $\bar{\Gamma}$ are the dependency matrices between power flow measurements and voltage angle states in $\mathcal{G}$ and $\mathcal{L}$, respectively, H^+ and $\bar{H}^+$ are the pseudoinverse of H and $\bar{H}$, respectively. Note that for target line l, both Γ_l and $\bar{\Gamma}_l$ have only two non-zero elements corresponding to the two end buses of l (denoted as buses l_f and l_t, respectively). In particular, $\Gamma_l^{l_f} = \bar{\Gamma}_l^{l_f} = -\Gamma_l^{l_t} = -\bar{\Gamma}_l^{l_t} = \frac{1}{x_l}$, where x_l is the line impedance of line l. Thus,

$$\bar{K}_l \bar{H} \bar{c}^* = \bar{\Gamma}_l \bar{c}^* = \Gamma_l c^* = K_l H c. \tag{4.47}$$

Therefore, the right-hand side of (4.44) is exactly equal to P_l^{ub}. This proves the upper bound in (4.42). Moreover, P_l^{ub} can be computed by the attacker, since it requires knowledge only of the local network $\mathcal{L}$ and the attack vector $\bar{c}^*$. □

Note that from (4.26) and (4.32), the attacker can compute the difference between the physical and cyber power flows on target line l ($\bar{P}_l^p$ and $\bar{P}_l^c$, respectively) solved with limited information attack optimization as

$$\bar{P}_l^p - \bar{P}_l^c = -\bar{K}_l \bar{H} \bar{c}^*. \tag{4.48}$$

Thus, the difference between physical and cyber power flows seen by the attacker and the system are the same

$$\bar{P}_l^p - \bar{P}_l^c = P_l^p - P_l^c. \tag{4.49}$$

4.2.3 Numerical Results

In this subsection, we illustrate the efficacy of the attacks designed with the method proposed in Sec. 4.2.1. To this end, we first compute the coefficient matrix with historical data using the multiple linear regression method. Subsequently, we solve the optimization problem to find the optimal attack vector $\bar{c}^*$ inside $\mathcal{L}$. Finally, we test the physical consequences of the attack vector $\bar{c}^*$ on the entire network $\mathcal{G}$. The test systems include the IEEE 24-bus reliability test system (RTS), IEEE 118-bus system, and Polish system from MATPOWER v4.1. In particular, the line rating data for IEEE 118-bus system is adopted from [49]. The whole network DC OPF and limited information attack algorithm is implemented with Matlab. The optimization problem is solved with CPLEX.

To model realistic power systems, we assume that there are congested lines prior to the attack and the attacker chooses one of them in $\mathcal{L}$ as the target to maximize

Table 4.2: Summary of The Test Systems

Test System	Scenario 1			Scenario 2		
	# of Congested Lines (n_c)	# of Buses in $\mathcal{E}_M$ ($n_{\mathcal{E}_M}$)	rank(B)	# of Congested Lines (n_c)	# of Buses in $\mathcal{E}$ ($n_{\mathcal{E}}$)	rank(B)
24-bus	2	2	2	2	15	3
118-bus	3	5	4	3	71	4
Polish	8	6	6	8	2357	10

power flow. This is achieved in simulation by uniformly reducing all line ratings by 50% for the IEEE 24-bus RTS and 45% for the IEEE 118-bus system.

We illustrate our results for the following choice of parameters: the weight of the l_1-norm of attack vector in (4.25), ζ, is set to 1% of the original power flow value of the target line; and the load shift factor in (4.28), τ, is set to 10%. We assume that the attacker can obtain 200 instances of historical data inside $\mathcal{L}$.

We focus on two scenarios for the historical data:

- **Scenario 1 - Constant Loads in $\mathcal{E}$:** In each instance of data, loads in $\mathcal{E}$ remain unchanged while loads in $\mathcal{L}$ vary as a percent p of the base load, where p is independent $\mathcal{N}(0, 10\%)$. That is, power injections vary only at buses with marginal generators (denoted $\mathcal{E}_M$). Therefore, in this scenario, the set $\mathcal{Y}$ is given by $\mathcal{Y} = \mathcal{E}_M$. The number of buses in $\mathcal{E}_M$ is denoted by $n_{\mathcal{E}_M}$

- **Scenario 2 - Varying Loads in $\mathcal{G}$:** In each instance of data, loads in both $\mathcal{L}$ and $\mathcal{E}$ vary as a percent p of the base load, with p chosen independently for each load as $\mathcal{N}(0, 10\%)$. In this scenario, power injections at all buses in $\mathcal{E}$ vary in the historical data, *i.e.,* $\mathcal{Y} = \mathcal{E}$.

Note that the data in both scenarios also satisfy the following assumptions: (i) the topology for all the historical data remains the same, (ii) the historical generation dispatches data in both scenarios satisfies OPF, and (iii) there exists a subset of buses $\mathcal{Z}$ in $\mathcal{E}$, for which power injections remain constant in the historical data. An example of attack is that an attacker hacks into the system and collects the data from 12:00 p.m. to 2:00 p.m. for the entire month of July and then launches a FDI attack at the end of the month.

Results for IEEE 24-bus RTS system

In this part, we present attack consequences on the IEEE 24-bus RTS system for Scenarios 1 and 2. The sub-network $\mathcal{L}$ is illustrated in Fig. 4.3. In each scenario, we compare the attack consequences on target line 28 determined by the optimization problems for two cases: (i) complete system knowledge as in [26] (identified as *global case*), and (ii) limited system knowledge (henceforth identified as *local case*). For local case, we compare the physical power flow P_l^p and the attacker-computed physical power flow $\bar{P}_l^p$. The results of attacks are illustrated in Fig. 4.5. We illustrate the difference between the physical and the attacker-computed pseudo-boundary injections in Fig. 4.6.

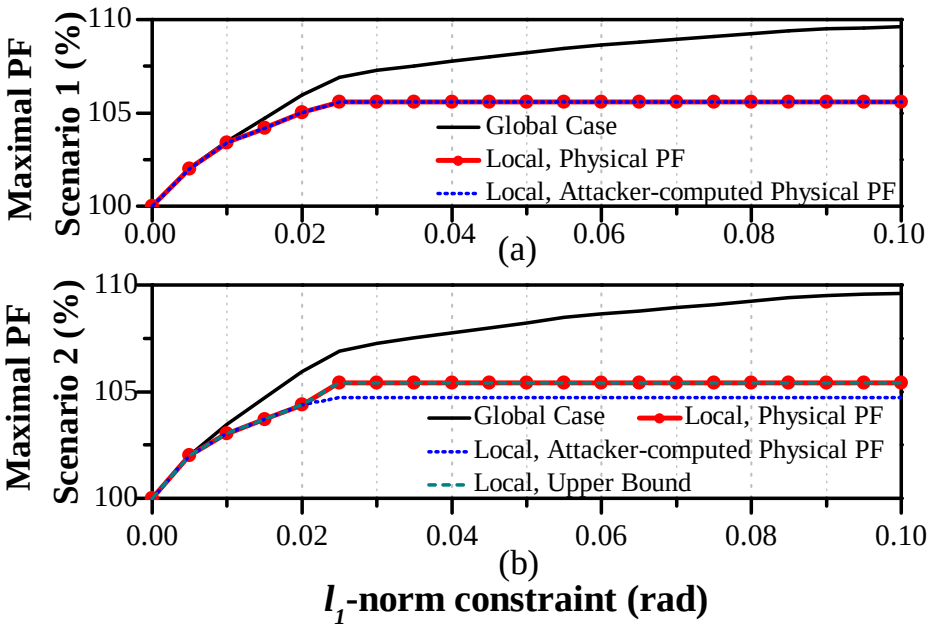


Figure 4.5: The Maximum Power Flow (PF) V.S. The l_1-Norm Constraint (N1) When Target Line Is 28 of IEEE 24-Bus System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

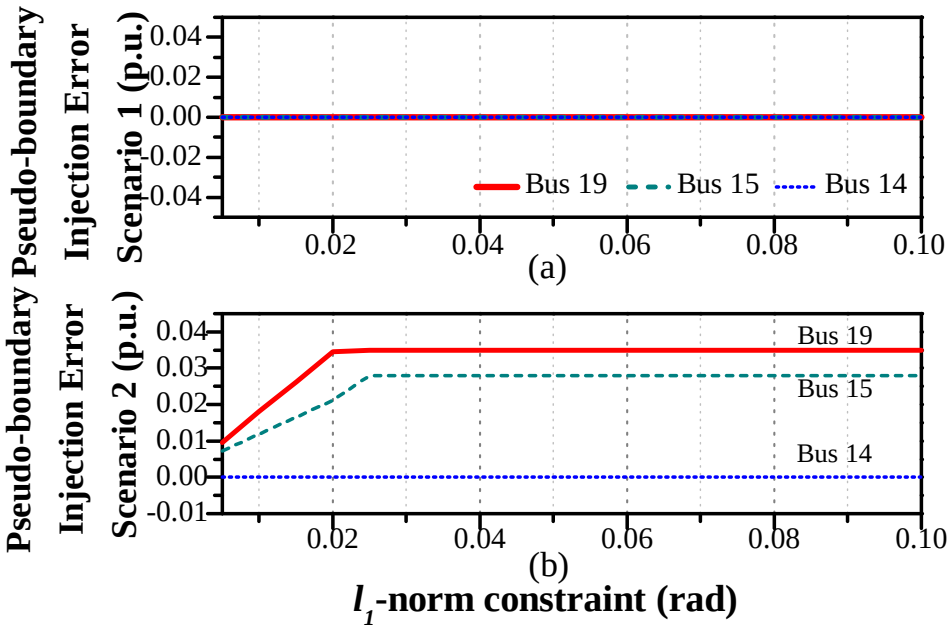


Figure 4.6: The Pseudo-Boundary Power Injection Error V.S. The l_1-Norm Constraint (N1) When Target Line Is 28 of IEEE 24-Bus System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

In Figs. 4.5(a) and (b), we note that the solutions for the local case is sub-optimal relative to that for the global case. The reason is that as N_1 is relaxed, getting a larger overflow on the target line requires measurements in both $\mathcal{L}$ and $\mathcal{E}$ to be modified. Therefore, the constraint on limited attack resources prevents any further increase in the maximal target line flow for the local case.

The parameters of the test system are summarized in Table 4.2. The historical data in Scenario 1 satisfies $\text{rank}(B) = n_{\mathcal{E}_M}$. Thus, by Theorem 1, the pseudo-boundary power injections are perfectly predicted by the multiple linear regression method, which explains why the attacker-computed system response post-attack is the same as the actual response, as illustrated in Figs. 4.5(a) and 4.6(a).

Furthermore, Table 4.2 shows that for the historical data in Scenario 2, $\text{rank}(B) < n_{\mathcal{E}}$. Thus, by Theorem 1, the predictions of pseudo-boundary injections by the multiple linear regression are not accurate and there will be mismatches between the

actual and the attacker-computed system response post-attack. This is verified by the non-zero pseudo-boundary power injection differences shown in Fig. 4.6(b). In Fig. 4.5(b), in addition to plotting the attacker-computed physical power flow, we also plot the upper bound on physical power flow. From Fig. 4.5(b), we observe that although there are mismatches between the actual and attacker-computed system response under attack, the upper bound found in Sec. 4.2.2 exactly matches the physical power flow.

Results for IEEE 118-bus System

In this part, we test the consequences of attacks on the IEEE 118-bus system. The details of sub-network $\mathcal{L}$ are listed in Table 4.3. The results of attacks designed with historical data in Scenarios 1 and 2 are illustrated in Fig. 4.7 with sub-plots (a) and (b), respectively. The difference between the physical and the attacker-computed pseudo-boundary injections at 3 of 18 boundary buses, buses 23, 70, and 80, for both scenarios are illustrated in Fig. 4.8.

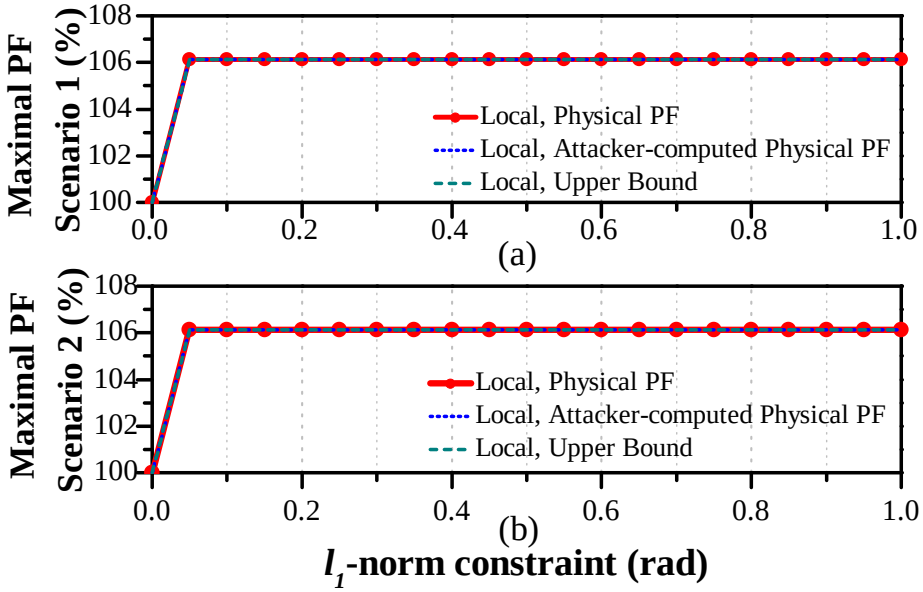


Figure 4.7: The Maximum Power Flow (PF) V.S. The l_1-Norm Constraint (N1) When Target Line Is 5 of IEEE 118-Bus System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

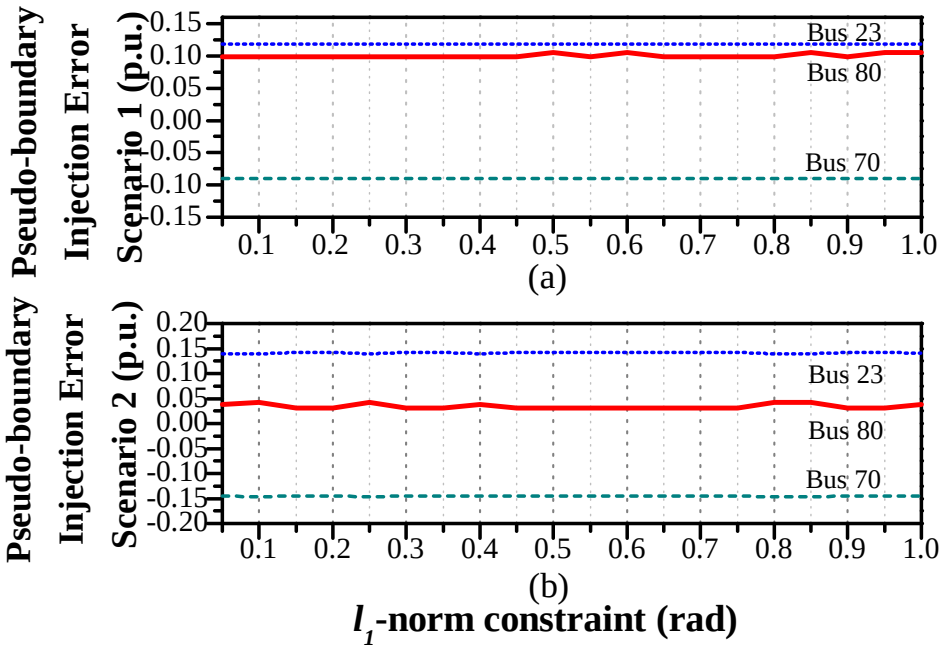


Figure 4.8: The Pseudo-Boundary Power Injection Error V.S. The l_1-Norm Constraint (N1) When Target Line Is 5 of IEEE 118-Bus System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

The parameters of the test system are also summarized in Table 4.2. Note that for historical data in both scenarios, B does not have full column rank. Therefore, Theorem 1 predicts a mismatch between physical and attacker-computed pseudo-boundary injections. This is verified by Fig. 4.8, which shows the pseudo-boundary injection error. In Figs. 4.7(a) and (b), we find that in both scenarios, both the attacker-computed physical power flow and the upper bound match the physical power flow. This case demonstrates that even though there are mismatches between physical and attacker-computed pseudo-boundary injections, the attacker-computed physical power flow can still be correct. Note that, in this case, both the cyber power flow and the attacker-computed cyber power flow reach the limit post-attack since the target line is congested before attack. Therefore, from (4.49), the attacker-computed physical power flow is the same as the physical power flow.

Table 4.3: Summary of The Attack Sub-network in IEEE 118-Bus System

Buses	1-14, 16, 17, 23, 25-27, 30, 33-35, 37-40, 47, 49, 59-66, 68-70, 75, 77, 80, 81, 116, 117
Lines	1-17, 20, 22, 31-33, 36-38, 47, 48, 50-55, 65, 88-100, 102, 104-108, 115, 116, 119, 120, 123, 124, 126, 127, 183, 184
Boundary Buses	13, 14, 17, 23, 27, 33-35, 40, 47, 49, 59, 62, 66, 70, 75, 77, 80

Results for Polish System

In this part, we test the consequences of attacks on the Polish system. The topology of the Polish system is illustrated in Fig. 4.9 where the sub-network $\mathcal{L}$ is highlighted with orange and the target line 1816 is highlighted with red. The results of attacks are demonstrated in Fig. 4.10. The difference between the physical and the attacker-computed pseudo-boundary injections at 3 of 7 boundary buses, buses 919, 1055, and 1215 for both scenarios are illustrated in Fig. 4.11.

