Hindawi
Security and Communication Networks
Volume 2017, Article ID 9602357, 12 pages
https://doi.org/10.1155/2017/9602357

Research Article

F-DDIA: A Framework for Detecting Data Injection Attacks in Nonlinear Cyber-Physical Systems

Jingxuan Wang,[1] Lucas C. K. Hui,[1] S. M. Yiu,[1] Gang Zhou,[2] and Ruoqing Zhang[1]

[1]*Department of Computer Science, University of Hong Kong, Pokfulam Road, Hong Kong*
[2]*Peking University, Beijing, China*

Correspondence should be addressed to Jingxuan Wang; hongkongwangjingxuan@gmail.com

Received 10 April 2017; Accepted 7 June 2017; Published 10 August 2017

Academic Editor: Leo Y. Zhang

Data injection attacks in a cyber-physical system aim at manipulating a number of measurements to alter the estimated real-time system states. Many researchers recently focus on how to detect such attacks. However, most of the detection methods do not work well for the nonlinear systems. In this paper, we present a compressive sampling methodology to identify the attack, which allows determining how many and which measurement signals are launched. The sparsity feature is used. Generally, our methodology can be applied to both linear and nonlinear systems. The experimental testing, which includes realistic load patterns from NYISO with various attack scenarios in the IEEE 14-bus system, confirms that our detector performs remarkably well.

1. Introduction

A cyber-physical system (CPS) is a dynamical system, which integrates the computational components (i.e., real-time operations) with its physical components (i.e., hardware facilities). Examples of CPS can be large-scale distributed systems, such as smart grid, transportation networks, railway control system, and medical monitoring. The design of CPS involves various of disciplines, such as control engineering, software engineering, and mechanics and networks. Particularly, control engineering is a communication network for transmitting sensor data (measurements) so that the system operator can in real-time monitor the production process. Among the control disciplines, a scheme called bad data detector (BDD) is applied to detect whether there exists a disruption of sensor data caused by the genetic malfunction or malicious attacks. The classical BDD technique is to utilize the "residual principle," which calculates the difference between the observed readings and the computed readings based on the estimated system states. When an attack is injected into the system, BDD will remove those readings (collected from the sensors), of which residuals are larger than a threshold.

As the increased vulnerabilities proposed by the recent discoveries of system malware, concerns about the security of CPS are arising. In 2011, a malware, known as Stuxnet [1], successfully penetrated the networks of Iran's uranium enrichment infrastructure via programmable logic controllers. From this instance, we can see that it is possible for an attacker to introduce errors on physical readings. Inspired by this attacking strategy, a class of attacks named *data injection attacks* are proposed in recent years, which can affect the system control algorithms and thus lead to abnormal operations [2, 3]. Hence, sufficient attention should be paid to the detection techniques against this attack, which is easy to be implemented by strong adversaries who are quite knowledgeable about the targeted systems.

To fight against this attack, existing works focus on the detection of data injection attacks and the protection of nonlinear measurements [4, 5]. Detectors utilizing the sparsity and low rank of the system topology are proposed in [6–8]. Greedy and game theory methods have been used for optimizing the placement of devices [9], to lower the possibility of the construction of data injection attacks. Applying the machine learning techniques to conduct the classification is proposed in [10]. They propose a "first difference aware" machine learning (FDML) classifier to detect the cyber attacks. A graph theory-based algorithm is proposed in [11] to determine which measurement signals an attacker will alter. However, we notice that all detection models except [11, 12]

are conducted in a constrained setting, by assuming that the functions from system states to measurements are linear. This assumption is too stringent to fit for some nonlinear systems, for example, alternative current (AC) model in power grids.

This paper investigates an alternative approach to detect data injection attacks in the nonlinear system. We propose a detector framework named F-DDIA to reconstruct the initial states of the plant from the corrupted observations, which formulates an error correction problem. In particular, we notice that, due to the property of data injection attacks, only a small fraction of the observations are supposed to be attacked at a given time instance. Thus, we formulate the error correction problem as a sparse optimization problem which can be solved with the general ℓ_1-minimization program technique. In this paper, we apply Douglas-Rachford techniques [13] among minimization techniques. Furthermore, we employ the "divide-and-conquer" principle to construct a compressive sensing model of a linear subspace, which is interesting in the general mathematical settings.

To validate and illustrate our algorithm, we use real-world CPS power grids as a case study. In particular, we use the data injection attacks model proposed in [2], where the attacks are directed by injecting false data into the sensors. Simulations based on IEEE 14-bus test systems validate the effectiveness of our methodology. The results show that the proposed algorithm can efficiently identify the data injection attacks (i.e., with high precision and recall values) and recover the initial system states (i.e., with small average phase error).

The rest of this paper is organized as follows. Section 2 presents the system model in a nonlinear system, including preliminaries related to a broad class of attacks. Section 3 states the problem and derives a theoretical justification of the efficacy of the security algorithm in a general cyber-physical system model. Section 4 analyzes the performance of the proposed approach through simulations. Section 5 gives concluding remarks.

2. Preliminaries

2.1. System Model and Bad Data Detector. A cyber-physical system is usually described by the following widely adopted discrete-time nonlinear dynamical model:

$$x[k+1] = \delta(x[k]) + Bu[k] + w[k], \tag{1a}$$

$$z[k] = h(x[k]) + e[k], \tag{1b}$$

where at time $k \in \mathcal{J} \triangleq \{0, \ldots, T-1\}$: $x[k] \in \mathbb{R}^n$ is the system state; $u[k] \in \mathbb{R}^l$ is the bounded input vector; $a[k] \in \mathbb{R}^m$ is the measurement vector (data collected by the sensors); $w[k]$ denotes the state noise (i.e., Gaussian with known statistics); and $v[k]$ denotes measurement errors. Here the matrix B is a constant matrix, $\delta : \mathbb{R}^n \to \mathbb{R}^n$ denotes the state transition function and $h : \mathbb{R}^n \to \mathbb{R}^m$ denotes the topology of the system, which are the nonlinear functions with respect to the states. The process of estimating system states from the measurements is called *state estimation*.

In traditional weighted least squares (WLS) state estimation, the system states are valid only if the measurement residual vector $r[k]$ is less than a threshold [14],

$$r[k] = \|z[k] - h(\hat{x}[k])\|_{\ell_2}, \tag{2}$$

where $\hat{x}[k]$ is the estimated system state after the process of state estimation. Specifically, the presence of bad measurements is inferred if $Jr[k] > \tau$, where τ is a chosen identification threshold. Upon detection of bad data, two kinds of methods, named the largest normalized residual test ($r_N^{\max}$) and hypothesis testing identification (HTI) method, are widely used to identify whether the measurements contain bad data.

2.2. Data Injection Attack. Data injection attacks are commonly known as *false data injection attacks* [2], *data framing attacks* [3, 15], in the sense of the following definition.

Definition 1. A vector $a[k]$ is called a (κ, m)-data injection attack if there exists an index set $i \in \mathcal{A}$, where $\mathcal{A}$ is the set of manipulated measurements and $\mathcal{A} \subset \mathcal{P} \triangleq \{1, \ldots, m\}$, such that

(i) $\|a[k]\|_{\ell_0} \leq \kappa$;

(ii) $a_i[k] = 0, \ \forall i \in \mathcal{P} \setminus \mathcal{A}$;

(iii) $a_i[k] \neq 0, \ \forall i \in \mathcal{A}$.