The parameters of the test system are also summarized in Table. 4.2. In this test system, the historical data in Scenario 1 satisfies $\text{rank}(B) = n_{\mathcal{E}_M}$. Thus, as stated in Theorem 1, the attacker can perfectly predict the pseudo-boundary injections with the multiple linear regression method. This can be verified by Figs. 4.10(a) and 4.11(a) which show the perfectly matched physical and attacker-computed power flow and pseudo-boundary injections, respectively. The matrix B for historical data in Scenario 2, on the other hand, does not have full column rank. Therefore, a mismatch between physical and attacker-computed pseudo-boundary injections is predicted by Theorem 1. This is verified by the non-zero pseudo-boundary injection errors shown in Fig. 4.11(b). In Fig. 4.10(b), we also find that in Scenario 2, both the attacker-computed

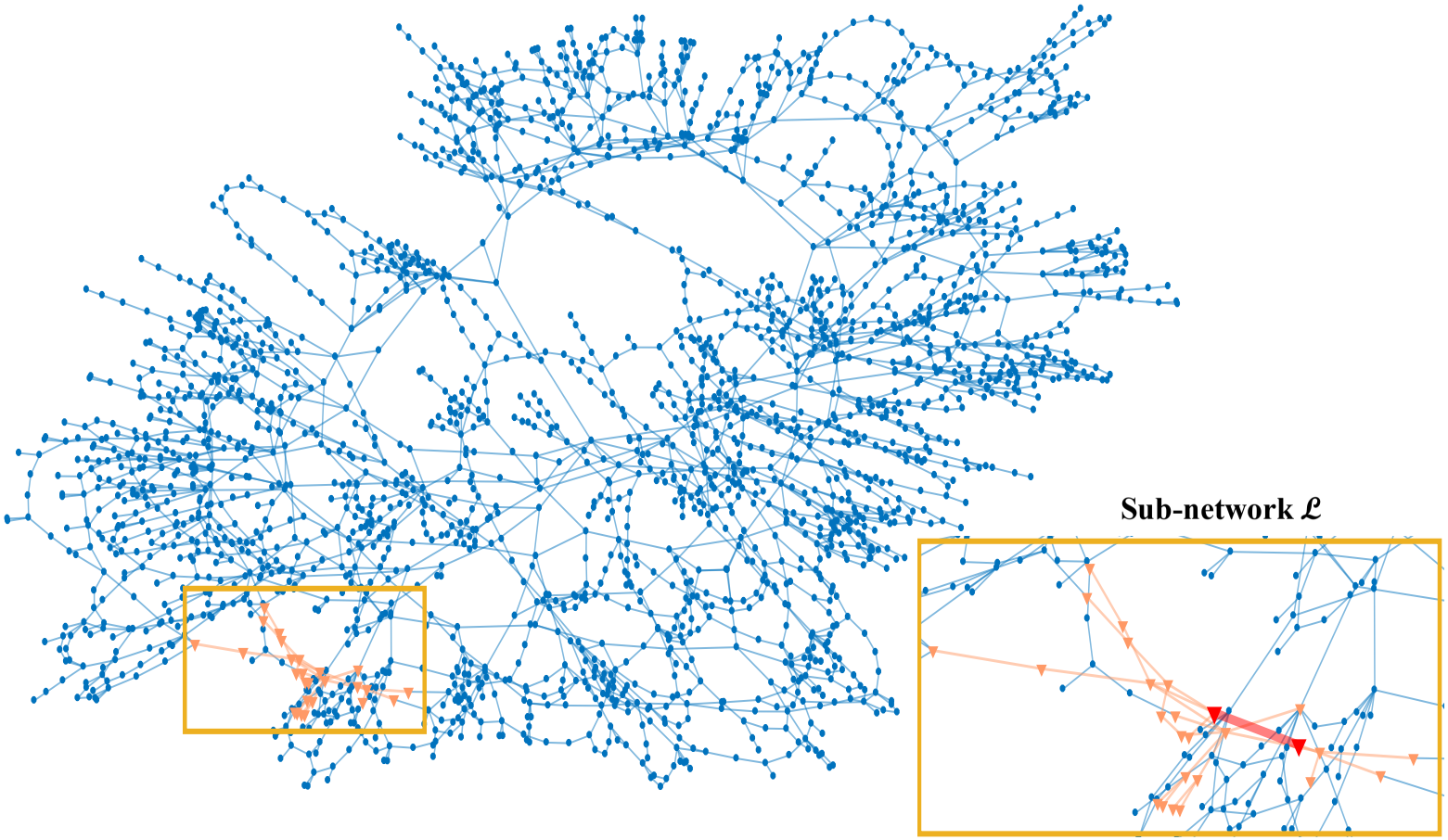


Figure 4.9: Polish System Decomposed into Attack Sub-Network and Attack External Network.

physical power flow and the upper bound match the physical power flow. This case further verifies our observation in the IEEE 118-bus system that even with prediction errors on pseudo-boundary injections, the attacker-computed physical power flow can still be correct when both the cyber and the attacker-computed cyber power flow reach the limit post-attack.

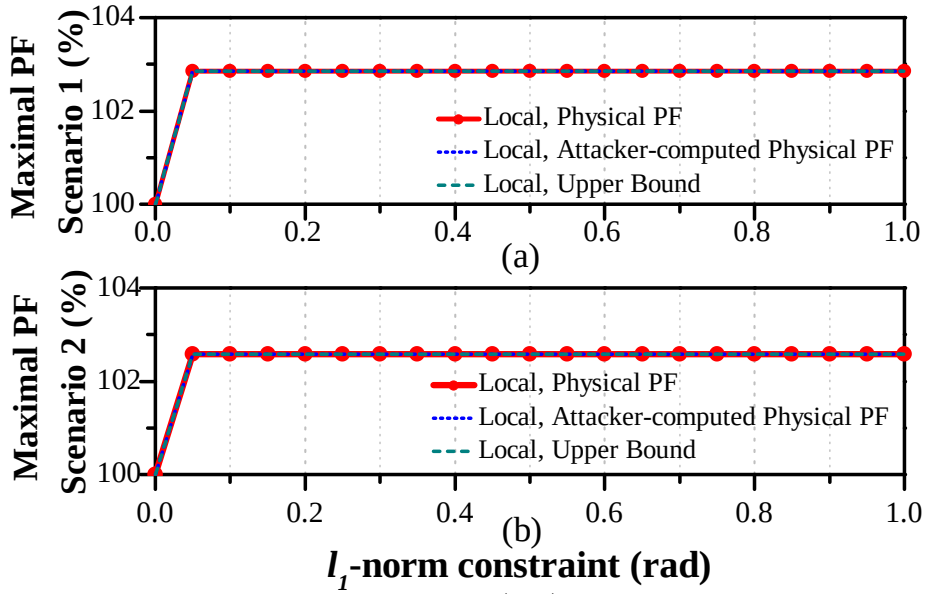


Figure 4.10: The Maximum Power Flow (PF) V.S. The l_1-Norm Constraint (N1) When Target Line Is 1816 of Polish System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

The attack sub-network studied here as well as in the IEEE 24-bus RTS and IEEE 118-bus systems are compared in Table. 4.4. From this table, it can be seen that compared to the attack sub-networks in the IEEE 24-bus and 118-bus systems, the size of the attack sub-network $\mathcal{L}$ studied here only covers 1% of the Polish system. This test case shows that even with a minuscule amount of system information, the attacker can still cause physical line overflow in a large-scale system.

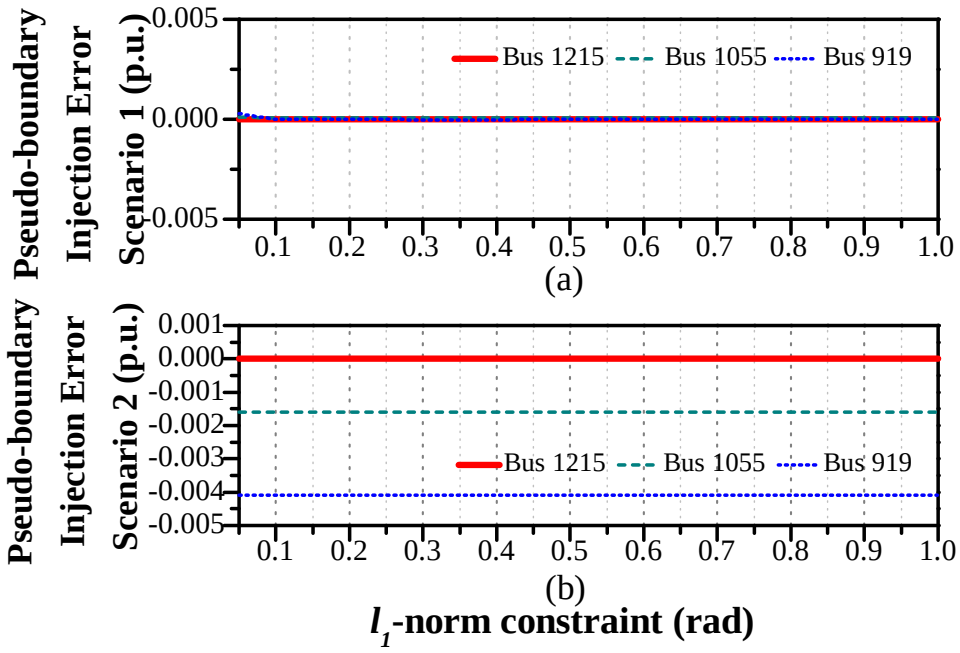


Figure 4.11: The Pseudo-Boundary Power Injection Error V.S. The l_1-Norm Constraint (N1) When Target Line Is 1816 of Polish System for (a) Scenario 1, and (b) Scenario 2 Historical Data.

Table 4.4: Comparison of the Attack Sub-networks in IEEE 24-Bus RTS, IEEE 118-Bus, and Polish Systems.

Test System	# of Buses	% of the Total Buses	# of Branches	% of the Total Branches	# of Generators	% of the Total Generators
24-bus	8	37.5%	12	31.58%	16	48.48%
118-bus	47	39.83%	63	33.87%	22	40.74%
Polish	26	1.09%	27	0.93%	5	1.53%

4.3 Sensitivity Analysis

In Sec. 4.2.3, several historical data assumptions have been made to ensure both the pre-attack and post-attack operation conditions have the similar characteristics

with each instance of historical data. Such characteristics include the congestion patterns, topologies, load distribution patterns, and marginal generation sets. In addition, linear (DC) power flow model is utilized for learning and attack design in Sec. 4.2. In this section, we denote such an instance of historical data as *perfect historical data*. We denote the historical dataset that consists of only perfect historical data as *perfect historical dataset*. Throughout, the perfect historical data for each test system satisfies the following assumptions: (i) in each instance of data, the system topology, the set of operating generators, and the cost of each generator are the same with the pre-attack operation condition, (ii) in each instance of data, the loads in $\mathcal{G}$ vary as a percent $p \sim \mathcal{N}(0, 10\%)$ of the base load, and (iii) the historical generation dispatches data satisfies OPF. Our simulation results in Sec. 4.2 have demonstrated that limited information attacks designed with perfect historical dataset can result in physical line overflow successfully. However, in reality, the attacker may obtain imperfect historical datasets which do not satisfy the perfect historical dataset definition. As a result, consequences of the attack may be reduced.

In this section, we perform sensitivity analysis of the attacks designed with the method proposed in Sec. 4.2.1 on multiple scenarios of imperfect historical datasets. The test systems include the IEEE 24-bus, 118-bus, and 2383-bus (Polish) systems. The following parameters are chosen to illustrate the results: the weight of the l_1-norm of attack vector in (4.25), ζ, is set to 1% of the original popwer flow value of the target line, the load shift factor in (4.28), τ, is set to 10%, and the l_1-norm constraint limit, N_1, are set to 0.05, 0.4, and 0.5 for the IEEE 24-bus, 118-bus, and Polish systems, respectively. We assume that the attacker can obtain 200 instances of historical data inside $\mathcal{L}$ for each of the three test systems. In particular, we only consider historical datasets with varying loads in the entire network $\mathcal{G}$ so as to demonstrate the worst-case performance.

We focus on the following four classes of imperfect historical datasets:

1. Dataset has different load varying percentage with the desired load shift τ (denoted as *load varying error dataset*).

2. Dataset includes topologies different from the target operation condition (denoted as *topology error dataset*).

3. Dataset has different generation dispatch plans (*e.g.,* some generators are shut down, or the outputs of some generators have been manually changed) (denoted as *dispatch error dataset*).

4. Dataset satisfies non-linear (AC) power flow model instead of DC power flow model (denoted as *AC power flow dataset*).

In the following subsections, we exhaustively test the performance of the linear regression method obtained with the 4 classes of imperfect historical datasets. To evaluate the impacts of different factors to the prediction accuracy and attack consequences, we ensure that the historical data in each subsection only has one of the above four mentioned errors. The other factors remain the same as the perfect historical dataset. For each choice of imperfect parameters, we randomly generate a corresponding imperfect historical dataset, compute the coefficient matrix with this dataset, solve the optimization problem to find the optimal attack, and test the physical consequences of the attack. We repeat this process 100 times and demonstrate the following statistical results: (i) the percentage of trials in which the target line has physical line overflow; (ii) the percentage of trials in which the target line physical power flow (PPF) equals to the attacker-computed physical power flow (CPF); (iii) the maximum, minimum, and median of the target line physical power flow values; (iv) the maximum positive and negative differences between the target line physical power

flows and attacker-computed physical power flows; and (v) the percentage of trials in which the pseudo-boundary injection prediction error decreased. Note that for statistical result (v), we use the error between the physical and attacker-computed pseudo-boundary injections to demonstrate the pseudo-boundary injection prediction accuracy. In particular, since there are multiple boundary buses in each test system, we compute the l_2-norm of the errors on all boundary buses and compare it with the average l_2-norm of the errors obtained from 100 perfect historical datasets. In the following subsections, we use *% of trails with target PPF overflow*, *% of trials with target PPF matching CPF*, *statistic results of the target PPF*, *max +/- target line PF difference*, and *% of trials with prediction error increasing* to denote statistical results (i)–(v) for short, respectively.

4.3.1 Load Varying Error Dataset

In this subsection, the impact of the historical data with load varying errors on the attack consequences has been studied. Compared to the perfect historical dataset, the imperfect historical dataset includes $n_o\%$ of imperfect historical data in which the loads in $\mathcal{G}$ vary as a percent $p \sim \mathcal{N}(0, \bar{\tau})$ of the base load, where $\bar{\tau} \neq 10\%$. All possible choices of $n_o\%$ imperfect historical data ranging between 20% to 100% with 20% increments and the standard deviation $\bar{\tau}$ of the load varying percent p ranging between 20% to 70% with 10% increments have been exhaustively studied to demonstrate the effect of the load varying errors.

Results for IEEE 24-bus RTS System

In this part, we demonstrate the sensitivity analysis results of the attacks in the IEEE 24-bus RTS system. The statistical results (i)–(v) are summarized in Tables. 4.5–4.9.

Table 4.5: Summary of The % of Trials with Target PPF Overflow for Load Varying Error Dataset in the IEEE 24-Bus System (Statistics (i)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	100%	100%	99%	96%	92%	93%
40%	100%	99%	96%	92%	93%	92%
60%	99%	99%	95%	91%	90%	91%
80%	100%	99%	90%	93%	93%	91%
100%	100%	99%	91%	92%	86%	87%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.6: Summary of The % of Trials with Target PPF Matching CPF for Load Varying Error Dataset in The IEEE 24-Bus System (Statistics (ii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	0	2%	5%	11%	19%	15%
40%	0	6%	13%	10%	13%	12%
60%	1%	4%	12%	17%	20%	20%
80%	0	3%	16%	12%	14%	21%
100%	2%	8%	17%	15%	19%	20%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.7: Summary of The Statistic Reuslts of The Target PPF for Load Varying Error Dataset in The IEEE 24-Bus System (Statistics (iii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
Statistics	20%	30%	40%	50%	60%	70%
Max	105.54%	105.50%	105.58%	105.55%	105.57%	105.59%
Min	102.04%	96.63%	95.10%	94.62%	94.43%	94.13%
Median	104.20%	104.20%	104.20%	104.20%	104.20%	104.20%

Table 4.8: Summary of The Max +/- Target Line PF Error for Load Varying Error Dataset in The IEEE 24-Bus System (Statistics (iv)).

Max Target PF Error	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
	20%	30%	40%	50%	60%	70%
+	5.96%	8.08%	9.42%	13.29%	14.6%	19.3%
−	1.99%	1.99%	2.63%	6.5%	6.4%	7.8%

Table 4.9: Summary of The % of Trials with Prediction Error Increasing for Load Varying Error Dataset in The IEEE 24-Bus System (Statistics (v)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	91%	97%	93%	83%	86%	88%
40%	99%	98%	88%	93%	94%	89%
60%	99%	93%	95%	94%	95%	96%
80%	100%	95%	97%	99%	100%	100%
100%	100%	97%	99%	100%	99%	100%

n_o%: Percentage of imperfect data in the historical dataset.

Results for IEEE 118-bus System

In this part, we demonstrate the sensitivity analysis results of the attacks in the IEEE 118 system. The statistical results (i)–(v) are summarized in Tables. 4.10–4.14.

Table 4.10: Summary of The % of Trials with Target PPF Overflow for Load Varying Error Dataset in The IEEE 118-Bus System (Statistics (i)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	100%	100%	100%	100%	100%	99%
40%	100%	100%	100%	100%	100%	100%
60%	98%	100%	100%	100%	100%	100%
80%	100%	100%	100%	100%	100%	100%
100%	100%	100%	100%	100%	100%	100%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.11: Summary of The % of Trials with Target PPF Matching CPF for Load Varying Error Dataset in The IEEE 118-Bus System (Statistics (ii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	83%	90%	98%	99%	100%	98%
40%	80%	97%	99%	100%	100%	100%
60%	71%	92%	99%	100%	100%	100%
80%	68%	89%	99%	100%	98%	98%
100%	59%	91%	95%	98%	100%	100%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.12: Summary of The Statistic Reuslts of The Target PPF for Load Varying Error Dataset in The IEEE 118-Bus System (Statistics (iii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
Statistics	20%	30%	40%	50%	60%	70%
Max	106.13%	106.13%	106.13%	106.13%	106.13%	106.13%
Min	97.98%	105.52%	104.01%	103.29%	102.97%	100.05%
Median	106.13%	106.13%	106.13%	106.13%	106.13%	106.13%

Table 4.13: Summary of The Max +/- Target Line PF Error for Load Varying Error Dataset in The IEEE 118-Bus System (Statistics (iv)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
Max Target PF Error	20%	30%	40%	50%	60%	70%
+	2.60%	8.05%	4.77%	3.96%	1.00%	1.00%
−	0	0	0	0	0	0

Table 4.14: Summary of The % of Trials with Prediction Error Increasing for Load Varying Error Dataset in IEEE The 118-Bus System (Statistics (v)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	27%	26%	43%	28%	34%	44%
40%	35%	34%	50%	57%	47%	42%
60%	41%	46%	57%	53%	52%	35%
80%	59%	50%	53%	56%	38%	28%
100%	41%	69%	40%	34%	22%	14%

n_o%: Percentage of imperfect data in the historical dataset.