To implement this class of attack, it requires the attacker to have the knowledge of either the measurements information (z) or the topology configuration ($h(\cdot)$). Specifically, data injection attack can be written in the form of

$$\overline{z}[k] = z[k] + a[k] = h(x[k]) + a[k], \tag{3}$$

where $a[k]$ is the injected false measurement data. There are many ways to generate this type of attacks. For example, if $h(\cdot)$ is available to the attacker, the attack a can be constructed in the following form (namely, false data injection attack in a linear system):

$$a = Hc, \tag{4}$$

where c is the error injected on the system state and $H = \partial h(x)/\partial x$ is the Jacobian matrix. However, to implement this attack, the attacker needs to take control of at least κ sensors, where $\kappa \leq m$.

2.3. Measurement Dynamics. We can use the polynomial regression approach to fit the measurement dynamics,

$$z[k+1] = \delta(x[k]) + Bu[k] + w[k] = \delta'(z[k]), \tag{5}$$

where $\delta' : \mathbb{R}^m \to \mathbb{R}^m$ denotes the dynamics of the measurements. Furthermore, we define $z_i[k]$ as the ith corrupted measurement at time k. That is, a polynomial regression model, which expresses the dynamics of the ith measurement can be given as follows:

$$\delta_i'(z_i[k]) = \gamma_{i,1}(z_i[k])^l + \cdots + \gamma_{i,l}(z_i[k]) + \gamma_{i,l+1}, \tag{6}$$

where l is called the degree of the polynomial and $i \in \mathscr{P}$. We denote $\gamma_i = (\gamma_{i,1}, \ldots, \gamma_{i,l+1}) \in \mathbb{R}^{l+1}$. As $\delta_i'(z_i[k])$ can be expressed in matrix form in terms of a response vector $z_i[k]$ and a parameter vector $\gamma_{i,j}$, where $1 \le j \le l+1$, we can rewrite $z_i[k+1]$ as a system of linear equations:

$$z_i[k+1] = X \begin{pmatrix} \gamma_{i,1} \\ \vdots \\ \gamma_{i,l+1} \end{pmatrix}, \tag{7}$$

where $X = ((z_i[k])^l \quad \cdots \quad z_i[k] \quad 1) \in \mathbb{R}^{l+1}$. Thus, the dynamical matrix γ can be estimated as

$$\widehat{\gamma}_i = \left(X^T X\right)^{-1} X^T z_i[k+1] \quad (i \in \mathscr{P}). \tag{8}$$

3. Our Methodologies

In this section, we formulate the detection problem as an error correction problem. We will further describe and explain why we can use ℓ_1-norm minimization technique (including Douglas-Rachford) to solve the detection problem.

3.1. Sparse Optimization Problem Formulation. In this paper, we consider the scenario that an attacker is limited to the resources of κ sensors and possesses the knowledge of system topology h, as well as the historical measurements $\overline{Z} = (\overline{z}[0]; \ldots; \overline{z}[T-1]) \in \mathbb{R}^{mT}$. Denote $Z = (z[0]; \ldots; z[T-1]) \in \mathbb{R}^{mT}$ as the initial measurements (without attacks) in time base. The obtained temporal observations $\overline{Z}$ can be expressed as

$$\overline{Z} = Z + \mathbb{A}, \tag{9}$$

where $\mathbb{A} = (a[0]; \ldots; a[T-1]) \in \mathbb{R}^{mT}$. Remark that, due to the property of data injection attacks, only a small fraction of the observations are supposed to be attacked at a given time instance. Hence, noticing the sparsity of vector $\mathbb{A}$, the detection problem can be converted to

$$\begin{aligned} \underset{\mathbb{A}}{\text{minimize}} \quad & \|\mathbb{A}\|_{\ell_0} \\ \text{subject to} \quad & \overline{Z} = Z + \mathbb{A}, \\ & \|a[k]\|_{\ell_0} \le \kappa, \ k \in \mathscr{I}, \end{aligned} \tag{10}$$

where κ is the maximum number of the meters that can be compromised. Under certain conditions which are explained above, we will focus on the problem of recovering the sparse vector $\mathbb{A}$ from $\overline{Z}$. And we denote the optimal solution of problem (10) as $\mathbb{A}^*$.

3.2. Subproblem Formulation. In the rest of this paper, we define the matrices $\mathbb{A} = [\mathbb{A}_1, \ldots, \mathbb{A}_{Tm+m}]$, $Z = [Z_1, \ldots, Z_{Tm+m}]$, and $\overline{Z} = [\overline{Z}_1, \ldots, \overline{Z}_{Tm+m}]$. We further define the matrices E, W, and $\overline{W}$ in the following forms:

$$\begin{aligned} E &= \begin{bmatrix} E_1^T \\ \vdots \\ E_m^T \end{bmatrix} = \begin{bmatrix} \mathbb{A}_1 & \mathbb{A}_{m+1} & \cdots & \mathbb{A}_{Tm+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{A}_m & \mathbb{A}_{2m} & \cdots & \mathbb{A}_{Tm+m} \end{bmatrix} \in \mathbb{R}^{m \times T}; \\ W &= \begin{bmatrix} W_1^T \\ \vdots \\ W_m^T \end{bmatrix} = \begin{bmatrix} Z_1 & Z_{m+1} & \cdots & Z_{Tm+1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_m & Z_{2m} & \cdots & Z_{Tm+m} \end{bmatrix} \in \mathbb{R}^{m \times T}; \\ \overline{W} &= \begin{bmatrix} \overline{W}_1^T \\ \vdots \\ \overline{W}_m^T \end{bmatrix} = \begin{bmatrix} \overline{Z}_1 & \overline{Z}_{m+1} & \cdots & \overline{Z}_{Tm+1} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{Z}_m & \overline{Z}_{2m} & \cdots & \overline{Z}_{Tm+m} \end{bmatrix} \in \mathbb{R}^{m \times T}. \end{aligned} \tag{11}$$

We can further obtain the following formulation among $E_i \in \mathbb{R}^T$, $W_i \in \mathbb{R}^T$, and $\overline{W}_i \in \mathbb{R}^T$:

$$\begin{aligned} \overline{W}_i &= W_i + E_i \quad (i \in \mathscr{P}), \\ \|\mathbb{A}\|_{\ell_0} &= \sum_{j=1}^{mT} \left\|\mathbb{A}_j\right\|_{\ell_0} = \sum_{i=1}^{m} \left\|E_i\right\|_{\ell_0} = \|E\|_{\ell_0}. \end{aligned} \tag{12}$$

We denote by $\mathrm{col}_{k \in \mathscr{I}}(E) \in \mathbb{R}^m$ the columns of the matrix E. Hence, problem (10) is equivalent to

$$\begin{aligned} \underset{E}{\text{minimize}} \quad & \|E\|_{\ell_0} \\ \text{subject to} \quad & \overline{W} = W + E, \\ & \left\|\mathrm{col}_k(E)\right\|_{\ell_0} \le \kappa, \ k \in \mathscr{I}. \end{aligned} \tag{13}$$

Note that $\|E\|_{\ell_0} = \sum_{i=1}^{m} \|E_i\|_{\ell_0}$; we can further solve problem (13) by seeking for the locally optimal choice for each E_i^* with the hope of finding a globally optimal solution (E^*):

$$\begin{aligned} \underset{E_i}{\text{minimize}} \quad & \left\|E_i\right\|_{\ell_0} \\ \text{subject to} \quad & \overline{W}_i = W_i + E_i, \\ & (i \in \mathscr{P}). \end{aligned} \tag{14}$$

The solution of this subproblem (14) will be given in Section 3.4. After solving m above optimization problems, the optimal solution E^* will be checked by the following constraints:

$$f\left(\mathrm{col}_k\left(E^*\right)\right) = \mathrm{sgn}\left(\kappa - \left\|\mathrm{col}_k\left(E^*\right)\right\|_{\ell_0}\right). \tag{15}$$

For any $k \in \mathscr{I}$, if $f(\mathrm{col}_k(E^*)) = 1$, there exists the attack; otherwise, there does not exist any data injection attack.