Results for the Polish System

In this part, we demonstrate the sensitivity analysis results of the attacks in the Polish system. The statistical results (i)–(v) are summarized in Tables. 4.15–4.19.

Table 4.15: Summary of The % of Trials with Target PPF Overflow for Load Varying Error Dataset in The Polish System (Statistics (i)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	100%	100%	100%	100%	100%	99%
40%	99%	99%	100%	100%	99%	98%
60%	100%	100%	99%	99%	99%	100%
80%	100%	100%	96%	99%	100%	98%
100%	100%	97%	99%	99%	100%	95%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.16: Summary of The % of Trials with Target PPF Matching CPF for Load Varying Error Dataset in The Polish System (Statistics (ii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	75%	97%	97%	94%	94%	97%
40%	83%	89%	96%	91%	94%	94%
60%	87%	89%	84%	69%	67%	71%
80%	83%	77%	59%	38%	33%	35%
100%	75%	56%	31%	7%	4%	7%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.17: Summary of The Statistic Reuslts of The Target PPF for Load Varying Error Dataset in The Polish System (Statistics (iii)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
Statistics	20%	30%	40%	50%	60%	70%
Max	102.86%	102.86%	102.85%	102.85%	102.85%	102.85%
Min	99.31%	98.80%	97.10%	98.25%	99.85%	99.07%
Median	102.56%	102.52%	102.50%	102.47%	102.45%	102.43%

Table 4.18: Summary of The Max +/- Target Line PF Error for Load Varying Error Dataset in The Polish System (Statistics (iv)).

Max Target PF Error	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
	20%	30%	40%	50%	60%	70%
+	1.79%	3.53%	8.36%	3.24%	2.84%	2.94%
−	0.31%	1.58%	5.67%	0.31%	0.50%	2.78%

Table 4.19: Summary of The % of Trials with Prediction Error Increasing for Load Varying Error Dataset in The Polish System (Statistics (v)).

	Standard Deviation $\bar{\tau}$ of Load Varying Percentage $p \sim \mathcal{N}(0, \bar{\tau})$					
n_o%	20%	30%	40%	50%	60%	70%
20%	59%	56%	72%	76%	77%	66%
40%	54%	63%	64%	68%	73%	74%
60%	56%	54%	67%	74%	78%	86%
80%	53%	59%	80%	87%	91%	90%
100%	58%	74%	83%	93%	95%	93%

n_o%: Percentage of imperfect data in the historical dataset.

The results in this subsection demonstrate that with load varying errors in historical data, the prediction errors on the pseudo-boundary injections will increase. For most of the test cases, the pseudo-boundary prediction errors will increase as the number of the imperfect data increase. The prediction errors seem to be irrelevant with the load varying factors. Table. 4.16 shows that for the Polish system, the prediction errors on the target line physical power flow also increase as the number of imperfect data increases. However, the attacker can still result in line overflow with such imperfect historical datasets. This can be verified with the data shown in Tables. 4.5, 4.10, and 4.15, for the IEEE 24-bus, IEEE 118-bus, and Polish systems, respectively, where physical line overflow occurs in over 95% of the test cases.

4.3.2 Topology Error Dataset

In this subsection, the impact of the historical data with topology errors on the attack consequences has been studied. We assume that the attacker is not aware of line outages in $\mathcal{E}$ when collecting historical data. We have studied two types of imperfect historical datasets as follows:

a. All instances of data in the imperfect historical dataset have uniform topology with a single line outage in $\mathcal{E}$.

b. The imperfect historical dataset consists of data under two different topologies. Each topology has a single line outage in $\mathcal{E}$.

The results of the above two types of imperfect historical datasets are demonstrated in Tables 4.20.

From this table, it can observed that if the attacker uses historical datasets with topology errors to design attacks, the pseudo-boundary prediction errors will increase. However, in over 90% of the test cases, the attacker can still cause line overflow with

Table 4.20: Summary of The Sensitivity Analysis Results for Topology Error Dataset.

	Type	# of Topologies	# of Total Trials	% of Trials with Target PPF Overflow	% of Trails with Target PPF Matching CPF	Statistic Reuslts of the Target PPF			Max Target Line PF Error		% of Trials with Prediction Error Increasing
						Max	Min	Med	+	−	
IEEE 24-bus	a	16	1600	100%	0.13%	105.42%	103.03%	104.2%	22.92%	1.99%	95.81%
	b	50	5000	100%	0.36%	104.22%	100.92%	104.2%	6.56%	1.99%	97.26%
IEEE 118-bus	a	118	11800	96%	94.75%	106.13%	96.34%	106.13%	1.59%	0	76.14%
	b	100	10000	95.90%	94.69%	106.13%	96.15%	106.13%	0.67%	0	77.51%
Polish	a	154	15400	96.68%	77.76%	102.85%	100.18%	102.81%	1.16%	0	95.16%
	b	100	10000	99.92%	79.12%	102.86%	99.61%	102.79%	2.31%	2.54%	94.45%

the inaccurate coefficient matrices. Specifically, in the IEEE 118-bus and Polish systems, the physical power flow and attacker-computed physical power flow matches in near 95% and 80% test cases, respectively.

4.3.3 Dispatch Error Dataset

In this subsection, we demonstrate the implication of the historical data with generation dispatch errors to the attack consequences. Specifically, we consider the following two types of dispatch errors:

1. Compared to the set of generators in the perfect historical data, all instances of data in the imperfect historical dataset have one generator outage in $\mathcal{E}$.

2. The generation dispatch in the imperfect historical dataset does not satisfy OPF result. This type of error is to mimic the manual generation dispatches in the historical dataset.

Generator Outage Error

In this part, we assume the attacker is not aware of the generator outages in $\mathcal{E}$ when collecting historical data. We have exhaustively tested the attack consequences resulting from the imperfect historical datasets with

a. All possible choices of a single generator outage in $\mathcal{E}$.

b. The imperfect historical dataset consists of data under two different generator outage patterns. Each pattern has a single generator outage in $\mathcal{E}$.

The results of the above two types of imperfect historical datasets are demonstrated in Tables 4.21.

Table 4.21: Summary of The Sensitivity Analysis Results for Generator Outage Error Dataset.

	Type	# of Generator Outage Patterns	# of Total Trials	% of Trials with Target PPF Overflow	% of Trails with Target PPF Matching CPF	Statistic Reuslts of the Target PPF			Max Target Line PF Error		% of Trials with Prediction Error Increasing
						Max	Min	Med	+	−	
IEEE 24-bus	a	17	1700	100%	0	104.2%	104.2%	104.2%	1.99%	0	100%
	b	50	5000	100%	0	104.2%	104.2%	104.2%	1.99%	0	100%
IEEE 118-bus	a	32	3200	96.88%	95.87%	106.13%	96.53%	106.13%	0.65%	0	79.5%
	b	50	5000	97.4%	96.34%	106.13%	99.35%	106.13%	0.8%	0	80.26%
Polish	a	322	32200	99.81%	74.14%	102.85%	97.6%	102.79%	0.98%	3.52%	94.42%
	b	50	5000	98.3%	81.4%	102.86%	99.98%	102.82%	0.42%	0	93.18%

The last column in this table demonstrates that prediction errors on the pseudo-boundary injections increase when historical datasets with generator outage errors are utilized to design attacks. However, in each test system, there are over 95% of the designed attacks resulting in physical line overflow. In particular, the physical consequences can be accurately predicted with near 97% and 82% inaccurate coefficient matrices in the IEEE 118-bus and Polish systems, respectively.

Manual Dispatch Error

In this part, we assume the some instances of the historical data do not satisfy the OPF results. This type of errors is to mimic the manual generation dispatches under some conditions (*e.g.,* to eliminate emergences). To simulate such errors, we randomly assign the generation cost to each generator in the test system and run OPF to obtain the results. Such results ensure that there are no violations in the system while the dispatches are totally different from the perfect historical dispatches. Compared to the perfect historical dataset, the imperfect historical dataset includes $n_o\%$ of imperfect historical data. We have studied the imperfect historical dataset with 20% to 100% imperfect data with 20% increments. The statistical results (i)–(v) are summarized in Tables. 4.22–4.26, respectively.

Table 4.22: Summary of The % of Trials with Target PPF Overflow for Manual Dispatch Error Dataset (Statistics (i)).

	Test Systems		
n_o%	24-bus	118-bus	Polish
20%	89%	98%	100%
40%	92%	93%	96%
60%	90%	98%	96%
80%	95%	96%	100%
100%	80%	96%	99%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.23: Summary of The % of Trials with Target PPF Matching CPF for Manual Dispatch Error Dataset (Statistics (ii)).

	Test Systems		
n_o%	24-bus	118-bus	Polish
20%	20%	97%	35%
40%	11%	93%	40%
60%	22%	98%	42%
80%	21%	96%	31%
100%	15%	96%	26%

n_o%: Percentage of imperfect data in the historical dataset.

Table 4.24: Summary of The Statistic Reuslts of The Target PPF for Manual Dispatch Error Dataset (Statistics (iii)).

	Test Systems		
Statistics	24-bus	118-bus	Polish
Max	105.64%	106.13%	102.85%
Min	95.00%	100.66%	101.76%
Median	105.64%	106.13%	102.76%

Table 4.25: Summary of The Max +/- Target Line PF Error for Manual Dispatch Error Dataset (Statistics (iv)).

Max Target	Test Systems		
PF Error	24-bus	118-bus	Polish
+	10.81%	11.42%	0.98%
−	6.5%	0	0

Table 4.26: Summary of The % of Trials with Prediction Error Increasing for Manual Dispatch Error Dataset (Statistics (v)).

	Test Systems		
n_o%	24-bus	118-bus	Polish
20%	92.59%	44%	54%
40%	95.65%	39%	52%
60%	100%	29%	53%
80%	100%	26%	54%
100%	100%	41%	49%

n_o%: Percentage of imperfect data in the historical dataset.

From Tables. 4.23 and 4.26, it can be seen that as the number of imperfect data with manual dispatch error increases, the attacker's prediction on both the target line physical power flow and the pseudo-boundary injections will be undermined. However, Table. 4.22 indicates that the attacker can still cause physical line overflow on the target line in most of the test trials.

4.3.4 AC Power Flow Dataset

In this subsection, the impact of the historical data with AC power flow on the attack consequences has been studied. We tested the performances of coefficient matrices learned form the historical dataset with load varying percent as $p \sim \mathcal{N}(0, 10\%)$, same topology, same on-line generators and costs with the pre-attack and post-attack operation conditions, but under AC power flow model. The statistical results (i)–(v) for all tested cases are summarized in Tables. 4.27.

From this table, it can be seen that historical datasets with non-linear power flow data can reduce the prediction accuracy of the pseudo-boundary injections. However, for all the test systems, 100% of the designed attacks can result in physical target line overflows.

Overall, the sensitivity analysis results in this section demonstrate the robustness of the limited information attack strategy introduced in Sec. 4.2.1 on imperfect datasets with load varying errors, topology errors, generation dispatch errors, and non-linear power flow data.

4.4 Discussion

In this chapter, several approximations are made to model ideal attackers and systems. In this section, the impacts of such approximations on the attack consequences in a realistic system are analyzed.

Table 4.27: Summary of The Sensitivity Analysis Results for AC Power Flow Dataset.

Test System	% of Trials with Target PPF Overflow	% of Trails with Target PPF Matching CPF	Statistic Reuslts of the Target PPF			Max Target Line PF Error		% of Trials with Prediction Error Increasing
			Max	Min	Med	+	−	
IEEE 24-bus	100%	30%	105.64%	104.2%	105.24%	1.54%	7.88%	23%
IEEE 118-bus	100%	100%	106.13%	106.13%	106.13%	0	0	100%
Polish	100%	100%	102.85%	102.85%	102.81%	0	0	95%

4.4.1 Approximation on Power Flow Limit Constraints

Throughout this chapter, the attack is defined to be successful if it can result in the physical active power flow on the target line exceeding the power flow limit. Note that only thermal limits are considered here, *i.e.,* $P_{\max}$ in (4.2) and (4.32) is the vector of thermal ratings. In practice, stability limit constraints for some specific lines may also be modeled to maintain the system synchronism and voltage stability. For such a transmission line, the stability line rating is generally less than the thermal rating. Therefore, if such a line is chosen as the target line, it is difficult for the attacker to cause physical line overflow which violates the thermal limit and overheats the line. However, the attacker can still use the attack optimization structure here to maximize the power flow on this line. Under this condition, FDI attacks obtained with the attack optimization problem may result in losing synchronism between the two end buses or voltage collapse problems.

In addition, the line thermal ratings here are adopted from MATPOWER test cases for the IEEE 24-bus, 118-bus, and Polish systesms with modifications. Such limits are assumed to be constant values in this chapter. However, in practice, thermal ratings of lines are calculated based on the maximum operating temperature of the conductor. There are typical continuous operation and emergency limits. These limits may vary druing summer and winter and depend on ambient temperature and wind speed. The ampacity of the conductor determines the limiting value and this in turn is dependent on a number of factors. Therefore, if the attacker neglects these factors and only models constant thermal limits (*e.g.,* some inaccurate values obtained from the historical data), the bi-level attack optimization problems proposed in this chapter can also result in erroneous estimation on the attack consequences. This, in turn, can undermine the physical damage caused by the designed attacks.

4.4.2 Approximation on PTDF Matrix

In this chapter, a static PTDF matrix is utilized to compute DC power flows. That is, the PTDF matrix is computed only with reactances of transmission lines. However, instead of the static PTDF matrix, a dynamic PTDF matrix which is linearized at the current operating condition is utilized in practice to compute power flow. Therefore, even when the system topology remains unchanged, PTDF matrices can vary under different operating conditions (*e.g.,* different load levels). Such an approximation can also undermine the attacker's estimation on the attack consequences.

Although the approximations discussed above can undermine the attack consequences on real power systems, attackers can always intensify the attacks by achieving more system knowledge and modeling a more accurate system response. Therefore, the system control center should take such vulnerabilities into consideration.

4.5 Concluding Remarks

This chapter studies the physical system consequences of two classes of limited information FDI attacks. In the first attack class, a bi-level optimization problem is formulated to maximize the power flow on a chosen target line with attacker's perfect information in a sub-network $\mathcal{L}$ as well as estimated information of marginal generators and PTDF out of $\mathcal{L}$. It is illustrated that with an appropriately chosen sub-network and perfect localized information, the attacker can overload transmission lines with limited load shifts in both linear and non-linear models in the test system.

In second attack class, we have introduced pseudo-boundary injections to represent the power flow delivered from the external network and developed a multiple linear regression model to learn the relationship between pseudo-boundary injections and the power injections inside $\mathcal{L}$. Furthermore, we have formulated a bi-level optimization

problem is to maximize the power flow on a chosen target line with attacker's perfect information in a sub-network $\mathcal{L}$ as well as the predicted pseudo-boundary injections. It is illustrated that the attacker can overload transmission lines with the proposed bi-level attack optimization problems with both perfect and inaccurate predictions of pseudo-boundary injections. Sensitivity analysis of this attack strategy is performed on imperfect historical datasets with load varying errors, topology errors, generation dispatch errors, and non-linear power flow data. It is demonstrated that such an attack strategy is robust to all the four types of errors.

Chapter 5

VULNERABILITY ANALYSIS FRAMEWORK FROM THE PERSPECTIVE OF SYSTEM

In Chapters 3–4, the feasibility and physical consequences of the FDI attacks designed with perfect and limited information have been analyzed from the perspective of the attacker. However, how can the control center perform the vulnerability analysis *apriori* is a question that remains to be answered. In this chapter, an off-line analysis method is proposed to identify the set of sub-networks in a test system that are more prone to FDI attacks. Throughout, such sub-networks are denoted as *key sub-networks*.

5.1 Off-line Vulnerability Framework

In the following, we propose an off-line vulnerability analysis framework to identify the key sub-networks in the power system.

Algorithm 1 Off-Line Vulnerability Analysis Algorithm to Identify Key Sub-Networks

Step 1 Assume the control center has a typical operation condition corresponding to a set of historical data. Such an operation condition can be obtained as follows:

1.1 Analyze system historical data including system topologies, loads, line power flows, generation cost, capacity, and dispatch, during a long period of time. Classify the historical data by load and generation dispatch values.

1.2 Identify the sets of historical data in which the power flow on at least one line reaches 90% of the line limit over 80% of the times. Denote such a set of historical data as a *congested historical dataset*.

1.3 For each congested historical dataset, create a typical operation condition by averaging the loads at each bus.

Step 2 Select a congested line l in a congested historical dataset to perform the following analysis:

2.1 Choose the end buses of this congested line as the center buses.

2.2 Identify the sub-graph $\mathcal{S}$ corresponding to the center buses. Denote the non-boundary buses in the sub-graph as $\mathcal{I}$. Check the number of buses (denoted as $n_{\mathcal{S}}$) inside $\mathcal{S}$. If $n_{\mathcal{S}}$ is greater than the maximum number of buses that the attacker can compromise (denoted as $n_{\max}$), *i.e.,* $n_{\mathcal{S}} > n_{\max}$, go to Step 3.

Algorithm 1 Off-Line Vulnerability Analysis Algorithm to Identify Key Sub-Networks

Step 2 (continued)

2.3 Solve the following bi-level attack optimization problem with the created typical operation condition to identify the attack vector c inside $\mathcal{S}$ that can maximize the physical power flow on line l.

$$\underset{c,P}{\text{maximize}} \;\; P_l \tag{5.1}$$

$$\text{subject to} \;\; c_i = 0, \; \forall i \notin \mathcal{I} \tag{5.2}$$

$$\text{and (3.4)–(3.8).}$$

If any line k inside $\mathcal{S}$ satisfies $|P_k^*| > P_{k,\max}$, record $\mathcal{S}$ as a key sub-network.