3.3. Solving Subproblem by ℓ_1-Minimization. Recall that the dynamical coefficients $(\gamma_1, \ldots, \gamma_m)$ are obtained (by polynomially fitting in Section 2.3). In view of adversary, $\overline{W}_i$ can be

rewritten as

$$\begin{aligned}\overline{W}_i &= \begin{pmatrix} \overline{z}_i[0] \\ \overline{z}_i[1] \\ \overline{z}_i[2] \\ \vdots \\ \overline{z}_i[T-1] \end{pmatrix} \\ &= \begin{pmatrix} z_i[0] + a_i[0] \\ z_i[1] + a_i[1] \\ z_i[2] + a_i[2] \\ \vdots \\ z_i[T-1] + a_i[T-1] \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ \gamma_{i,l} \\ (\gamma_{i,l})^2 \\ \vdots \\ (\gamma_{i,l})^{T-1} \end{pmatrix} z_i[0] \\ &\quad + \begin{bmatrix} 1 & & & \\ \gamma_{i,l} & 1 & & \\ (\gamma_{i,l})^2 & \gamma_{i,l} & 1 & \\ \vdots & \vdots & & \ddots \\ (\gamma_{i,l})^{T-1} & \cdots & \gamma_{i,l} & 1 \end{bmatrix} E_i \\ &\quad + \begin{pmatrix} 0 \\ g_i(\overline{z}[0]) \\ g_i(\overline{z}[1]) + \gamma_{i,l} g_i(z[0]) \\ \vdots \\ g_i(\overline{z}[T-1]) + \gamma_{i,l} g_i(z[T-2]) + \cdots \end{pmatrix}. \end{aligned} \tag{16}$$

Then we use the notation $\widetilde{W}_i$ as follows:

$$\begin{aligned}\widetilde{W}_i &= \overline{W}_i \\ &\quad - \begin{pmatrix} 0 \\ g_i(\overline{z}[0]) \\ g_i(\overline{z}[1]) + \gamma_{i,l} g_i(z[0]) \\ \vdots \\ g_i(\overline{z}[T-1]) + \gamma_{i,l} g_i(z[T-2]) + \cdots \end{pmatrix} \\ &= \Gamma_i z_i[0] + \Psi_i E_i, \end{aligned} \tag{17}$$

where the matrices $\Gamma_i \in \mathbb{R}^T$ and $\Psi_i \in \mathbb{R}^{T \times T}$ are

$$\begin{aligned}\Gamma_i &= \begin{pmatrix} 1 \\ \gamma_{i,l} \\ (\gamma_{i,l})^2 \\ \vdots \\ (\gamma_{i,l})^{T-1} \end{pmatrix}, \\ \Psi_i &= \begin{bmatrix} 1 & & & \\ \gamma_{i,l} & 1 & & \\ (\gamma_{i,l})^2 & \gamma_{i,l} & & \\ \vdots & \vdots & \ddots & \\ (\gamma_{i,l})^{T-1} & \cdots & & 1 \end{bmatrix}. \end{aligned} \tag{18}$$

In this paper, We have an approximation $g_i(z[k]) \doteq g_i(\overline{z}[k])$. The reason we take this approximation is that the difference of $z[k]$ and $\overline{z}[k]$ is

$$\begin{aligned} g_i(\overline{z}[k]) - g_i(z[k]) &= \gamma_{i,1} \overline{z}_i[k]^l + \cdots + \gamma_{i,l-1} \overline{z}_i[k]^2 \\ &\quad - \gamma_{i,1} z_i[k]^l - \cdots \\ &\quad - \gamma_{i,l-1} z_i[k]^2. \end{aligned} \tag{19}$$

For example, $g_i(\overline{z}[k]) - g_i(z[k]) = \gamma_{i,1} a_i[k](1 + a_i[k])$ when $l = 2$. Since the values of $\gamma_{i,1}$ are small ($i = 1, \ldots, m$), $g_i(z[k]) \doteq g_i(\overline{z}[k])$. We have done experiments about this fact, and the experimental result supports our approximation claim. Then, $\widetilde{W}_i$ in (17) can be updated as

$$\begin{aligned}\widetilde{W}_i &\doteq \overline{W}_i \\ &\quad - \begin{pmatrix} 0 \\ g_i(\overline{z}[0]) \\ g_i(\overline{z}[1]) + \gamma_{i,l} g_i(\overline{z}[0]) \\ \vdots \\ g_i(\overline{y}[T-1]) + \gamma_{i,l} g_i(\overline{z}[T-2]) + \cdots \end{pmatrix}. \end{aligned} \tag{20}$$

We can further take the QR decomposition of $\Gamma_i \in \mathbb{R}^T$ [16]:

$$\Gamma_i = S \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = \begin{bmatrix} S_{i1} & S_{i2} \end{bmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \tag{21}$$

where $S \in \mathbb{R}^{T \times T}$, $S_{i1} \in \mathbb{R}^T$, $S_{i2} \in \mathbb{R}^{T \times (T-1)}$, $R_{i1} \in \mathbb{R}^1$, and $\begin{bmatrix} S_{i1} & S_{i2} \end{bmatrix}$ is orthogonal. Before multiplying (17) by $\begin{bmatrix} S_{i1} & S_{i2} \end{bmatrix}^T$, we can have

$$\begin{bmatrix} S_{i1}^T \\ S_{i2}^T \end{bmatrix} \widetilde{W}_i = \begin{pmatrix} R_{i1} \\ 0 \end{pmatrix} z_i[0] + \begin{bmatrix} S_{i1}^T \\ S_{i2}^T \end{bmatrix} \Psi_i E_i. \tag{22}$$

By using the second block row, we can solve the following problem to obtain the sparse solution E, instead of $\mathbb{A}$:

$$S_{i2}^T \widetilde{W}_i = S_{i2}^T \Psi_i E_i \quad (i \in \mathscr{P}). \tag{23}$$

Hence, the problem is reduced to reconstruct a sparse vector E_i from the observations $S_{i2}^T \widetilde{W}_i$. Problem (14) is equivalent to the following problem:

$$\begin{aligned} \underset{E_i}{\text{minimize}} \quad & \|E_i\|_{\ell_0} \\ \text{subject to} \quad & S_{i2}^T \Psi_i E_i = S_{i2}^T \widetilde{W}_i, \end{aligned} \tag{24}$$

where $E_i \in \mathbb{R}^T$. As is discussed above, solving problem (24) is in general NP-hard since it requires searches over all subsets of columns of $S_{i2}^T \Psi_i$, a procedure which has exponential complexity. To overcome this problem, a frequently discussed approach considers a similar program in the ℓ_1-norm:

$$\begin{aligned} \underset{E_i}{\text{minimize}} \quad & \|E_i\|_{\ell_1} \\ \text{subject to} \quad & S_{i2}^T \Psi_i E_i = S_{i2}^T \widetilde{W}_i. \end{aligned} \tag{25}$$

This operation is common and can be found in [13, 17, 18]. Throughout this paper, we consider Douglas-Rachford splitting algorithm [13] in the context of above ℓ_1-minimization.

3.4. Theoretical Guarantee. In this paper, we are also interested in studying the theoretical conditions under which obtaining the solution of the problem is guaranteed. It is well known that an inverse problem of finding the solution to the compressive sensing problem involves mathematical questions on the existence, uniqueness, and stability of the solution. On the other hand, the equivalence of the solution between (13) and (25) is not very clear and proof may be needed. We therefore consider two questions for a given $S_{i2}^T \Psi_i$ and signal $S_{i2}^T \widetilde{W}_i$ ($i \in \mathscr{P}$): (i) *uniqueness*: under which conditions a possible sparsest solution is necessarily unique to problem (13)/(25)? and (ii) *equivalence*: under which conditions a sparse solution to problem (13) is also equivalent to the solution of problem (25)?

3.4.1. Uniqueness. As is described in Section 3.3, solving problem (24) requires exhaustive searches over all subsets of columns of $S_{i2}^T \Psi_i$. Actually, it is a combinatorial procedure in nature and has exponential complexity. Inspired by [7, 17], Theorem 3 provides a sufficient condition for a unique solution to problem (24). It guarantees obtaining a unique sparse vector (i.e., E) from the corrupted observations (i.e., $\overline{Z}$) for the ℓ_0 minimization. We denote by $\text{row}_{i \in \mathscr{P}}(E) \in \mathbb{R}^T$ the rows of the matrix E. Before giving the theorem, we need to first introduce the following definition [17].

Definition 2 (see [17, Definition 1.1]). Let $S_{i2}^T \Psi_i$ be the matrix with the finite collection of vectors $\text{col}(S_{i2}^T \Psi_i)_{k \in \mathscr{J}} \in \mathbb{R}^m$ as columns. For every integer $1 \le \nu \le |\mathscr{J}|$, we define the ν-restricted isometry constants ρ_ν to be the smallest quantity such that $S_{i2}^T \Psi_i$ obeys

$$(1 - \rho_\nu) \|E_i\|^2 \le \left\| S_{i2}^T \Psi_i E_i \right\|^2 \le (1 + \rho_\nu) \|E_i\|^2, \tag{26}$$

for all real coefficients $E_{i \in \mathscr{P}}$.

The number ρ_ν measures how close the vectors $\text{row}_i(S_{i2}^T \Psi_i)$ are to behave. In particular, for $\nu = 1$, we can have

$$1 - \rho_1 \le \left\| \text{row}_i \left(S_{i2}^T \Psi_i \right) \right\|^2 \le 1 + \rho_1, \quad \text{for } \forall i \in \mathscr{P}. \tag{27}$$

To see the relevance of ρ_ν to the error recovery problem, we consider the following theorem.

Theorem 3. *In a cyber-physical system, let S_{i2}, $\widetilde{W}_i$, Ψ_i, ν, κ, and $\mathscr{J}$ be specified as above. A sparse solution E can be uniquely recovered from solving the optimization problem (13), if $\rho_{2\nu} < 1$, and $\|\text{col}_{k \in \mathscr{J}}(E)\|_{\ell_0} \le \kappa$.*

Proof. We first prove that if $\rho_{2\nu} < 1$, there exists a unique E_i to problem (24). Suppose for the sake of contradiction that the solution is not unique; then there exist two solutions $E^{\text{opt1}} \ne E^{\text{opt2}}$. Thus, there exists at least one variable i ($1 \le i \le m$) such that

$$\begin{aligned} S_{i2}^T \Psi_i \text{row}_i \left(E^{\text{opt1}} \right) &= S_{i2}^T \widetilde{W}_i, \\ S_{i2}^T \Psi_i \text{row}_i \left(E^{\text{opt2}} \right) &= S_{i2}^T \widetilde{W}_i, \end{aligned} \tag{28}$$

where $\|\text{row}_i(E^{\text{opt1}})\|_{\ell_0} = \|\text{row}_i(E^{\text{opt2}})\|_{\ell_0} = \nu$. Then we can have

$$S_{i2}^T \Psi_i \left(\text{row}_i \left(E^{\text{opt1}} \right) - \text{row}_i \left(E^{\text{opt2}} \right) \right) = 0. \tag{29}$$

By construction $\text{row}_i(E^{\text{opt1}}) - \text{row}_i(E^{\text{opt2}})$ is of size less than or equal to 2ν. Applying (27) and the hypothesis $\rho_{2\nu} < 1$, we conclude that $\|\text{row}_i(E^{\text{opt1}}) - \text{row}_i(E^{\text{opt2}})\|^2 = 0$, contradicting the hypothesis that $\text{row}_i(E^{\text{opt1}})$ and $\text{row}_i(E^{\text{opt2}})$ are distinct.

Then we prove that E is unique to problem (13). Given the proof that E_i, or equivalently $\text{row}_i(E)$, can be uniquely obtained by solving problem (24) and $E = [E_1; \ldots; E_m]$, we conclude that E is unique to the following problem:

$$\begin{aligned} \underset{E}{\text{minimize}} \quad & \|E\|_{\ell_0} \\ \text{subject to} \quad & \overline{W} = W + E. \end{aligned} \tag{30}$$

And given the condition that $\|\text{col}(E)_{k \in \mathscr{J}}\|_{\ell_0} \le \kappa$, we can conclude that E is also unique to problem (13). □

In the literature, a lot of efforts have been made to determine how sparse the desired corrected error must be for equivalence to hold. As we consider to use ℓ_1-minimization instead of ℓ_0 (to obtain the desired error), the conditions in the above lemma may not be guaranteed. Thus, Theorem 4 gives a general condition, which guarantees a unique solution E_i for ℓ_1-minimization problem.

Theorem 4. *In a cyber-physical system, let S_{i2}, $\widetilde{W}_i$, and Ψ_i be specified as above. A sparse solution E can be uniquely recovered from solving the optimization problem*

$$\begin{aligned} \underset{E}{\text{minimize}} \quad & \|E\|_{\ell_1} \\ \text{subject to} \quad & \overline{W} = W + E, \\ & \left\| \text{col}\,(E)_k \right\|_{\ell_1} \le \kappa, \quad k \in \mathscr{J}, \end{aligned} \tag{31}$$

if, for all $E^* \neq E$, we have $\|(E-E^*)_J\|_{\ell_1} - \|(E-E^*)_{\bar{J}}\|_{\ell_1} < 0$ and $\|\mathrm{col}(E)_{k\in\mathcal{J}}\|_{\ell_0} \leq \kappa$, where J and $\bar{J}$ are the support of vectors E and $E^* - E$, respectively.

Proof. We prove that given any $E^{\mathrm{opt1}} \neq E^{\mathrm{opt2}}$ and $\|(E^{\mathrm{opt2}} - E^{\mathrm{opt1}})_J\|_{\ell_1} - \|(E^{\mathrm{opt2}} - E^{\mathrm{opt1}})_{\bar{J}}\|_{\ell_1} < 0$ and $\|\mathrm{col}_k(E)\|_{\ell_0} \leq \kappa$ ($k \in \mathcal{J}$), we can always uniquely recover E^* from (31). Suppose for the sake of contradiction that the solution is not unique; then there exist two instinct solutions that $E^{\mathrm{opt1}} \neq E^{\mathrm{opt2}}$ but $\|E^{\mathrm{opt1}}\|_{\ell_1} = \|E^{\mathrm{opt2}}\|_{\ell_1}$. We use the vectors $\mathbb{A}^{\mathrm{opt1}} = [\mathrm{row}_1(E^{\mathrm{opt1}}); \ldots; \mathrm{row}_m(E^{\mathrm{opt1}})] \in \mathbb{R}^{mT}$ and $\mathbb{A}^{\mathrm{opt2}} = [\mathrm{row}_1(E^{\mathrm{opt2}}); \ldots; \mathrm{row}_m(E^{\mathrm{opt2}})] \in \mathbb{R}^{mT}$ instead of E^{opt1} and E^{opt2}, respectively.