2.4 Set all the buses inside $\mathcal{S}$ as center buses and go to Step 2.2.

Step 3 Repeat Step 2 for all congested lines in all congested historical datasets to exhaustively identify all the key sub-networks.

In the following section, we evaluate the vulnerability of the IEEE 24-bus, IEEE 118-bus, and Polish systems to FDI attacks with the proposed framework. In particular, for medium- or large-scale test systems such as the IEEE 118-bus and the Polish systems, the attack optimization problem in Step 2.3 may become intractable due to the increasing number of constraints and their associated binary variables. To overcome this difficulty, our prior work [27, 28] introduces four computationally efficient algorithms to provide upper and/or lower bounds on the objective value. These algorithms include:

1. Row generation for line limit constraints (RG), which reduces the number of line

limit constraints and their associated binary variables of the equivalent single level mixed integer linear programming (MILP) problem of the bi-level attack optimization problem using row generation.

2. Row and column generation for line and generator limit constraints (RCG), which further reduces the number of binary variables by judiciously eliminating generation limit constraints using column generation.

3. Cyber-physical-difference maximization (DM) which provides upper and lower bounds via a linear program (LP) that maximizes the difference between target line cyber and physical power flows.

4. Modified Benders' decomposition for bi-level programs (MBD) that uses Benders' decomposition to solve the original bi-level optimization problem.

In Sec. 5.2, one or more of these algorithms are used to ensure the tractability of the optimization problem in Algorithm 1.

5.2 Numerical Results

In this section, we illustrate the efficacy of the framework proposed in Algorithm 1. In particular, we assume the base case operation conditions in the modified IEEE 24-bus, IEEE 118-bus, and Polish systems (see Sec. 4.2.3) are the typical operation conditions learned from the historical datasets. The load shift factor is chosen to be 10%. The maximum number of buses that the attacker can compromise, $n_{\max}$, is set as $\frac{2}{3}$ of the total number of buses in each test system, *i.e.,* $n_{\mathcal{S}} \leq \frac{2}{3} n_{\mathcal{G}}$. In particular, RG and RCG methods are utilized to ensure the tractability of the framework in the IEEE 118-bus and Polish systems, respectively.

The vulnerability analysis framework in Algorithm 1 identifies 5, 8, and 28 key sub-networks in the IEEE 24-bus, IEEE 118-bus, and Polish systems, respectively.

The parameters of the key sub-networks in each test systems are summarized in Table. 5.3.

From the table, it can be seen that for each test system, there are key sub-networks covering a small portion of the entire system. It further verifies our observation in Chapter 4 that FDI attacks can cause line overflow even within a minuscule amount of system information. In addition, we can also see that in general, as the size of the key sub-network expands, not only the maximum line overflow value on the target line, but also the number of lines with violation resulting from FDI attacks increase. However, an opposite example of the two key sub-networks corresponding to the target line 11 in the IEEE 24-bus system shows that the attack consequences will not get worse as the key sub-network expands. It can also be verified by the second and third key sub-networks corresponding to the target line 155 in the IEEE 118-bus system. Such examples indicate that the load redistribution on buses inside the smaller key sub-network plays an essential role in causing target line overflow.

Table 5.3: Summary of the Key Sub-networks in the IEEE 24-bus System, IEEE 118-bus System, and Polish System.

	Target Line	# of Buses	# of Branches	# of Violations	Max PF (%)
IEEE 24-bus System	11	4	3	1	114.29%
		14	18	1	114.29%
	23	12	15	2	104.78%
	28	8	11	1	100.67%
		15	19	2	107.78%
IEEE 118-bus System	5	7	6	1	100.85%
		43	46	1	104.62%
	11	8	8	1	100.13%
		48	51	1	103.09%
	104	76	98	2	104.79%
	155	12	12	1	101.62%
		24	34	1	104.95%
		63	82	1	104.95%
Polish System	24	813	911	2	100.46%
		1378	1595	3	101.38%
	292	321	352	1	100.32%
		829	931	1	100.93%
		1386	1609	2	101.67%
	1381	9	8	1	100.13%

continued on next page

continued from previous page

	Target Line	# of Buses	# of Branches	# of Violations	Max PF (%)
Polish System	1381	33	33	1	100.64%
		184	194	2	101.41%
		775	874	2	102.00%
		1483	1686	2	102.14%
	1382	21	20	1	100.37%
		178	186	1	100.59%
		775	873	2	101.22%
		1483	1686	2	102.14%
	1816	11	11	1	100.78%
		121	128	1	101.67%
		688	774	1	101.67%
		1316	1502	3	101.99%
	2109	8	7	1	100.21%
		317	350	1	101.06%
		839	939	2	101.80%
		1403	1622	4	103.02%
	2110	4	3	1	100.88%
		11	10	1	101.12%
		321	354	1	102.08%
		843	944	3	103.10%

continued on next page

continued from previous page

	Target Line	# of Buses	# of Branches	# of Violations	Max PF (%)
Polish System	2110	1407	1627	3	104.74%
	2239	1152	1321	1	100.2%

*$\mathcal{S}$: Sub-network, PF: Power flow

To better protect the system, at least one secure measurement should be placed inside the smallest key sub-networks for each congestion line in each test system. Furthermore, when a specific pattern of congested lines occurs, the control center should closely monitor the load varying patterns inside the key sub-networks, so as to identify the anomalies in time.

The computation time for each test system is summarized in Table. 5.4. We can observe that Algorithm 1 can assess the vulnerability of all the test systems in less than 7 seconds. These results demonstrate the computational efficiency of Algorithm 1, which further indicates its potential to be employed as an on-line vulnerability analysis tool.

Table 5.4: Summary of The Computation Times.

Test System	IEEE 24-bus	IEEE 118-bus	Polish
Times (s)	2.00	6.31	3.19

5.3 Concluding Remarks

In this chapter, we focused on the vulnerability analysis of power systems to FDI attacks from the perspective of the system control center. We proposed an off-line

vulnerability analysis framework to analyze historical data, identify patterns of congested lines, assess the vulnerability of the sub-network surrounding each congested line layer by layer, and finally identify key sub-networks that are prone to FDI attacks. It is demonstrated that this framework is both accurate and efficient in assessing the vulnerability of the test system apriori. How to identify key measurements to keep secure and to analyze load varying behaviors inside the identified key sub-networks is crucial future work that needed to further protect the system from FDI attacks.

Chapter 6

FALSE DATA INJECTION ATTACKS ON PHASOR MEASUREMENTS THAT BYPASS LOW-RANK DECOMPOSITION

In this chapter, we focus on the vulnerability of PMUs to FDI attacks.

Recently, using measurements obtained from deployed PMUs in the grid, [50] and [51] illustrate the low-rank nature of PMU data. These approaches suggest that PMU measurements can be modeled as a matrix to capture both the temporal aspects (*e.g.*, via the rows of the matrix) and the spatial aspects (for each time instant via the columns).

As reviewed in Sec. 1.2, in [38, 39, 40], low-rank decomposition (LD) has been proposed to detect FDI attacks on the electric power system using a block of consecutive measurement data. On the other hand, the FDI attacks of most interest are those in which the attacker is not omniscient and omnipresent — this limited knowledge and limited capabilities of FDI attacks are often captured (see, for *e.g.*, [12, 14, 52, 26, 29, 45, 27, 53]) by restricting attacker knowledge to a subset of the network and restricting counterfeits to a small number of meters, respectively. This latter restriction along with the above mentioned low-rank properties of a block of PMU data suggests that the resulting counterfeit PMU measurement matrix can be viewed as a linear combination of a low-rank (actual) measurement matrix and a sparse attack matrix (counterfeit additions to measurement).

In [38], the authors propose a LD approach (introduced in [41] for arbitrary sparse datasets), for temporal SCADA data; specifically, they demonstrate that attacks designed without knowledge of the temporal correlations of the SCADA measurements can be detected by solving an LD problem. Furthermore, their model assumes that

while the FDI attack matrix is sparse in each time instant, the attacker attacks a different set of measurements. While such a model is quite general, for attacks designed with a specific effect (financial or physical damage), sustaining attacks over time on the same meters can have more impact. Focusing on such sustained attacks, for PMU data, the authors of [39, 40] show that an LD-based detector can identify column sparse FDI attack matrix where the column sparsity is a result of the assumption that the attacker attacks the same set of PMU measurements over time.

Following [39, 40] we model PMU data as a low-rank matrix. Furthermore, focusing on impactful FDI attacks, our attack model involves sustained attacks on the same meters over time, *i.e.*, column sparse attacks (using the nomenclature that rows and columns indicate spatial and temporal data, respectively). Although the LD detector shows good performance in detecting column sparse unobservable FDI attacks on both synthetic data and some field PMU data [39, 40], a question that needs to be addressed is the following: if an attacker has knowledge of the time correlation of the PMU data, can it take advantage of such knowledge and design FDI attacks that can bypass the detector? In this chapter, we assume the attacker has the ability to predict the system dynamics, and we introduce a new class of FDI attacks that can bypass the LD detector. These attacks are designed with a convex optimization problem. We prove that the LD detector cannot identify the exact set of states that are modified by the attacker. We demonstrate that such attacks are unobservable for both traditional bad data detectors and the LD detector on both the IEEE 24-bus and IEEE 118-bus systems.

6.1 Preliminaries

In this section, we introduce the models for the SE with phasor measurements, FDI attacks on PMU, and the LD detector. Throughout, we assume there are n_b

buses, n_{br} branches, n_g generators, and n_z measurements in the system.

6.1.1 State Estimation with Phasor Measurements

PMUs collect complex bus voltage and branch current measurements. The reporting rate of the PMU measurements is usually 30 times per second [54]. These measurements have a linear relationship with the complex bus voltage states. At each time instant t, the PMU measurement model can be written as

$$z_t = Hx_t + e_t \tag{6.1}$$

where at time instant t, z_t is the $n_z \times 1$ measurement vector; x_t is the state vector of complex bus voltage; e_t is an $n_z \times 1$ noise vector assumed to be composed of independent Gaussian random variables; the complex matrix H is the $n_z \times n_b$ dependency matrix between measurements and states. Note that the state can be estimated based on PMU measurements via a single weighted least squares (WLS) [54], unlike traditional SCADA-based SE which requires multiple iterations due to the nonlinearity of the measurement function [42].

One possible way to process PMU data is to collect over a block of time (*e.g.,* 5 to 20 seconds) and then process them as a batch (see for example [55]). We adopt this approach and write the PMU measurements as a matrix where each row vector corresponds to PMU measurements at one time instant and each column vector consists of the measurements collected in the same channel over a period of times. The PMU measurements in (6.1) over N time instants can then be collected as

$$Z = XH^T + E \tag{6.2}$$

where matrices $Z = \left[z_1^T;\ z_2^T; \ldots;\ z_N^T\right]$, $X = \left[x_1^T;\ x_2^T; \ldots; x_N^T\right]$, and $E = \left[e_1^T;\ e_2^T; \ldots; e_N^T\right]$ are PMU measurement matrix, state matrix, and noise matrix, respectively. Note that

z_t^T, x_t^T, and e_t^T for $t = 1, 2, \ldots, N$ are the transpose of the measurement, state, and noise column vectors, respectively, in (6.1).

6.1.2 Unobservable FDI Attack on PMU

Assume the attacker has control of the measurements in a subset $\mathcal{S}$ of the network, denoted as the attack subgraph. As in Chapter 3, we first distinguish between two types of buses in the network: *load buses* that have load directly connected to them, and *non-load buses* with no load. We assume $\mathcal{S}$ is bounded by load buses. The set of measurements in $\mathcal{S}$ are denoted as $\mathcal{J}$. In the absence of noise, an attack is defined to be unobservable if

$$\tilde{Z} = Z + D = Z + CH^T = (X + C) H^T + E \tag{6.3}$$

where $\tilde{Z}$ is the $N \times n_z$ post-attack measurement matrix, D is the $N \times n_z$ attacked measurements matrix such that $D = CH^T$, and C is the $N \times n_b$ attack matrix. In the following, we define the set of non-zero columns in a matrix as its column support, written as $supp\,(\cdot)$. Note that the attacker is constrained to inject false data only in the measurements in $\mathcal{J}$. Thus, D is a column sparse matrix where $supp\,(D) \subseteq \mathcal{J}$. One natural way to form a column sparse D is to choose a column sparse C.

Prior work [12, 14, 52, 26, 29, 45, 27] considers a special case of (6.3) with only one time instant, *i.e.,* $N = 1$. These works show that traditional bad data detectors based on measurement residuals cannot detect such FDI attacks.

6.1.3 Prior Work: Attack Detection Based on Low-Rank Matrix Decomposition

Traditional bad data detectors based on measurement residuals cannot detect the FDI attacks introduced in (6.3). However, exploiting the low-rank nature of the high-dimensional PMU data matrix Z, the authors in [39] propose a new attack detection

mechanism based on LD so as to separate the low-rank matrix Z and column sparse matrix CH^T in (6.3). We now briefly review their attack assumptions and detection methodology.

Given a measurement matrix $\tilde{Z}^{(\mathrm{LD})}$, the measurement matrix without attack, $Z^{(\mathrm{LD})}$, and the attack matrix $C^{(\mathrm{LD})}$ can be identified by solving the following convex optimization problem:

$$\underset{Z^{(\mathrm{LD})}\in\mathbb{C}^{N\times n_z}, C^{\mathrm{LD}}\in\mathbb{C}^{N\times n_b}}{\text{minimize}} \quad \|Z^{(\mathrm{LD})}\|_* + \lambda\|C^{(\mathrm{LD})}\|_{1,2} \tag{6.4}$$

$$\text{subject to} \quad \tilde{Z}^{(\mathrm{LD})} = Z^{(\mathrm{LD})} + C^{(\mathrm{LD})}\tilde{H}^T \tag{6.5}$$

where $\|Z^{(\mathrm{LD})}\|_*$ is the nuclear norm of $Z^{(\mathrm{LD})}$; $\|C^{(\mathrm{LD})}\|_{1,2}$ is the $l_{1,2}$-norm of $C^{(\mathrm{LD})}$, *i.e.,* the sum of l_2-norm of columns in $C^{(\mathrm{LD})}$; λ is a weight factor; and $\tilde{H}$ is the normalized dependency matrix, where for each row vector H_i, $\tilde{H}_i = H_i/\|H_i\|$. The objective (6.4) is to minimize the rank of $Z^{*(\mathrm{LD})}$ (captured by its nuclear norm) and the column sparsity of $C^{*(\mathrm{LD})}$ (captured by its $l_{1,2}$-norm).

After obtaining the optimal solution, $\left(Z^{*(\mathrm{LD})}, C^{*(\mathrm{LD})}\right)$ for (6.4)–(6.5), the set of attacked measurements and states, $supp\left(C^{*(\mathrm{LD})}\tilde{H}^T\right)$ and $supp\left(C^{*(\mathrm{LD})}\right)$, respectively, can be identified as the column support of $C^{*(\mathrm{LD})}\tilde{H}^T$ and $C^{*(\mathrm{LD})}$. Assume there exists unobservable attacks in $\tilde{Z}^{(\mathrm{LD})}$, such that $\tilde{Z}^{(\mathrm{LD})} = Z + C\tilde{H}^T$. The authors prove that for a specific range of λ, *i.e.,* $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, the optimization in (6.4) can successfully identify $supp(C)$, *i.e.,* $supp\left(C^{*(\mathrm{LD})}\right) = supp\left(C\right)$, under the assumption that every nonzero column of $C\tilde{H}^T$ does not lie in the column space of Z.

6.2 FDI Attack Exploiting Low-Rank Property of PMU Measurement Matrix

In this section, we introduce a class of FDI attacks that cannot be detected by the LD detector in (6.4)–(6.5). We assume that the attacker has the following knowledge and capabilities:

1. The attacker has full system topology information.

2. The attacker can perfectly predict the measurements in the following N instances and has the capability to estimate the predicted states.

3. The attacker has control of the measurements in a subset $\mathcal{S}$ of the network.

6.2.1 Attack Strategy

Given a PMU measurement matrix Z and the potential attacked states set $\mathcal{I}$, we propose the following optimization problem to design FDI attacks:

$$\underset{C\in\mathbb{C}^{N\times n_b}}{\text{minimize}} \quad \|Z + C\tilde{H}^T\|_* \tag{6.6}$$

$$\text{subject to} \quad supp(C) \subseteq \mathcal{I} \tag{6.7}$$

where $\|\cdot\|_*$ denotes the nuclear norm. For optimal solution C^*, the optimal post-attack measurement matrix denoted as $\tilde{Z}^*$ can be written as

$$\tilde{Z}^* = Z + C^*\tilde{H}^T. \tag{6.8}$$

The goal of the attacker is to ensure that the attacked measurement matrix $\tilde{Z}^*$ is low-rank when Z is low-rank. This can be approximated by minimizing the nuclear norm of $\tilde{Z}^*$ as (6.6). Constraint (6.7) ensures that the attacker can only attack states in $\mathcal{I}$, *i.e.,* C^* is a column sparse matrix.

In the following, we prove that either $\tilde{Z}^*$ bypasses the LD detector (*i.e.,* results in $C^{*(LD)} = 0$), or the LD detector identifies at least one measurement as corrupted that is not.

Theorem 3. *Assume the attack-free measurement matrix Z can bypass the LD detector, i.e., for $\tilde{Z}^{(LD)} = Z$, $\left(Z^{*(LD)}, C^{*(LD)}\right) = (Z, \mathbf{0})$. Assume the solution C^* of (6.6)–(6.7) is non-zero. Then using $\tilde{Z}^*$ in the LD detector, the resulting $C^{*(LD)}$ satisfies that either $C^{*(LD)} = \mathbf{0}$, or supp $\left(C^{*(LD)}\right) \not\subseteq$ supp (C^*).*

Proof. First, we prove that for a given Z, $\|\tilde{Z}^*\|_* \leq \|Z\|_*$. For a given Z, $C = \mathbf{0}$ is always a feasible solution for (6.6)–(6.7). For $C = \mathbf{0}$, the objective $\|\tilde{Z}\|_* = \|Z + C\tilde{H}^T\|_* = \|Z\|_*$. Since we minimize (6.6), the objective of C^* is less than or equal to that of the feasible solution $\mathbf{0}$. That is, $\|\tilde{Z}^*\|_* \leq \|Z\|_*$ always holds.

Suppose Z can bypass the LD detector. That is, for input $\tilde{Z}^{(\mathrm{LD})} = Z$, $\left(Z^{*(\mathrm{LD})},\ C^{*(\mathrm{LD})}\right) = (Z, \mathbf{0})$. As we just proved

$$\|Z + C^*\tilde{H}^T\|_* = \|\tilde{Z}^*\|_* \leq \|Z\|_* = \|Z^{*(\mathrm{LD})}\|_*. \tag{6.9}$$

Thus,

$$\|Z + C^*\tilde{H}^T\|_* \leq \|Z\|_* \leq \|Z\|_* + \lambda\|C^*\|_{1,2}. \tag{6.10}$$

Let $C^{*(\mathrm{LD})}$ be the optimal solution of the LD detector for $\tilde{Z}^{(\mathrm{LD})} = \tilde{Z}^*$. The objective (6.4) for $\tilde{Z}^*$ satisfies

$$\|\tilde{Z}^* - C^{*(\mathrm{LD})}\tilde{H}^T\|_* + \lambda\|C^{*(\mathrm{LD})}\|_{1,2} \leq \|\tilde{Z}^*\|_* \leq \|Z\|_*. \tag{6.11}$$

Note that $\|\tilde{Z}^* - C^{*(\mathrm{LD})}\tilde{H}^T\|_*$ can be rewritten as $\|Z + \left(C^* - C^{*(\mathrm{LD})}\right)\tilde{H}^T\|_*$.

If $supp\left(C^{*(\mathrm{LD})}\right) \subseteq \mathcal{I}$, then

$$\|Z + \left(C^* - C^{*(\mathrm{LD})}\right)\tilde{H}^T\|_* \geq \|Z + C^*\tilde{H}^T\|_* \tag{6.12}$$

since C^* is the optimal solution for (6.6)–(6.7). That is, $\tilde{Z}^*$ and $C^{*(\mathrm{LD})}$ satisfy

$$\|\tilde{Z}^* - C^{*(\mathrm{LD})}\tilde{H}^T\|_* + \lambda\|C^{*(\mathrm{LD})}\|_{1,2} \geq \|\tilde{Z}^*\|_*. \tag{6.13}$$

Therefore, the only solution that can satisfy both (6.11) and (6.13) is $C^{*(\mathrm{LD})} = \mathbf{0}$. □

6.2.2 Numerical Results

In this subsection, we illustrate the efficacy of the unobservable FDI attacks introduced in Sec. 6.2.1. To this end, we first solve the attack optimization problem

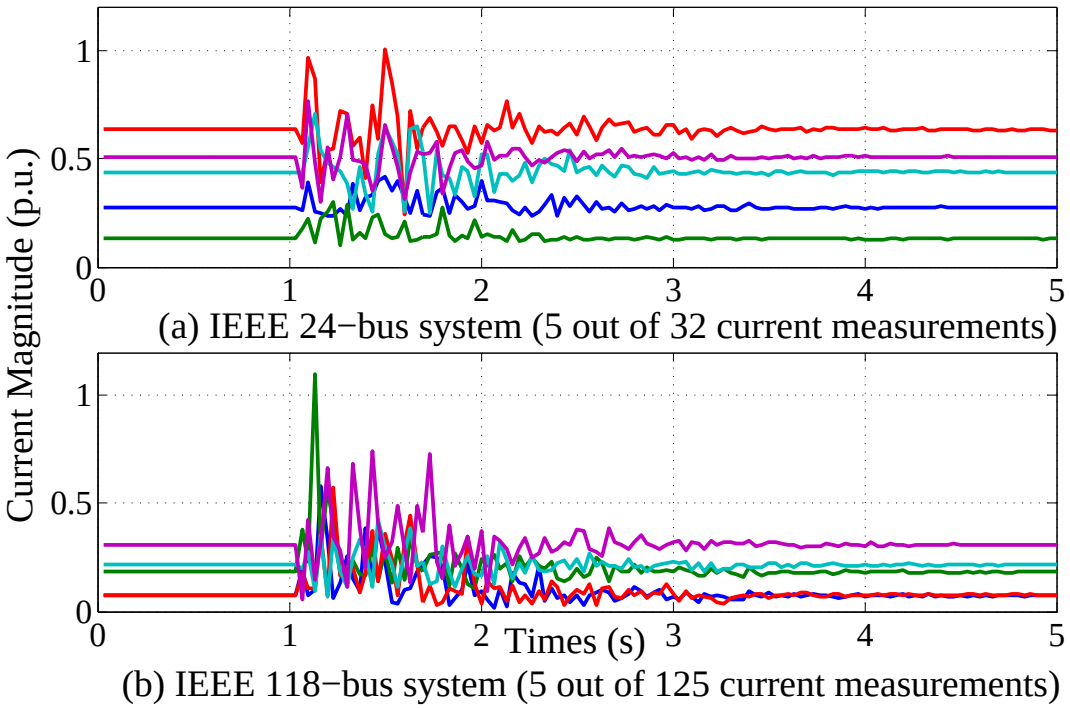


Figure 6.1: Current Magnitudes of The Synthetic PMU Data.

in (6.6)–(6.7) to find the optimal attack matrix C^*. Subsequently, we construct the post-attack measurement matrix $\tilde{Z}^*$ with C^* as (6.8). Finally, we solve the LD detection optimization problem (6.4)–(6.5) for $\tilde{Z}^*$ to check if the attack matrix C^* is detected. Throughout, we assume that the LD detector selects 2 seconds worth of PMU measurements data, *i.e.,* $N = 60$, while the attacker injects bad data. The test systems include the IEEE 24-bus reliability test system (RTS) and IEEE 118-bus system. The convex optimization problems for LD detection and attack design are solved with MOSEK. In the LD detection optimization problem, the weight λ is chosen to be 1.05 for both the IEEE 24-bus and the IEEE 118-bus systems.

We assume both test systems are fully observable with PMU measurements. This is achieved by solving an optimal PMU placement problem as introduced in [56]. The details of the PMU locations and measurements for both test systems are summarized in Table 6.2. In [39], the authors demonstrate an actual PMU dataset, which we do

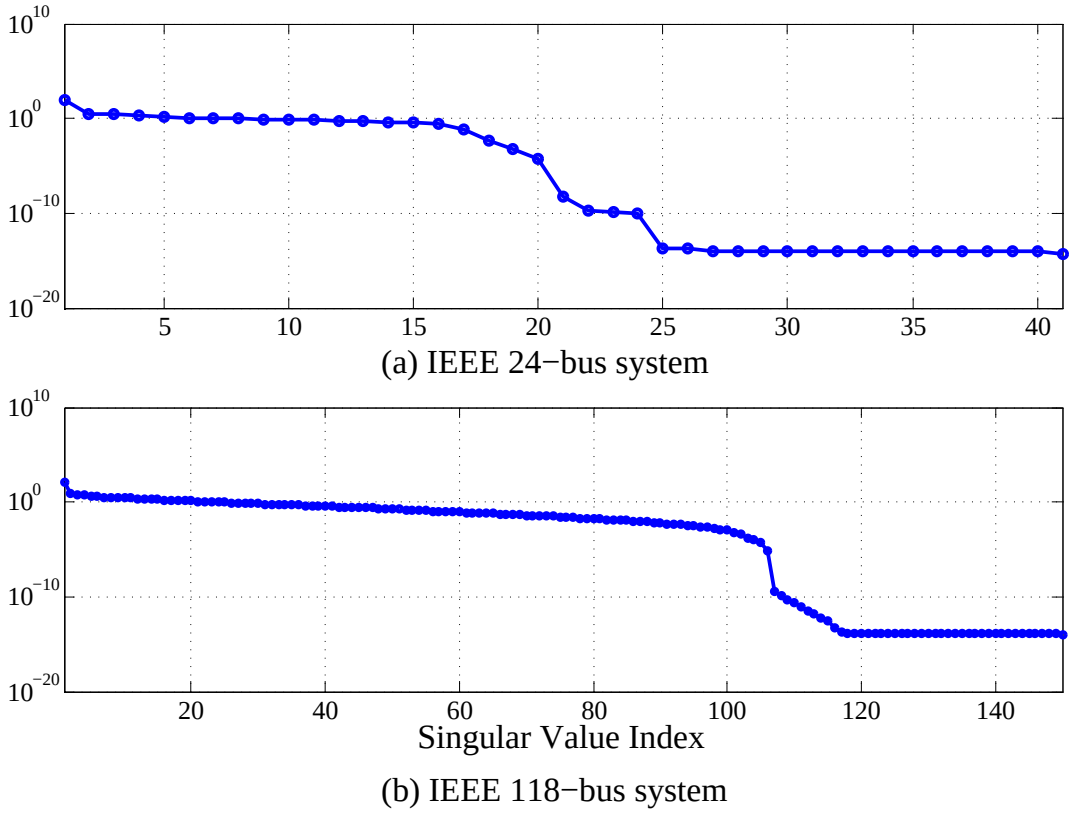


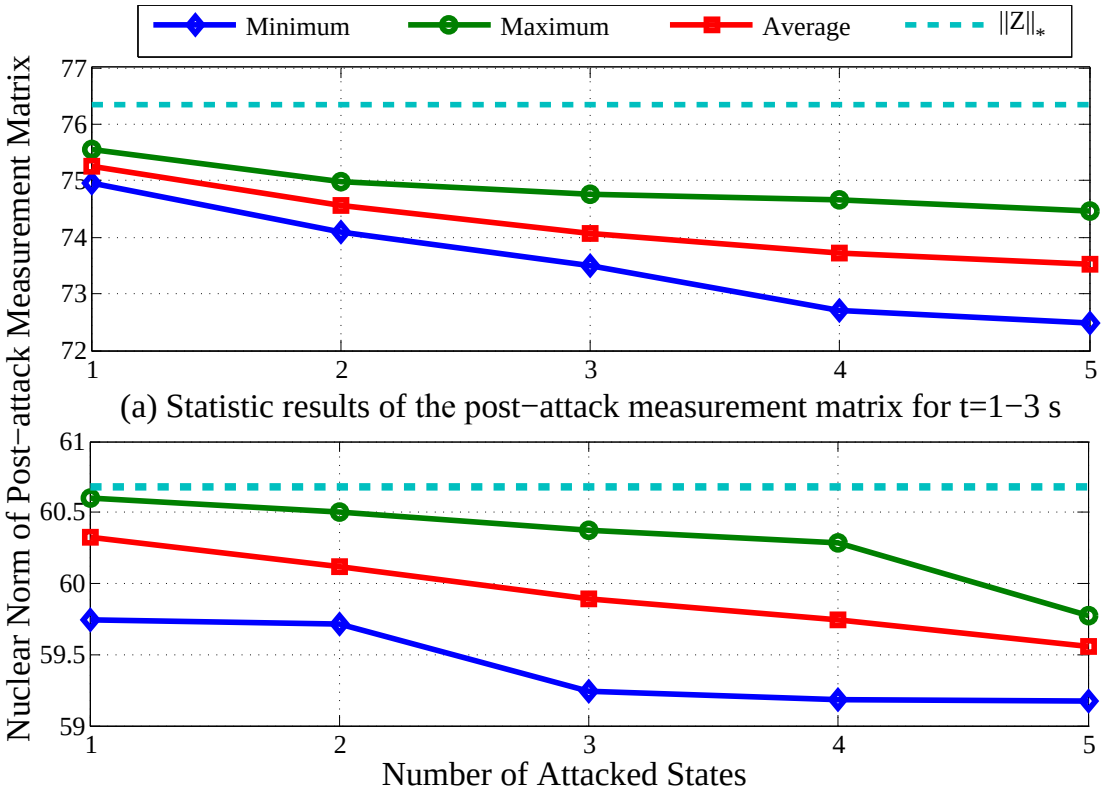
Figure 6.2: Singular Values of The Synthetic PMU Data Matrix in Decreasing Order.

not have access to. Therefore, to make a fair comparison, we generate synthetic PMU data over 5 seconds in each test system. To model realistic data with a disturbance, at the first time instant t after 1 second, we change the load at each bus by adding a random value d to the base load, such that $d \sim \mathcal{N}\left(0, \frac{60}{1.1^{(t-31)}}\right)$. We then solve an AC power flow to obtain the measured phasors of bus voltage and branch current as measurements at time instant t. The resulting synthetic measurements for the IEEE 24-bus system and the IEEE 118-bus system are partially illustrated in Fig 6.1. The singular values for the synthetic measurement matrices for the IEEE 24-bus system and the IEEE 118-bus system are illustrated in Fig 6.2. It can be seen that these synthetic measurements have the same low-rank property as the actual PMU data as illustrated in [39]. The synthetic data is then broken into two parts for testing,

one for t =1–3 seconds, the other for t =3–5 seconds. Observe that the measurement matrix for t=1–3 seconds has more disturbances than that for t=3–5. Furthermore, we assume noiseless measurements, *i.e.*, $E = \mathbf{0}$ in (6.3).

Table 6.1: Statistic Results of $\|\tilde{Z}^*\|_*$ in The IEEE 118-Bus System.

Time Period	$\|\tilde{Z}^*\|$			$\|Z\|$
	Min	Max	Ave	
1–3 second	116.1	116.7	116.5	116.8
3–5 second	56.9	57.1	57.0	57.1



(a) Statistic results of the post−attack measurement matrix for t=1−3 s

(b) Statistic results of the post−attack measurement matrix for t=3−5 s

Figure 6.3: Statistic Results of $\|\tilde{Z}^*\|_*$ in The IEEE 24-Bus System.

We exhaustively generate the unobservable attacks with all potential attacked

state sets $\mathcal{I}$ for which $1 \leq |\mathcal{I}| \leq 5$ in the IEEE 24-bus system; for tractability we consider only $|\mathcal{I}| = 1$ in the IEEE 118-bus system. Specifically, as observed in our prior work [26, 29], unobservable attacks must be designed inside a subgraph $\mathcal{S}$ which is bounded by load buses. In $\mathcal{S}$, the states of all the non-bounded buses (including load and non-load buses) have to be changed. In this subsection, the attacked state sets $\mathcal{I}$ are selected according to this rule.

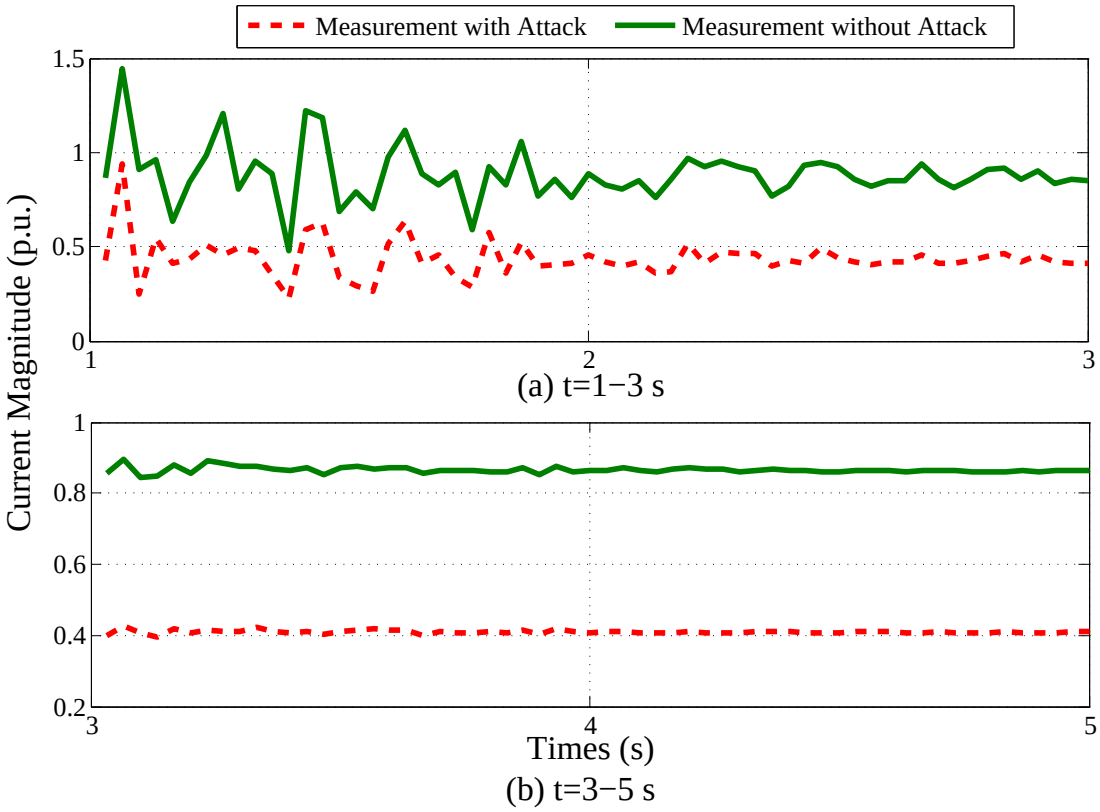


Figure 6.4: Magnitude of The From Side Current Measurement on Line 12 in The IEEE 24-Bus System with $\mathcal{I} = \{8\}$.

For every attack we tested, the LD detector is completely bypassed, *i.e.*, $C^{*(\mathrm{LD})} = \mathbf{0}$. We summarize the statistic results including maximum, minimum, and average values of $\|\tilde{Z}^*\|_*$ for the IEEE 24-bus system and the IEEE 118-bus system in Fig. 6.3 and Table 6.1, respectively. From these results, it can be seen that for every attack we tested, $\|\tilde{Z}^*\|_* \leq \|Z\|_*$ always holds. In addition, in Fig. 6.3, we also find that for

the IEEE 24-bus system, $\|\tilde{Z}^*\|_*$ gradually decreases as the number of attacked state increases.