$$\begin{aligned}
\left\|\mathbb{A}^{\mathrm{opt1}}\right\|_{\ell_1} &= \left\|\overline{Z} - Z^{\mathrm{opt1}}\right\|_{\ell_1} = \left\|Z^{\mathrm{opt1}} + \mathbb{A}^{\mathrm{opt1}} - Z^{\mathrm{opt2}}\right\|_{\ell_1} \\
&= \left\|\left(\mathbb{A}^{\mathrm{opt2}} - Z^{\mathrm{opt1}} + Z^{\mathrm{opt2}}\right)_J\right\|_{\ell_1} \\
&\quad + \left\|\left(Z^{\mathrm{opt1}} - Z^{\mathrm{opt2}}\right)_{\bar{J}}\right\|_{\ell_1} \\
&\geq \left\|\mathbb{A}^{\mathrm{opt2}}{}_J\right\|_{\ell_1} - \left\|\left(Z^{\mathrm{opt1}} - Z^{\mathrm{opt2}}\right)_J\right\|_{\ell_1} \\
&\quad + \left\|\left(Z^{\mathrm{opt1}} - Z^{\mathrm{opt2}}\right)_{\bar{J}}\right\|_{\ell_1} \\
&= \left\|\mathbb{A}^{\mathrm{opt2}}{}_J\right\|_{\ell_1} - \left\|\left(\mathbb{A}^{\mathrm{opt2}} - \mathbb{A}^{\mathrm{opt1}}\right)_J\right\|_{\ell_1} \\
&\quad + \left\|\left(\mathbb{A}^{\mathrm{opt2}} - \mathbb{A}^{\mathrm{opt1}}\right)_{\bar{J}}\right\|_{\ell_1} > \left\|\mathbb{A}^{\mathrm{opt2}}{}_J\right\|_{\ell_1} \\
&= \left\|\mathbb{A}^{\mathrm{opt2}}\right\|_{\ell_1},
\end{aligned} \tag{32}$$

contradicting the hypothesis that $\mathbb{A}^{\mathrm{opt1}} \neq \mathbb{A}^{\mathrm{opt2}}$. Therefore, we conclude that $\mathrm{row}_i(E)$ is unique to problem (25). Equivalently, E is unique to the following problem:

$$\begin{aligned}
&\underset{E}{\text{minimize}} \quad \|E\|_{\ell_1} \\
&\text{subject to} \quad \overline{W} = W + E.
\end{aligned} \tag{33}$$

Furthermore, given the condition that $\|\mathrm{col}(E)_{k\in\mathcal{J}}\|_{\ell_0} \leq \kappa$, we conclude that E is unique to problem (31). □

In conclusion, Theorems 3 and 4 show that the hypothesis of our theorem holds provided that the sparse error can be uniquely corrected. Naturally, if the assumption does not hold, then neither does (13) or (31).

3.4.2. Equivalence. Next, we will discuss the conditions under which it is theoretically possible to use ℓ_1-minimization to obtain the sparse solution E (or $\mathbb{A}$) instead of ℓ_0-minimization. We derive an algorithm for precisely verifying ℓ_0-ℓ_1 equivalence. We can use the following definition and proposition proposed in [19].

Definition 5 (see [19, Definition 2]). We define $\mathcal{SK}_d(B_1)$ as the collection of all d-dimensional faces of the ℓ_1-ball B:

$$\mathcal{SK}_d(B_1) \doteq \left\{\mu \in \mathbb{R}^{mT} : \|\mu\|_{\ell_1} = 1, \|\mu\|_{\ell_0} \leq d+1\right\}, \tag{34}$$

where $B_1 \doteq \{\mu \in \mathbb{R}^{mT} : \|\mu\|_{\ell_1} \leq 1\}$.

Proposition 6 (see [19, Proposition 3]). *In a cyber-physical system, let S_{i2}, $\widetilde{W}_i$, and Ψ_i be specified as above. For every $E_i \in \mathbb{R}^T$ and $S_{i2}^T\widetilde{W} \in \mathbb{R}^{T-1}$, the following implication holds:*

$$\begin{aligned}
\left\|S_{i2}^T\widetilde{W} - S_{i2}^T\Psi_i E_i^*\right\|_{\ell_0} \leq \frac{1}{2}\mathscr{C}_i \Longrightarrow E_i^* \\
= \underset{E_i}{\mathrm{argmin}} \left\|S_{i2}^T\widetilde{W} - S_{i2}^T\Psi_i E_i\right\|_{\ell_1},
\end{aligned} \tag{35}$$

if and only if $\forall \mu \in \mathcal{SK}_{\mathscr{C}_i-1}(B_1), \forall S^T\widetilde{W} \in \mathbb{R}^T \setminus 0$ and $\|\mu + S_{i2}^T\Psi_i S^T\widetilde{W}\|_{\ell_1} > 1$, where $\mathscr{C}_i$ = (number of columns of $S_{i2}^T\Psi_i$ that are linearly independent).

Proof. See Proposition 3 in [19]. □

Note that implication (35) is the condition that we want to verify. As we need to deal with high-dimensional matrices (e.g., $E \in \mathbb{R}^{m\times T}$), we need to give asymptotic guarantees of equivalence, which is described in Proposition 6. In our experiments, it is confirmed that we can benefit from this equivalence, even when the matrices are in high dimensions.

4. Experimental Results

4.1. Case Study: Power Network. We employ a real-world power grid system as the test system we used. A state-space control model in a smart grid consists of buses connected to transmission lines. We use the IEEE 14-bus system as the test system [20]. Moreover, we use the real load data in year 2016 from New York Independent System Operator (NYISO). The NYISO load data include the 11 regions (namely, A-H). Similar to [12], the following procedures are used to estimate 5-minute system state (x) using load pattern from NYISO.

(1) Link each load bus of IEEE 14-bus system to one region of NYISO using the following matrix:

$$\begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 9 & 10 & 11 & 12 & 13 & 14 \\ \mathrm{F} & \mathrm{C} & \mathrm{I} & \mathrm{B} & \mathrm{G} & \mathrm{K} & \mathrm{E} & \mathrm{H} & \mathrm{J} & \mathrm{D} & \mathrm{A} \end{pmatrix}. \tag{36}$$

The first row of the matrix is the bus number of IEEE 14-bus system and the second row represents the corresponding NYISO region index.

(2) Normalize the load data collected from NYISO to the initial real and reactive load of the corresponding IEEE 14-bus system. Due to lack of reactive load information in NYISO database, we use the direct current (DC) power flow model to estimate system states. This condition can be relaxed when the reactive load data is available.

(3) Add the normalized load data on the IEEE 14-bus system.

(4) Estimate the system state ($\hat{x}$) from the solution of power flow analysis for benchmarking purpose. In this paper, we apply Newton-Raphson algorithm for estimating $\hat{x}$.

Table 1: Regression coefficients for the predicting at 11:55 pm, Jun 30, 2016.