We now illustrate a typical case in the IEEE 24-bus system in with $\mathcal{I} = \{8\}$. In Fig. 6.4, the magnitudes of the from side current measurement on line 12 are demonstrated, before and after attack. From this case, we can see that the designed attack accurately captures the temporal correlation of the PMU measurements.

These observations are consistent with Theorem 3. In fact, they are stronger than Theorem 3 since we did not find any case where the LD detector results in $C^{*(\mathrm{LD})}$ such that $C^{*(\mathrm{LD})} \neq \mathbf{0}$ and $supp\left(C^{*(\mathrm{LD})}\right) \not\subseteq \mathcal{I}$.

Table 6.2: Monitored PMU Measurements in Both The IEEE 24-Bus System and The IEEE 118-Bus System.

IEEE 24-bus System	Voltage (Buses with PMU)	1, 2, 7, 9, 10, 11, 15, 17, 20
	Current (From Side)	1, 2, 3, 4, 5, 11, 14, 15, 16, 17, 18, 19, 24, 25, 26, 27, 30, 31, 36, 37
	Current (To Side)	1, 6, 8, 9, 10, 12, 13, 14, 16, 28, 34, 35
IEEE 118-bus System	Voltage (Buses with PMU)	2, 5, 10, 12, 15, 17, 21, 25, 29, 34, 37, 41, 45, 49, 53, 56, 62, 64, 72, 73, 75, 77, 80, 85, 87, 91, 94, 101, 105, 110, 114, 116
	Current (From Side)	5, 11, 13, 17, 20, 21, 23, 26, 28, 33, 39, 40, 44, 49, 50, 52, 53, 58, 60, 62, 68, 70, 71, 74, 75, 76, 80, 82, 85, 86, 95, 97, 98, 99, 100, 101, 106, 120, 121, 123, 124, 128, 133, 135, 136, 143, 147, 148, 150, 151, 152, 153, 155, 162, 169, 170, 171, 176, 177, 178, 182, 184, 185
	Current (To Side)	1, 3, 4, 8, 9, 12, 13, 14, 15, 18, 19, 21, 22, 27, 31, 32, 35, 36, 45, 47, 48, 50, 51, 56, 61, 65, 66, 67, 68, 69, 73, 78, 79, 91, 92, 94, 111, 112, 113, 115, 116, 117, 118, 119, 120, 123, 124, 125, 127, 131, 132, 134, 140, 145, 146, 160, 166, 168, 174, 175, 180, 183

6.3 Worst-case Physical Line Overflow Attacks

In this section, we propose a heuristic method to study the physical consequences of worst-case unobservable FDI attacks on phasor measurements. Specifically, we assume the system response at each time instance satisfies OPF. Although this assumption is far fetched now, it is made here for evaluation of the worst-case attack consequences. The FDI attacks studied in this section are able to (i) bypass the LD detector, and (ii) result in physical line overflow which cannot be found in the cyber layer. Besides the knowledge and capabilities introduced in Sec. 6.2, we assume that the attacker has additional knowledge as follows:

1. The attacker has knowledge of load distribution, generation costs, and line thermal limits of the system.

2. The attacker has perfect prediction of the load varying patterns in the following N instances.

6.3.1 Attack Strategy

In this subsection, we introduce a bi-level attack optimization problem with one leader (first level) and multiple independent followers (second level) to formulate the worst-case line overflow FDI attacks. The leader problem formulates the attacker's limitation. The t^{th} independent follower problem formulates DC OPF under attacks as post-attack system responses at time instance t, $t = 1, ..., N$. In order to ensure the tractability of the problem, we only consider DC power flow model and formulate the matrix of the voltage angle states, Θ, instead of that of the complex voltage states, X. The voltage angle attack matrix is denoted with C' so as to distinguish with the

complex voltage attack matrix C. The formulation is as follows:

$$\underset{\Theta,P}{\text{maximize}} \quad \sum_{t=1}^{N} P_{l,t} \tag{6.14}$$

$$\text{subject to} \quad \|\Theta^* + C'\|_* \leq N_* \tag{6.15}$$

$$\mathcal{P}_{\mathcal{I}_a}(C') = C' \tag{6.16}$$

$$-\tau P_D^T \leq HC'^T \leq \tau P_D^T \tag{6.17}$$

$$P = \Gamma \Theta^{*T} \tag{6.18}$$

$$\{P_{G,t}^*, \Theta_t^*\} = \arg \left\{ \min_{P_{G,t}, \Theta_t} \quad \sum_{g=1}^{n_g} f_g P_{Gg,t} \right\} \quad t = 1, ..., N \tag{6.19}$$

$$\text{subject to} \quad GP_{G,t}^T - H\Theta_t^T = P_{D,t}^T \qquad (\lambda_t) \tag{6.20}$$

$$-P_{\max} \leq \Gamma(\Theta_t + C_t')^T \leq P_{\max} \qquad \left(\mu_t^{\mp}\right) \tag{6.21}$$

$$P_{G,\min} \leq P_{G,t}^T \leq P_{G,\max} \qquad \left(\alpha_t^{\mp}\right) \tag{6.22}$$

where for any matrix A, A_t represents the tth row vector of A, P is the $t \times n_{br}$ real power flow matrix, P_G is the $t \times n_g$ real power generation output matrix, P_D is the $t \times n_b$ real power load matrix, $P_{\max}$ is the $n_{br} \times 1$ vector of thermal limits, $P_{G,\max}$ and $P_{G,\min}$ are the $n_g \times 1$ vectors of maximum and minimum generation limits, respectively, G is the $n_b \times n_g$ generator-to-bus connectivity matrix, f is the $n_g \times 1$ generation cost vector, H is the $n_b \times n_b$ dependency matrix between power injection and voltage angle state, Γ is the $n_{br} \times n_b$ dependency matrix between power flow and voltage angle state, τ is the load shift factor, N_* is the nuclear norm constraint limit, N is the number of time instances formulated in the problem, λ_t, $\mu_t^{\mp}$, and $\alpha_t^{\mp}$ are the vectors of dual variables for the second level node balance, thermal limit, and generation limit constraints at time instance t, respectively.

The objective of the optimal problem is to maximize the sum of physical power flows on target line l over N time instances. In the first level, the attack vector is

chosen subject to the nuclear norm constraint of the attack matrix in (6.15), attacked states limitation in (6.16), and the load shift limitation in (6.17). In the second level, the system responses to the attack matrix over N time instances determined in the first level are modeled via N independent DC OPF problems in (6.19)−(6.22).

The bi-level optimization problem introduced above is non-linear and non-convex. For tractability, a modified Benders' decomposition method is used to solve this problem. This method is first introduced in our prior work [28] to solve a bi-level attack optimization problem with one leader and one follower. The problem in [28] is to determine an attack vector at one time instance to maximize the physical power flow on the target line. This bi-level attack optimization problem is first converted to an equivalent single level problem by replacing the second level problem with its optimal conditions. After then, the equivalent single level problem is decomposed into a master problem (MP) and a slave problem (SP). In this section, we further extend this method to solve the one leader multiple independent followers problem to determine an attack matrix for N time instances. That is, we decompose the original bi-level attack optimization problem into one MP and N independent SPs.

The MP takes the following form:

$$\begin{aligned} &\underset{C',\gamma}{\text{minimize}} \;\; \gamma \\ &\text{subject to} \;\; (6.15)-(6.17) \end{aligned} \tag{6.23}$$

where γ is a variable introduced to represent $\sum_{t=1}^{N} P_{l,t}^{*}$, which will be updated by adding cuts. Note that in MP, the state matrix Θ is fixed as a constant matrix.

The SP at t instance takes the following form:

$$\underset{\Theta_t, P_{G,t}}{\text{minimize}} \;\; -\Gamma\Theta_t^T \tag{6.24}$$

$$\text{subject to} \;\; GP_{G,t}^T - H\Theta_t^T = P_{D,t}^T \qquad \left(\tilde{\lambda}_t\right) \tag{6.25}$$

$$-P_{\max} \leq \Gamma(\Theta_t + C'_t)^T \leq P_{\max} \qquad \left(\tilde{\mu}_t^{\mp}\right) \tag{6.26}$$

$$P_{G,\min} \leq P_{G,t}^T \leq P_{G,\max} \qquad \left(\tilde{\alpha}_t^{\mp}\right) \tag{6.27}$$

$$\begin{aligned}\sum_{g=1}^{n_g} f_g P_{Gg,t} = \lambda_t^T P_{D,t}^T + \alpha_t^{+T} P_{G,\max} - \alpha_t^{-T} P_{G,\min} \\ + \mu_t^{+T}\left(P_{\max} - \Gamma C_t'^T\right) \\ + \mu_t^{-T}\left(P_{\max} + \Gamma C_t'^T\right)\end{aligned} \tag{6.28}$$

$$G^T\lambda_t + \alpha_t^+ - \alpha_t^- = f_g \qquad (\omega_t) \tag{6.29}$$

$$-H^T\lambda_t + \Gamma^T\mu_t^+ - \Gamma^T\mu_t^- = 0 \tag{6.30}$$

$$[\mu_t^+ \;\; \mu_t^- \;\; \alpha_t^+ \;\; \alpha_t^-] \leq 0 \tag{6.31}$$

where $\tilde{\alpha}_t$, $\tilde{\mu}_t^{\mp}$, $\tilde{\alpha}_t^{\mp}$, and ω_t are dual variable vectors of the corresponding constraints. The objective of SP in (6.24) is to maximize the physical power flow at time instance t subject to the optimal conditions of the DC OPF. Note that, the attack vector C'_t is fixed at C'^*_t obtained from MP. The optimal conditions include the primal feasibility constraints (6.25)–(6.27), the strong duality constraint (6.28), and the dual feasibility constraints (6.29)–(6.31).

At the optimal solution of the SP at time instance t satisfies

$$\begin{aligned}P_{l,t}^* = \Gamma\Theta_t^{T*} = \tilde{\lambda}_t^T P_{D,t} + \tilde{\mu}_t^{+T}\left(P_{\max} - \Gamma C_t'^{*T}\right) \\ + \tilde{\mu}_t^{-T}\left(P_{\max} + \Gamma C_t'^{*T}\right) + \tilde{\alpha}_t^{+T} P_{G,\max} \\ - \tilde{\alpha}_t^{-T} P_{G,\min} + \omega_t^T f_g.\end{aligned} \tag{6.32}$$

Therefore, one way to add an optimality cut in the MP is to take the summation of the right hand sides of (6.32) for $t = 1, ..., N$, such as

$$\begin{aligned}\gamma \geq -\sum_{t=1}^{N}\Big[\tilde{\lambda}_t^{*T} P_{D,t} + \tilde{\mu}_t^{+*T}\left(P_{\max} - \Gamma C_t'^T\right) \\ + \tilde{\mu}_t^{-*T}\left(P_{\max} + \Gamma C_t'^T\right) + \tilde{\alpha}_t^{+*T} P_{G,\max}\end{aligned} \tag{6.33}$$

$$-\tilde{\alpha}_t^{-*T} P_{G,\min} + \omega_t^{*T} f_g \big] \, .$$

Note that the fixed constant C'^* is replaced with the variable matrix C' and the dual variable vectors obtained from SP are fixed at the optimal solutions $\tilde{\alpha}_t^*$, $\tilde{\mu}_t^{\mp *}$, $\tilde{\alpha}_t^{\mp *}$, and ω_t^* for $t = 1, ..., N$. The MP and SP can then be solved iteratively, with the MP updating C'^* and the SP updating cuts in each iteration. The method is summarized in Algorithm 2.

Algorithm 2 Modified Benders' Decomposition Algorithm

Step 0: **Initialization:**

- Set the initial iteration counter k, the initial matrix C', and the lower bound of the objective function, $\mathrm{ZD}^{(k)}$, as $k = 1$, $C'^{(k)} = \mathbf{0}$, and $\mathrm{ZD}^{(k)} = -10^5$, respectively.

Step 1: **SP solutions:**

- Fix C'_t with $C'^{(k)}_t$ and solve all SPs for $t = 1, ..., N$;
- Update the upper bound of the objective function, $\mathrm{ZU}^{(k)}$, as $\mathrm{ZU}^{(k)} = \sum_{t=1}^{N} \Gamma_l \Theta_t^{*(k)} = \sum_{t=1}^{N} P_{l,t}^{*(k)}$.

Step 2: **Convergence checking:**

- If $\frac{|\mathrm{ZU}^{(k)} - \mathrm{ZD}^{(k)}|}{|\mathrm{ZD}^{(k)}|} \leq \varepsilon$, the solution with a level of accuracy ε of the objective function is $C'^{*} = C'^{(k)}$, $\Theta^{*} = \Theta^{(k)}$, and $P_G^{*} = P_G^{(k)}$. Otherwise, the algorithm continues with the next step.

Step 3: **MP solution:**

- Add an optimality cut (6.33) in MP. Note that the dual variable vectors in (6.33) are fixed with $\tilde{\alpha}_t^{*(k)}$, $\tilde{\mu}_t^{\mp *(k)}$, $\tilde{\alpha}_t^{\mp *(k)}$, and $\omega_t^{*(k)}$ for $t = 1, ..., N$, and the Θ in (6.15) is fixed with $\Theta^{*(k)}$.

Algorithm 2 Modified Benders' Decomposition Algorithm (Continued)

Step 3: **MP solution:** (continued)

- Update the iteration counter, $k \leftarrow k+1$, and solve the MP. Note that at every iteration a new constraint is added. The solution of the MP provides $C'^{*(k)}$ and $\gamma^{*(k)}$.
- Update the objective function lower bound, $\text{ZD}^{(k)} = \gamma^{(k)}$. The algorithm continues in Step 1.

In particular, since the original bi-level optimization problem is non-convex, Algorithm 2 is not guaranteed to give the global optimal solution of the original attack optimization problem [57]. Therefore, the optimal solution obtained with Algorithm 2, denoted as $\sum_{t=1}^{N} P_{l,t}^{*(\text{BD})}$, is a lower bound of the global optimal solution $\sum_{t=1}^{N} P_{l,t}^{*}$.

Since the LD detector is designed for non-linear measurement matrix, we need to construct the post-attack non-linear measurement matrix with the attack matrix C'. Prior work [26] has introduced a method to construct a non-linear post-attack measurement vector with a DC attack vector. We now modify this method to construct the complex post-attack measurement matrix with the DC attack matrix as follows:

Algorithm 3 Convert DC Attack to AC Attack

1. Identify the attack subgraph with $\mathcal{I}_a$.
2. At time instance t, collect measurements in $\mathcal{S}$, perform local SE and obtained the estimated states as $\hat{X}_{t,i} = \hat{V}_{t,i}\angle\hat{\Theta}_{t,i}$, $\forall i \in \mathcal{S}$.
3. For load bus $i \in \mathcal{S}$, set the voltage angle as $\Theta_{t,i}^{(a)} = \hat{\Theta}_{t,i} + C'_{t,i}$, and compute the post-attack complex voltage state as $X_{t,i}^{(a)} = \hat{V}_{t,i}\angle\hat{\Theta}_{t,i}^{(a)}$.

Algorithm 3 Convert DC Attack to AC Attack (Continued)

4. For non-load bus u in $\mathcal{S}$, the state should be updated to ensure the power flow balance on the bus as

$$\sum_{s\in\Omega_u} I_{us} = X_{t,u} \sum_{s\in\Omega_u} X_{t,s} \left(G_{us} + jB_{us}\right) \forall j \in \mathcal{S}. \tag{6.34}$$

   where $r_{us} + jx_{us}$ is the reactance of line connecting bus u and s, Ω_u is the set of buses connecting to bus u. Note that the attacker has already known $X_{t,s}^{(a)}$. That is, the attacker can solve (6.34) to update $X_{t,u}^{(a)}$.

5. Compute the post-attack measurement $\tilde{Z}_t$ as $\tilde{Z}_{t,i}^T = Z_{t,i}^T,\ \forall i \notin \mathcal{S}$, and $\tilde{Z}_{t,i}^T = H_i X_t^{(a)T},\ \forall i \in \mathcal{S}$.

6. Repeat 1–5 for $t = 1, ..., N$.

Note that, since Algorithm 2 is for DC attack design, when converting to a non-linear post-attack matrix, it is not guaranteed that such a post-attack measurement matrix can always bypass the LD detector. The performance of this attack model is demonstrated in Sec. 6.3.2.

6.3.2 Numerical Results

In this subsection, we illustrate the efficacy of the worst-case physical line overflow FDI attacks introduced in Sec. 6.3.1. To this end, we first solve the attack optimization problem in (6.14)–(6.22) with Algorithm 2 to find the optimal attack matrix C'^*. Subsequently, we construct the post-attack measurement matrix $\tilde{Z}^*$ with Algorithm 3. Finally, we solve the LD detection optimization problem (6.4)–(6.5) for $\tilde{Z}^*$ to check if the attack matrix C'^* is detected. Same as in Sec. 6.2.2, we assume that the LD detector selects 2 seconds worth of PMU measurements data, *i.e.,* $N = 60$, while the

attacker injects bad data. The test system is the IEEE 24-bus reliability test system (RTS). To model realistic power systems, we assume that there are congested lines prior to the attack and the attacker chooses one line in $\mathcal{L}$ as the target to maximize power flow. This is achieved in simulation by uniformly reducing all line ratings by 50%. The convex optimization problems for the LD detection and attack design are solved with MOSEK. In the LD detection optimization problem, the weight λ is chosen to be 1.05. The rules to place PMU and generate phasor measurements are the same as introduced in Sec. 6.2.2. The resulting synthetic measurements before attack are partially illustrated in Fig 6.5.