	$\widehat{\gamma}_1$	$\widehat{\gamma}_2$	$\widehat{\gamma}_3$
Zone A	-7.12×10^{-6}	1.02	−14.62
Zone B	-2.66×10^{-6}	1.01	−2.47
Zone C	-1.63×10^{-5}	1.05	−41.82
Zone D	-1.5×10^{-3}	2.39	−314.79
Zone E	-2.14×10^{-5}	1.03	−12.99
Zone F	-8.72×10^{-6}	1.02	−16.77
Zone G	-4.03×10^{-6}	1.01	−3.64
Zone H	-6.22×10^{-5}	1.04	−4.19
Zone I	-2.67×10^{-5}	1.04	−13.75
Zone J	-1.76×10^{-6}	1.02	−474.37
Zone K	-1.27×10^{-6}	1.01	−18.61

Similar to [12], we estimate T operating points of the system state (x) by adding the normalized 5-minute load data on the MATPOWER IEEE 14-bus case file [21]. In this paper, we use one-day NYISO data as the testing set. Thus, on one day, there will be 288 operating points. So, we set $T = 288$ to construct the F-DDIA method. Second, we prepare the *attacked* samples as follows. We let the parameter κ range from 1 to $m = 54$ in the IEEE 14-bus test system. For each κ, we simulate κ-specific meters to attempt the attack construction ($a = Hc$) with a randomly injected error c. Thus, at most, a total of 6564 labeled samples, which includes 6017 attack samples and 547 initial samples (without attacks), are prepared.

4.2. Parameters in Load Fitting. According to Section 2.3, the γ in (6) is the parameter of the measurement (load) dynamical model for power grid system. We estimate γ by polynomial regression using data traces of $\overline{z}[k+1] - \overline{z}[k]$. The historical load data in NYISO and attack samples prepared in previous session are used to construct the matrix $\widehat{\gamma}$ (i.e., polynomial regression in order of l) in (6). The measurement dynamics at each time k are estimated by the data of 24 hours prior to the time. For example, if we want to estimate the load dynamics at 0:05 am Jun 30, Zone F, the load data samples (which may contain attacks) during 0:05 am, Jun 29–0:00 am, Jun 30 are used.

We are concerned about what the regression order l is appropriate for fitting the dynamics of the system. The experimental results show that $l = 2$ is a suitable regression order. As the increase of l will improve the load fitting accuracy at the cost of computation time, we will use $l = 2$ in the rest of our experiments. Table 1 gives the regression results for predicting the dynamical model by using the load data on Jun 30, 2016.

Specifically, we take Zone F for an example; Figure 1 shows a quadratic polynomial fit of load in Zone F with 95% confidence bounds (the 95% interval indicates that we have a 95% chance that a new observation will fall within the bounds.). We collect the hourly data to fit the model, where the blue "+" represents the actual hourly load, and the green curve describes the fitting model.

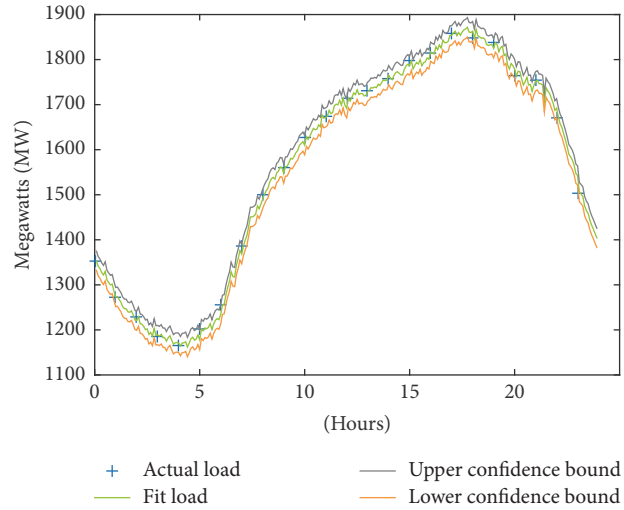

Figure 1: The quadratic polynomial fit of the load data in Zone F with 95% confidence bounds on Jun 30, 2016, when $l = 2$.

4.3. Performance Matrices. When $\mathbb{A}$ is calculated by our detector, we set the following rule to identify whether the system is attacked:

$$\mathscr{D}_i[k] = \begin{cases} 1 & \left|\mathbb{A}_{i+11k}\right| \geq \sigma_{\text{ob}} \times \left|\overline{Z}_{i+11k}\right| \\ 0 & \text{otherwise,} \end{cases} \tag{37}$$

where σ_{ob} is the observation threshold when detecting data injection attacks. The parameter σ_{ob} will be discussed later in this section. We denote the user-defined threshold $\mathscr{D}_i[k] = 1$ when $\overline{z}_i[k]$ is identified as attacked. Then, we identify whether $\overline{z}[k]$ is attacked by aggregating the values of $\mathscr{D}_i[k]$ ($i \in \mathscr{P}$). We predict $\overline{z}[k]$ as *attacked* (denoted as Label$[k] = 1$) if the sum of $\mathscr{D}_i[k]$ is larger than the all-users-defined threshold $\mathscr{N}_a$, and *secure* (denoted as Label$[k] = 0$) otherwise:

$$\text{Label}[k] = \begin{cases} 1 & \sum_{i=1}^{m} \mathscr{D}_i[k] > \mathscr{N}_a \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

In smart grid networks, the major concern is not only the detection of attack cases but also that of the secure cases. In other words, after following the rule (38), we need to be careful of the samples with high precision and recall performance in order to avoid false alarms. Therefore, we utilize precision and recall metrics, which are commonly used for classification tasks [10]. Specifically, as Table 2 defines, we denote CA as the number of attacked samples, which we identified as *attacked*, WA as the number of secure samples, which we identified as *attacked*, CS as the number of secure samples, which we identified as *secure*, and WS as the number of attacked samples, which we identified as *secure*.

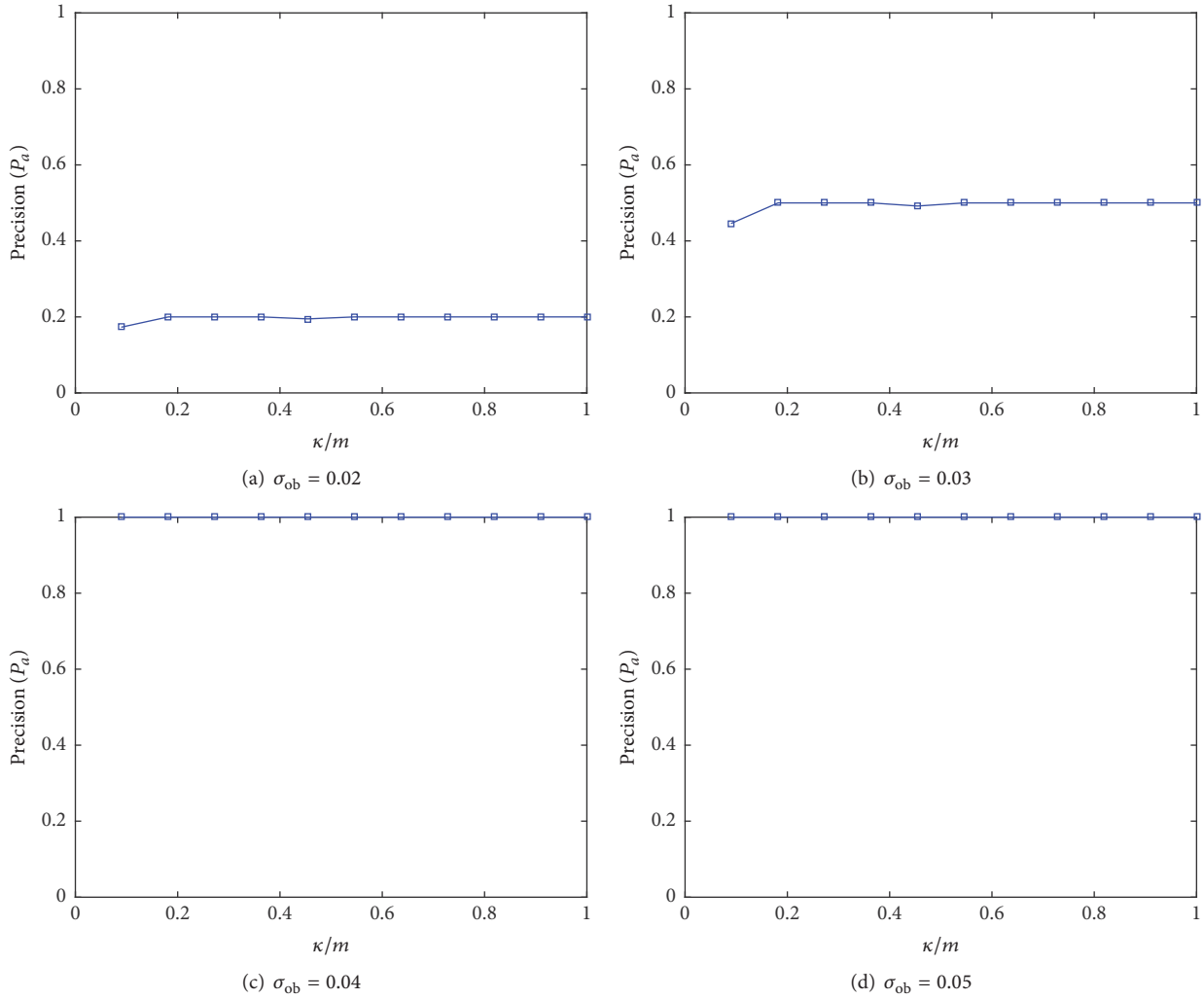

Figure 2: Precision of *attacked* samples for the IEEE 14-bus system.

Table 2: Denotations for defining evaluation metrics.

	Attacked	Secure
Classified as attacked	CA	WA
Classified as secure	WS	CS

In addition, the performance of the proposed detector can be measured by the precision and recall metrics:

$$P_a = \frac{\text{CA}}{\text{CA} + \text{WA}},$$
$$R_a = \frac{\text{CA}}{\text{CA} + \text{WS}},$$
$$P_s = \frac{\text{CS}}{\text{CS} + \text{WS}}, \quad (39)$$
$$R_s = \frac{\text{CS}}{\text{WA} + \text{CS}},$$

where P_a (P_s) and R_a (R_s) indicate the precision and recall values for the class *attacked (secure)*, respectively. Precision values give information about the decision performance of the algorithms among identified class. And recall values measure the degree of attack retrieval.

4.4. Performance on Detecting Attacks. We first analyze the performance of the proposed algorithm against the attacks, which are made from a set of false data injection attacks when $\kappa = 1$. In the experiments, we observe that the selection of threshold parameter σ_{ob} does affect the precision and recall performances. Table 3 shows the comparison for different σ_{ob} values. P_a and R_s increase as σ_{ob} increases and remain 100% when $\sigma_{ob} \geq 0.04$. In addition, R_a and P_s decrease as σ_{ob} increases. Note that the precision value at $\sigma_{ob} = 0.01$ is 7.14% and the recall value at $\sigma_{ob} > 0.06$ is lower than 50% for class *attacked*. Thus, the optimal σ_{ob} value should be in range $[0.02, 0.06]$. Note that the performance at $\sigma_{ob} = 0.05$ is quite similar to that at $\sigma_{ob} = 0.06$; thus we do not draw the performance at $\sigma_{ob} = 0.06$ in Figures 2, 3, 4, and 5 to avoid unreadability.

The performance of different σ_{ob} values for identifying *attacked* samples is compared in Figures 2 and 3, where

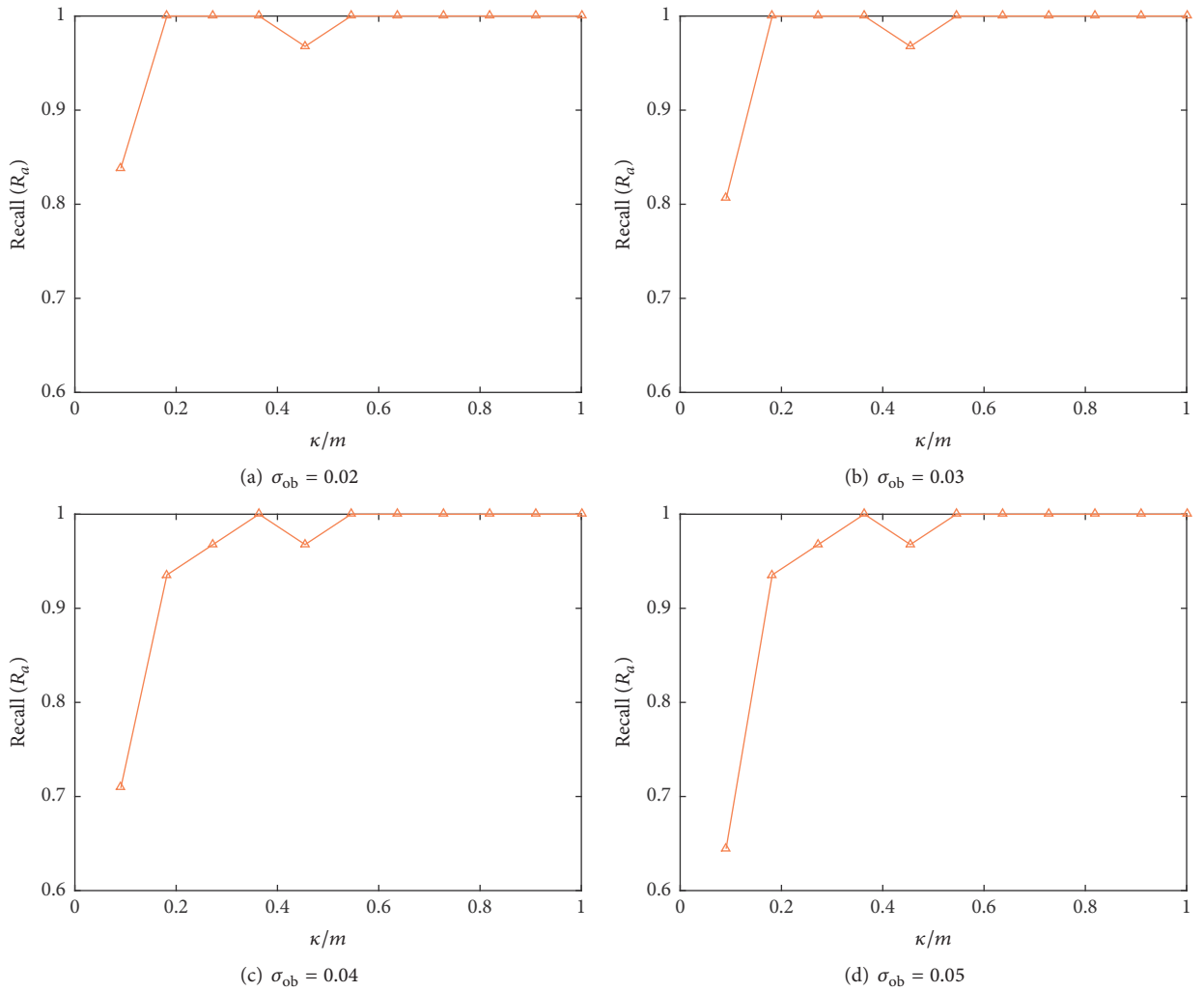

Figure 3: Recall of *attacked* samples for the IEEE 14-bus system.

Table 3: Performance of proposed detector against multiperiod attacks for IEEE 14-bus system, $\kappa = 1$.