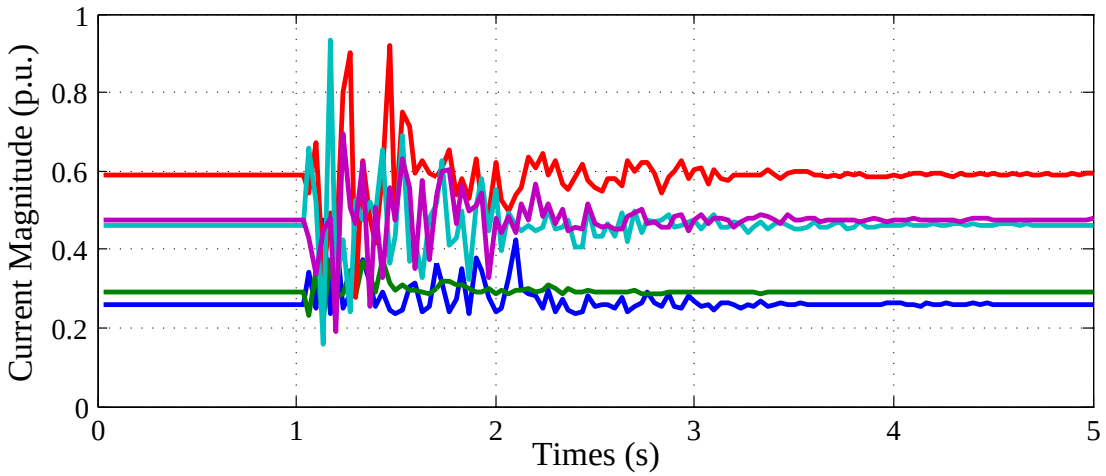


Figure 6.5: IEEE 24-Bus System 5 Out of 24 Current Measurement Magnitudes of The Synthetic PMU Data for Testing Worst-Case Attack Optimization Problem.

We choose the attacked states set as $\mathcal{I}_a = \{16, 17, 18, 21, 22\}$, the load shift factor τ as 10%, the convergence threshold ε as 6×10^{-4}, the maximum iteration number in Algorithm 2 as 100.

The converge behavior for attacks at t =1–3 seconds and t =3–5 seconds are demonstrated in Figs. 6.6(a) and (b), respectively. The upper and lower bounds variation for attacks at t =1–3 seconds and t =3–5 seconds are illustrated in Figs. 6.7(a) and (b), respectively.

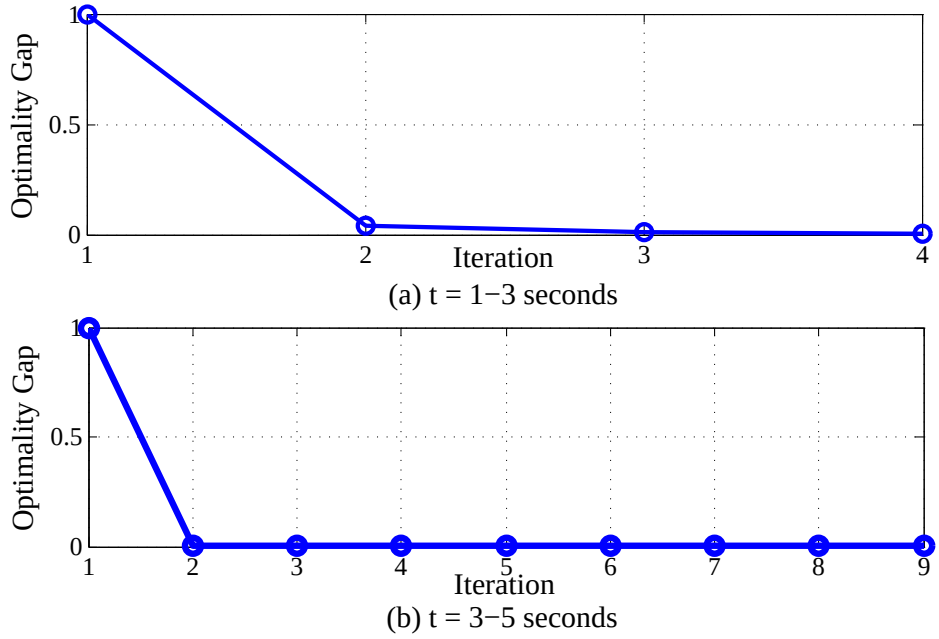


Figure 6.6: The Converge Behavior of Algorithm 2 for The Test Case.

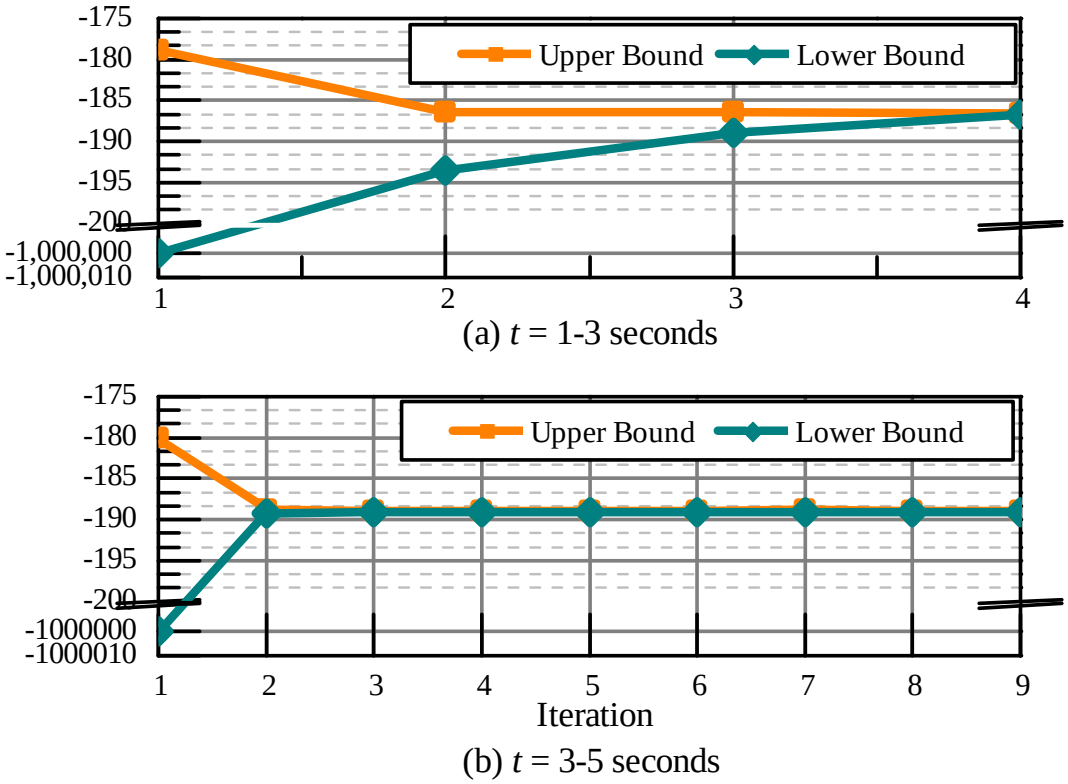
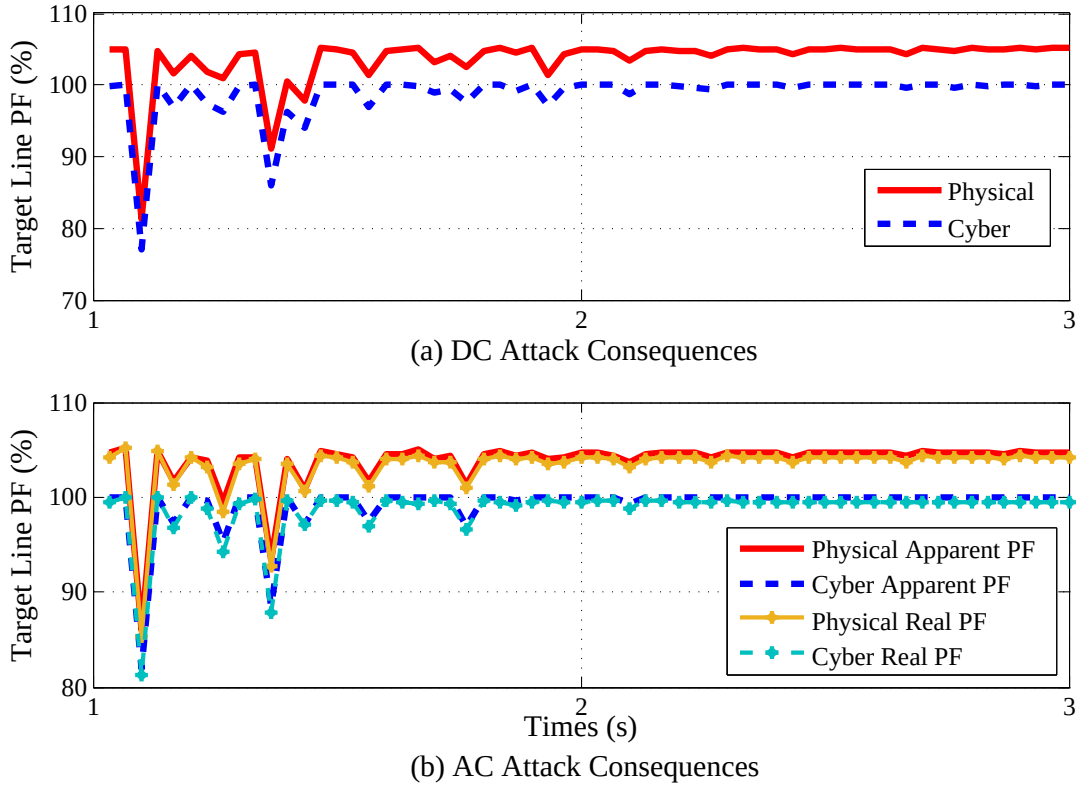


Figure 6.7: The Upper and Lower Bounds Variation of Algorithm 2 for The Test Case.

It can be observed from Figs. 6.6 and 6.7 that for both simulation time periods, Algorithm 2 can converge within 10 iterations. This result indicates the good convergence behavior of Algorithm 2 on the test case.

The post-attack cyber and physical power flow for t =1–3 seconds under both DC (from the bi-level attack optimization problem) and AC power flow models are demonstrated in Figs. 6.8(a) and 6.8(b), respectively. The results for t =3–5 seconds are illustrated in Fig. 6.9.



(a) DC Attack Consequences

(b) AC Attack Consequences

Figure 6.8: The Post-Attack Power Flow (PF) When The Target Line Is 28 of IEEE 24-Bus System for t =1–3 Seconds.

Figs. 6.8(a) and 6.9(a) demonstrate that the attack designed with the bi-level attack optimization problem can lead to unobservable physical target line overflow at 118 out of 200 time instances during $t = 1 - 5$ seconds. Figs. 6.8(b) and 6.9(b) illustrate that although the attack matrices are solved by linear optimization problems, they can still cause overflows in the non-linear system model.

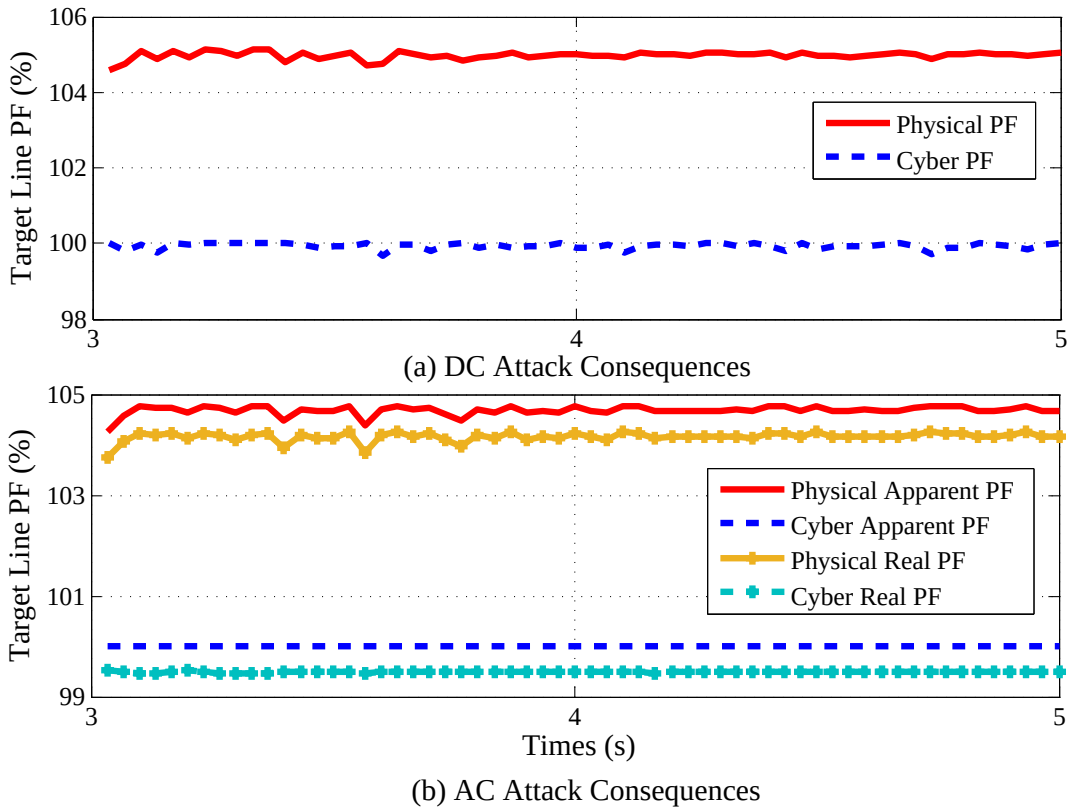


Figure 6.9: The Post-Attack Power Flow (PF) When The Target Line Is 28 of IEEE 24-Bus System for t =3–5 Seconds.

Table 6.3: Summary of The Nuclear Norm Results for Both DC and AC Attacks at t =1–3 Seconds and t =3–5 Seconds.

Time Period	DC Attack		AC Attack	
	$\|\Theta^* + C'^*\|$	$\|\Theta^*\|$	$\|\tilde{Z}^*\|$	$\|Z\|$
1–3 second	7.61	7.65	79.95	80.88
3–5 second	6.22	6.27	63.46	64.41

The post-attack measurement matrices $\tilde{Z}^*$ constructed with Algorithm 3 at both t =1–3 seconds and t =3–5 seconds completely bypass the LD detector, *i.e.*, $C^{*(\mathrm{LD})} = \mathbf{0}$. We summarize the detection results in Table. 6.3. From these results, it can be observed that (i) $\|\Theta^* + C'^*\|_* \leq \|\Theta\|_*$ holds for both DC attacks designed with

Algorithm 2; and (ii) $\|\tilde{Z}^*\|_* \leq \|Z\|_*$ holds for both AC attacks constructed with Algorithm 3.

6.4 Concluding Remarks

In this chapter, we have studied the vulnerability of phasor measurement units to FDI attacks. Prior work demonstrated that unobservable FDI attacks that can bypass traditional bad data detectors based on measurement residuals can be identified by the LD detector. In this work, we have shown that a more sophisticated attacker that understands the temporal correlation of PMU data can exploit it to design unobservable FDI attacks that cannot be detected by the LD detector. Moreover, we have illustrated that the attacker can further developed a single leader multiple followers problem to design a worst-case FDI attacks that can both bypass the LD detector and cause physical line overflow problems.

Chapter 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

The cyber layer of the electrical power system is vulnerable to FDI attacks. In this dissertation, the vulnerability and physical consequences of FDI attacks on SCADA and PMUs are analyzed. From the perspective of the attacker, two classes of limited information FDI attacks on SCADA measurements and two classes of FDI attacks on PMU data are introduced. For each class of attacks, the attack model is proposed; the attacker's knowledge is identified; the physical consequences of attacks are demonstrated via one or more IEEE test systems. From the perspective of the system control center, an off-line vulnerability analysis framework is proposed as the first step to thwart FDI attacks. The major conclusions are drawn as follows:

1. The physical consequences of FDI attacks desgined with perfect knowledge in an attack sub-network $\mathcal{L}$ and limited estimated information outside of the sub-network are analyzed. A bi-level optimization problem is formulated to maximize the power flow on a chosen target line with the attacker's perfect information in $\mathcal{L}$, as well as estimated information of marginal generators and PTDF outside of $\mathcal{L}$. It is illustrated that with an appropriately chosen sub-network and perfect localized information, the attacker can overload transmission lines with limited load shifts in both linear and non-linear models in the test system. The incomplete congestion knowledge, inaccurate external marginal generation, and inaccurate PTDF matrix in the external network $\mathcal{E}$ can undermine the attacker's evaluation of attack consequences. However, even designed with

inaccurate estimated knowledge, FDI attacks can still cause line overflows.

2. A class of FDI attacks designed with perfect information inside an attack sub-network $\mathcal{L}$ and no information outside of $\mathcal{L}$ is introduced. Pseudo-boundary injections are introduced to represent the power flows delivered from the external network. A multiple linear regression model is developed for the attacker to learn the relationship between pseudo-boundary injections and power injections inside $\mathcal{L}$. A bi-level optimization problem is formulated to maximize the power flow on a chosen target line with the attacker's perfect information in the attack sub-network, as well as the predicted pseudo-boundary injections. It is proved that the attacker can perfectly predict the pseudo-boundary injections with power injections in $\mathcal{L}$ under certain circumstances and can still compute the upper bounds on the attack consequences even with inaccurate predictions. Numerical results illustrate that the attacker can overload transmission lines with the proposed bi-level attack optimization problems. A sensitivity analysis on multiple scenarios of imperfect historical datasets shows the robustness of this attack strategy. It can be seen that even with a minuscule amount of system information and imperfect historical datasets, the attacker can still learn the system-wide re-dispatch behavior with a simple multiple linear regression model. In conclusion, one must be concerned that even with limited information, an attacker with access to historical data can take advantage of it to the detriment of reliable system operations.

3. For the two classes of limited information attacks, approximations including constant thermal line limits and static PTDF matrix are made to model ideal attacker and system. Although such approximations may undermine the attack consequences in actual power systems, the attack consequences still indicate

that such attacks are credible threats in power systems.

4. An off-line vulnerability analysis framework from the perspective of the control center is developed to analyze historical data, identify patterns of congested lines, assess the vulnerability of sub-networks surrounding each congested line layer by layer, and finally identify key sub-networks that are prone to FDI attacks. It is demonstrated that this framework is both accurate and efficient in assessing the vulnerability of the test system apriori and has the potential to be employed as an on-line analysis tool to identify real-time vulnerability. The identified key sub-networks indicate that the load redistribution on a small set of key buses play an essential role in causing target line overflow. In conclusion, how to identify key measurements to keep secure and to analyze load varying behaviors inside the identified key sub-networks is crucial to further protect the system from FDI attacks

5. The vulnerability of phasor measurement units to FDI attacks is analyzed. A convex optimization problem that ensures the low-rank post-attack measurements matrix is introduced to design the FDI attacks on PMU data that can bypass the LD detector. A bi-level single leader multiple followers attack optimization problem is proposed to design a worst-case FDI attack that can both bypass the LD detector and result in physical line overflow problems. It can be concluded that a more sophisticated attacker that understands the temporal correlation of PMU data can exploit it to design unobservable FDI attacks that can cause physical overflow on the system, while it cannot be detected by the LD detector. A better detection mechanism that can further identify FDI attacks that tract the temporal correlation of the PMU measurements are needed.

7.2 Future Work

To further study the implications of FDI attacks on power system operations, the following work is suggested for the future.

7.2.1 Generalization of the Attack Optimization Structure to Achieve Other Attack Consequences

In this dissertation, we focus on understanding the class of FDI attacks that can result in physical line overflow violations. However, other attack objectives can also be achieved with both the perfect and limited information bi-level attack optimization structures. The potential attack consequences of interest are listed as follows:

1. Maximize the total operation costs of the attack-induces re-dispatch. For perfect information attacks, this can be simply achieved by changing the objective in (3.1) as

$$\text{Maximize } \sum_{g=1}^{n_g} C_g \left(P_{G,g}\right). \tag{7.1}$$

   For limited information attacks, these attack consequences can be achieved by replacing (4.25) with

$$\text{Maximize } \sum_{g\in\mathcal{L}} C_g \left(\bar{P}_{G,g}\right). \tag{7.2}$$

2. Maximize the load shedding in the post-attack system re-dispatch. To achieve this goal, a new vector of load shedding variables (denoted as P_S) is introduced in the second level DC OPF problem. The perfect information bi-level attack optimization problem can be rewritten as:

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{n_b} P_{S,i} - \zeta \left\|c\right\|_0 \qquad (7.3)\\ &\text{subject to } (3.3)-(3.4) \end{aligned}$$

$$\{\theta^*, P_G^*\} = arg\left\{\min_{\theta, P_G} \sum_{g=1}^{n_g} C_g\left(P_{Gg}\right) + \sum_{i=1}^{n_b} C_{S,i} P_{S,i}\right\} \tag{7.4}$$

$$\text{subject to} \quad GP_G - H\theta = P_D - P_S \tag{7.5}$$

$$(3.7) - (3.8)$$

where $C_{S,i}$ is the cost of the shedding load at bus i. This attack optimization problem can be further modified as a limited information attack optimization problem.

3. Maximize the physical interface power flow to result in voltage collapse. In practice, the net active power flows on key transmission lines that connect one region and the other should be within a interface power flow limit to ensure the voltage stability [58]. In general, such a limit is lower than the sum of the thermal limits on these lines, since operating at the thermal limit may result in voltage collapse. The attack optimization problem introduced in Chapter 3 can be modified to maximize the physical interface power flow by replacing (3.1) with

$$\text{Maximize} \sum_{k \in \Omega_{int}} P_k \tag{7.6}$$

where Ω_{int} represents the set of key transmission lines connecting two regions.

7.2.2 Generalization of the Limited Information Attack Strategy to Cyber-Physical Attacks

In previous work [29], the author and her collaborator introduce a class of unobservable cyber-physical topology attacks. By implementing these attacks, the attacker can open a circuit breaker as well as change measurements to result in line overflow in an unobservable way. It is useful to evaluate the vulnerability of the limited in-

formation cyber-physical topology attacks that combines methods introduced in [29] and Chapter 4.

7.2.3 FDI Attacks Considering Real-time Contingency Analysis Module

The EMS model considered in this dissertation as well as in [26, 29] neglect the contingency analysis module. In practice, SE is followed by contingency analysis. Security constraints of lines and generators that have post-contingency violations are then included in the subsequent economic dispatch module. In fact, if such security constraints are not modeled in the bi-level attack optimization problem, it will undermine the attackers' evaluation of attack consequences. Therefore, consequences of line overflow FDI attacks modeled with security constraints should be evaluated. The attack optimization problem can be modified as

$$\text{maximize} \quad P_l - \zeta \left\| c \right\|_0 \tag{7.7}$$

$$\text{subject to} \quad (3.21) - (3.23) \tag{7.8}$$

$$\{P_G^*\} = \arg \left\{ \sum_{g=1}^{n_g} C_g \left(P_{Gg} \right) \right\} \tag{7.9}$$

$$\text{subject to} \quad (3.6), (3.8) \tag{7.10}$$

$$P^a = K(GP_G - P_D + Hc) \tag{7.11}$$

$$-P_{\max} \leq P^a \leq P_{\max} \tag{7.12}$$

$$-P_{k,\max} \leq P_k^a + L_{k,o} P_o^a \leq P_{k,\max} \quad k, o \in \{1, 2, ..., n_{br}\}, \quad k \neq o \tag{7.13}$$

where P^a is the post-attack power flow vector in the cyber layer, $L_{k,o}$ is the line outage distribution factor which represents the change of power flow in the line k post to the outage of line o. Same as the bi-level attack optimization structure introduced in Chapter 3, the first and second levels model the attacker and system response to the attack, respectively. In particular, besides DC OPF, $N-1$ line outage constraints

are formulated with (7.13) in the second level.

A related problem is to understand the feasibility and consequences of FDI attacks that can mask contingencies. Assume that line outage contingencies exist before attacks. These contingencies can be masked by FDI attacks designed with the following optimization problem:

$$\underset{c}{\text{minimize}} \quad \|c\|_0 \tag{7.14}$$

$$\text{subject to} \quad \|c\|_0 \leq N_0 \tag{7.15}$$

$$-\tau P_D \leq Hc \leq \tau P_D \tag{7.16}$$

$$P^a = K(GP_G - P_D + Hc) \tag{7.17}$$

$$-P_{\max} \leq P^a \leq P_{\max} \tag{7.18}$$

$$-P_{k,\max} \leq P_k^a + L_{k,o}P_o^a \leq P_{k,\max} \quad k, o \in \{1, 2, ..., n_{br}\}, \quad k \neq o \tag{7.19}$$

where the objective is to minimize the size of the attack subgraph.

7.2.4 Design of the Detection Mechanisms

Besides analyzing the vulnerability of other classes of FDI attacks, another avenue is to design detection mechanisms to identify anomalies caused by FDI attacks.

It can be observed that all classes of unobservable FDI attacks studied in this dissertation lead to an inevitable load redistribution inside a subgraph. Although the net loads in both the subgraph and the entire system remain unchanged, the load varying behavior at each bus may be inconsistent with that observed in the historical data. For example, after analyzing the historical data from 12:00 p.m. to 2:00 p.m. during the summer, operators found that at this time period, the loads at two adjacent buses vary at the same direction over 95% of times. When the loads at the two buses are observed to vary at different directions during this time period, one can suspect that there are FDI attacks inside the system. In Chapter 5, it is also

illustrated that the load redistribution at some specific buses play an essential role in causing target line overflow. If the buses with abnormal load varying behaviors also fall into this set of essential buses, one can conclude the existence of FDI attacks.

In addition, FDI attacks may result in anomalies on congested line power flow in the cyber layer. It can be seen from the simulation results in Chapters 4–6 that to cause a physical power flow violation, the attacker has to first reduce the cyber power flow on the target congested line with FDI attacks. In the following system re-dispatch, power outputs from cheaper generators will increase due to the relaxation of the thermal limit constraint on the target line. This in turn, result in physical overflow violation in the target line. Such anomalies can supplement the load varying anomaly detection to better diagnose FDI attacks.

REFERENCES

[1] A. G. Phadke, J. S. Thorp, and K. J. Karimi, "State estimlatjon with phasor measurements," *IEEE Transactions on Power Systems*, vol. 1, pp. 233–238, Feb 1986. 1.1

[2] K. Sun, S. Likhate, V. Vittal, V. S. Kolluri, and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and decision trees," *IEEE Transactions on Power Systems*, vol. 22, pp. 1935–1943, Nov 2007. 1.1

[3] Y. V. Makarov, P. Du, S. Lu, T. B. Nguyen, X. Guo, J. W. Burns, J. F. Gronquist, and M. A. Pai, "PMU-based wide-area security assessment: Concept, method, and implementation," *IEEE Transactions on Smart Grid*, vol. 3, pp. 1325–1332, Sept 2012. 1.1

[4] J. D. L. Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Transactions on Smart Grid*, vol. 1, pp. 20–27, June 2010. 1.1

[5] Y. Zhang, P. Markham, T. Xia, L. Chen, Y. Ye, Z. Wu, Z. Yuan, L. Wang, J. Bank, J. Burgett, R. W. Conners, and Y. Liu, "Wide-area frequency monitoring network (FNET) architecture and applications," *IEEE Transactions on Smart Grid*, vol. 1, pp. 159–167, Sept 2010. 1.1

[6] M. Zeller, "Myth or reality – does the Aurora vulnerability pose a risk to my generator?," in *64th Annual Conference for Protective Relay Engineers*, pp. 130–136, April 2011. 1.1

[7] "The Stuxnet worm: A cyber-missile aimed at Iran," tech. rep., The Economist, 24 September 2010. 1.1

[8] T. Espiner, "Siemens: Stuxnet infected 14 industrial plants." http://www.zdnet.com/article/siemens-stuxnet-infected-14-industrial-plants/, September 2010. 1.1

[9] S. Kelly, "Homeland security cites sharp rise in cyber attacks." http://security.blogs.cnn.com/2012/07/04/homeland-security-cites-sharp-rise-in-cyber-attacks/, July 2012. 1.1

[10] S. Toppa, "The national power grid is under almost continuous attack, report says." http://time.com/3757513/electricity-power-grid-attack-energy-security/, March 2015. 1.1

[11] K. Zetter, "Inside the cunning, unprecedented hack of Ukraine's power grid." http://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/, March 2016. 1.1

[12] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM Conference on Computer and Communications Security*, CCS '09, (Chicago, Illinois, USA), pp. 21–32, 2009. 1.2, 6, 6.1.2

[13] A. T. H. Sandberg and K. H. Johansson, "On security indices for state estimators in power networks," in *1st workshop secure control system*, 2010. 1.2

[14] A. Teixeira, S. Amin, H. Sandberg, K. Johansson, and S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 5991–5998, Dec 2010. 1.2, 6, 6.1.2

[15] G. Dan and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 214–219, Oct 2010. 1.2

[16] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "On malicious data attacks on power system state estimation," in *Universities Power Engineering Conference (UPEC), 2010 45th International*, pp. 1–6, 2010. 1.2

[17] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012. 1.2, 3.2, 3.2

[18] J. Liang, O. Kosut, and L. Sankar, "Cyber-attacks on AC state estimation: Unobservability and physical consequences," in *IEEE PES General Meeting*, (Washington, DC), July 2014. 1.2

[19] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011. 1.2

[20] J. Kim and L. Tong, "On topology attack of a smart grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, (Washington, DC), pp. 1–6, February 2013. 1.2

[21] M. Rahman, E. Al-Shaer, and R. Kavasseri, "Impact analysis of topology poisoning attacks on economic operation of the smart power grid," in *Distributed Computing Systems (ICDCS), 2014 IEEE 34th International Conference on*, pp. 649–659, June 2014. 1.2

[22] A. Ashok and M. Govindarasu, "Cyber attacks on power system state estimation through topology errors," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, July 2012. 1.2

[23] J. Zhang and L. Sankar, "Implementation of unobservable state-preserving topology attacks," in *North American Power Symposium (NAPS), 2015*, pp. 1–6, Oct 2015. 1.2

[24] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *Smart Grid, IEEE Transactions on*, vol. 2, pp. 382–390, June 2011. 1.2

[25] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, pp. 1731–1738, Sept 2012. 1.2

[26] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Transactions on Power Systems*, vol. 31, pp. 3864–3872, Sept 2016. 1.2, 1.3, 3, 3.1, 3.2, 3.2, 4.1.1, 4.1.1, 4.1.3, 4.2, 4.2.2, 4.2.3, 6, 6.1.2, 6.2.2, 6.3.1, 7.2.3

[27] Z. Chu, J. Zhang, O. Kosut, and L. Sankar, "Evaluating power system vulnerability to false data injection attacks via scalable optimization," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 260–265, Nov 2016. 1.2, 4.2.2, 5.1, 6, 6.1.2

[28] Z. Chu, J. Zhang, O. Kosut, and L. Sankar, "Vulnerability assessment of power systems to false data injection attacks," *IEEE Transactions on Smart Grid*, 2017. under review. 1.2, 5.1, 6.3.1

[29] J. Zhang and L. Sankar, "Physical system consequences of unobservable state-and-topology cyber-physical attacks," *IEEE Transactions on Smart Grid*, vol. 7, pp. 2016–2025, July 2016. 1.2, 3.2, 4.2.2, 6, 6.1.2, 6.2.2, 7.2.2, 7.2.3

[30] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 3153–3158, Dec 2012. 1.2, 1.3

[31] X. Liu and Z. Li, "Local load redistribution attacks in power systems with incomplete network information," *IEEE Transactions on Smart Grid*, vol. 5, pp. 1665–1676, July 2014. 1.2, 1.3

[32] X. Liu and Z. Li, "False data attacks against AC state estimation with incomplete network information," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–10, 2016. 1.2, 1.3

[33] H. Lin, Y. Deng, S. Shukla, J. Thorp, and L. Mili, "Cyber security impacts on all-PMU state estimator - a case study on co-simulation platform GECO," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pp. 587–592, Nov 2012. 1.2

[34] C. Beasley, G. K. Venayagamoorthy, and R. Brooks, "Cyber security evaluation of synchrophasors in a power system," in *2014 Clemson University Power Systems Conference*, pp. 1–5, March 2014. 1.2

[35] J. Kim and L. Tong, "On phasor measurement unit placement against state and topology attacks," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 396–401, Oct 2013. 1.2

[36] H. Sedghi and E. Jonckheere, "Statistical structure learning of smart grid for detection of false data injection," in *2013 IEEE Power Energy Society General Meeting*, pp. 1–5, July 2013. 1.2

[37] D. Lee and D. Kundur, "Cyber attack detection in PMU measurements via the expectation-maximization algorithm," in *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 223–227, Dec 2014. 1.2

[38] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Transactions on Smart Grid*, vol. 5, pp. 612–621, March 2014. 1.2, 1.3, 6

[39] M. Wang, P. Gao, S. G. Ghiocel, J. H. Chow, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, "Identification of "unobservable" cyber data attacks on power grids," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 830–835, Nov 2014. 1.2, 1.3, 6, 6.1.3, 6.2.2

[40] P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, "Identification of successive "unobservable " cyber data attacks in power systems through matrix decomposition," *IEEE Transactions on Signal Processing*, vol. 64, pp. 5557–5570, Nov 2016. 1.2, 1.3, 6

[41] H. Xu, C. Caramanis, and S. Sanghavi, "Robust PCA via outlier pursuit," *IEEE Transactions on Information Theory*, vol. 58, pp. 3047–3064, May 2012. 1.2, 6

[42] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York: CRC Press, 2004. 2.1, 2.1, 6.1.1

[43] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004. 2

[44] L. Jia, R. Thomas, and L. Tong, "On the nonlinearity effects on malicious data attack on power system," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, July 2012. 3.2

[45] J. Zhang, Z. Chu, L. Sankar, and O. Kosut, "False data injection attacks on power system state estimation with limited information," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, July 2016. 3.2, 4.2.2, 6, 6.1.2

[46] "PSERC final project report: Electricity market structures to reduce seams and enhance investment," February 2010. 4.2

[47] M. Farrokhseresht, N. Farrokhseresht, and M. R. Hesamzadeh, "Optimizing the net transfer capacity of a cross-border interconnector by reactive power planning," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, July 2016. 4.2

[48] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*. Wiley, 5th ed., 2012. 4.2.1

[49] "IEEE 118-bus, 54-unit, 24-hour system, unit and network data." 4.2.3

[50] N. Dahal, R. L. King, and V. Madani, "Online dimension reduction of synchrophasor data," in *PES T&D 2012*, pp. 1–7, May 2012. 6

[51] Y. Chen, L. Xie, and P. R. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *2013 IEEE Power Energy Society General Meeting*, pp. 1–5, July 2013. 6

[52] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011. 6, 6.1.2

[53] J. Zhang, Z. Chu, L. Sankar, and O. Kosut, "Can attackers with limited information exploit historical data to mount successful false data injection attacks on power systems?," *IEEE Transactions on Power Systems*, 2017. submitted. 6

[54] A. G. Phadke, J. S. Thorp, R. F. Nuqui, and M. Zhou, "Recent developments in state estimation with phasor measurements," in *2009 IEEE/PES Power Systems Conference and Exposition*, pp. 1–7, March 2009. 6.1.1, 6.1.1

[55] M. Wang, J. H. Chow, P. Gao, X. T. Jiang, Y. Xia, S. G. Ghiocel, B. Fardanesh, G. Stefopolous, Y. Kokai, N. Saito, and M. Razanousky, "A low-rank matrix approach for the analysis of large amounts of power system synchrophasor data," in *2015 48th Hawaii International Conference on System Sciences*, pp. 2637–2644, Jan 2015. 6.1.1

[56] B. Xu and A. Abur, "Observability analysis and measurement placement for systems with PMUs," in *IEEE PES Power Systems Conference and Exposition, 2004.*, pp. 943–946 vol.2, Oct 2004. 6.2.2

[57] N. V. Sahinidis and I. E. Grossmann, "Convergence properties of generalized Benders' decomposition," *Computers and Chemical Engineering*, vol. 15, p. 481, 1991. 6.3.1

[58] B. Lee, H. Song, S.-H. Kwon, G. Jang, J.-H. Kim, and V. Ajjarapu, "A study on determination of interface flow limits in the kepco system using modified continuation power flow (mcpf)," *IEEE Transactions on Power Systems*, vol. 17, pp. 557–564, Aug 2002. 3