σ_{ob}	P_a	R_a	P_s	R_s
0.01	7.14%	100%	100%	95.47%
0.02	**17.33%**	**83.87%**	**99.94%**	**98.61%**
0.03	**44.64%**	**80.65%**	**99.93%**	**99.65%**
0.04	**100%**	**70.97%**	**99.90%**	**100%**
0.05	**100%**	**64.52%**	**99.88%**	**100%**
0.06	100%	63.64%	99.87%	100%
0.07	100%	41.67%	99.80%	100%
0.08	100%	45.45%	99.81%	100%
0.09	100%	41.67%	99.80%	100%
0.10	100%	38.10%	99.78%	100%
0.20	100%	9.10%	99.68%	100%

$\kappa/m \in [0, 1]$. We observe that P_a increases and R_a decreases when σ_{ob} increases. The precision value of *attacked* class is approximately 100% when $\sigma_{ob} = 0.04$ and $\sigma_{ob} = 0.05$. The recall value of the *attacked* class increases with rising κ/m values and is approximately 100% when κ/m is larger than 54.55%. Although the proposed algorithm at $\sigma_{ob} = 0.02$ and $\sigma_{ob} = 0.03$ may correctly detect the *attacked* samples as κ/m increases, the *secure* variables are incorrectly labeled as *attacked* and therefore give more false alarms.

Meanwhile, the performance of identifying *secure* samples is compared in Figures 4 and 5. Both values (precision and recall) of the *secure* class are high (i.e., near 100%). Summing up, the above experimental results show that if we choose the parameter $\sigma_{ob} \in [0.04, 0.06]$, our methodology can efficiently detect the data injection attacks.

4.5. Performance on Recovering System States. In this part, we compare the performances of our detector and the residual-based approach with the performance of recovering the initial systems states. We first introduce how we evaluate the performances of an algorithm. In IEEE 14-bus system, the state vector x will have 14 bus voltage magnitudes and 13 phase angles, where the phase angle of one reference bus is set as the reference. If the system is observable [14], the state vector x can be represented as follows:

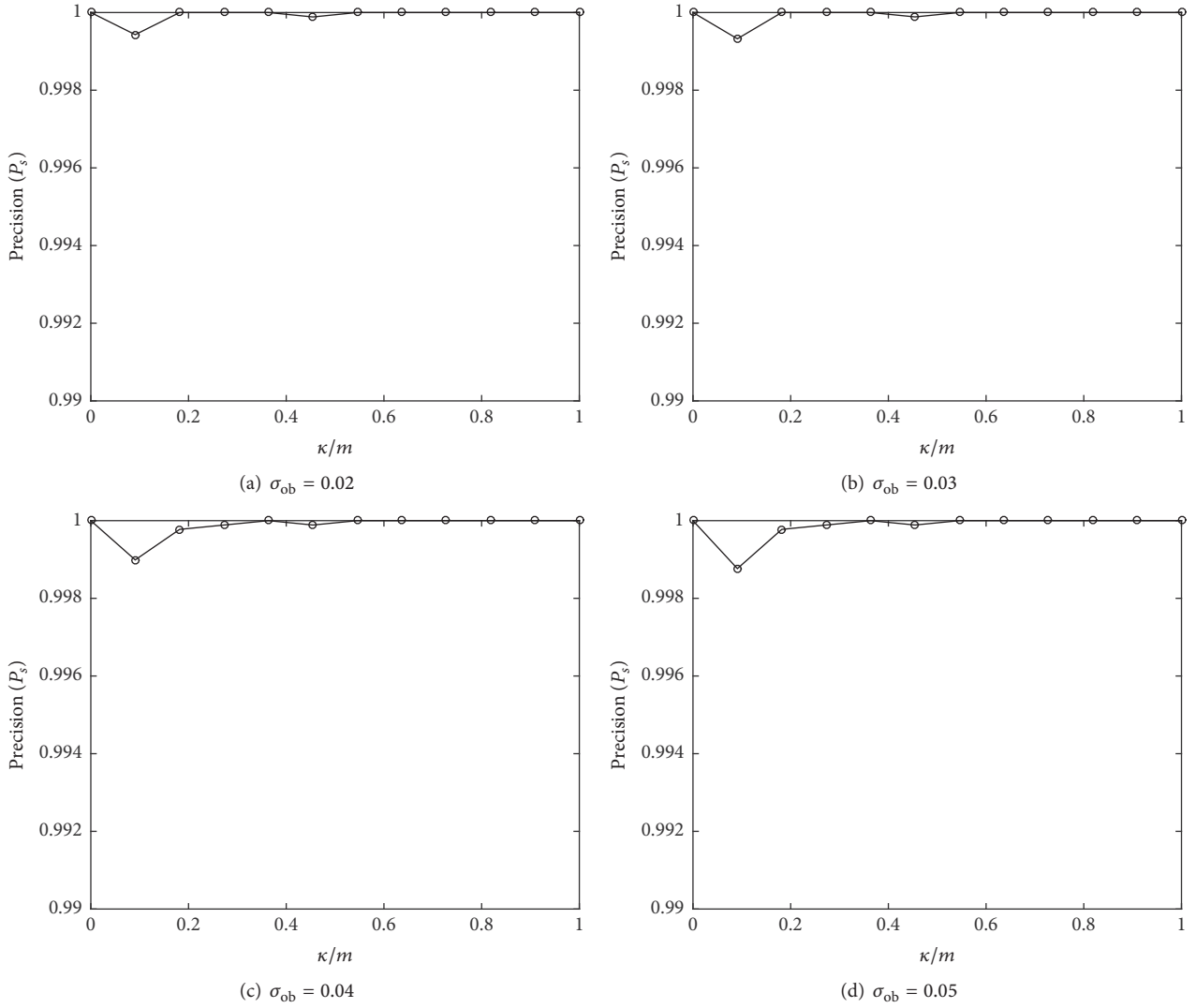

Figure 4: Precision of *secure* samples for the IEEE 14-bus system.

$x = (V_1, V_2, \ldots, V_{14}, \theta_2, \theta_3, \ldots, \theta_{13})^T$, where V_i, θ_i is voltage magnitude and voltage angle at bus i. Therefore, the average absolute phase error for bus i, denoted as $\zeta_{\theta_i[k]}$, can be described as follows:

$$\zeta_{\theta_i[k]} = \frac{1}{\vartheta}\sum_{j=1}^{\vartheta}\zeta_{\theta_i^j[k]} = \frac{1}{\vartheta}\sum_{j=1}^{\vartheta}\left|\frac{\widehat{\theta}_i^j[k] - \theta_i[k]}{\theta_i[k]}\right|, \quad (40)$$

where

ϑ is number of testing samples;

$\zeta_{\theta_i^j[k]}$ ith is bus absolute phase error at time k when under the jth attack in the testing samples;

$\widehat{\theta}_i^j[k]$ ith is bus recovered phase angle at time k when under the jth attack in the testing samples;

$\theta_i[k]$ ith is bus true phase angle at time k.

The proposed algorithm and residual-based algorithm have been tested under various attack scenarios (i.e., $\kappa = 1, 2, \ldots, 11$). Table 4 presents the results when $\kappa = 1$ and $\kappa = 11$, respectively. We can first see the superiority of our methodology, comparing with the residual-based algorithm. For example, when $\kappa = 11$, the average phase error of our proposed algorithm on bus 12 is 0.2886, whereas the error is 4.2793 for the residual-based algorithm, which is 13 times larger than our algorithm. Second, we can see that the average phasor errors for $\kappa = 1$ are in general smaller than those for $\kappa = 11$, which means that the performances of both algorithms work better when κ is small. Third, we see that the F-DDIA result of bus 11 (or bus 13) is quite different from the that of bus 12 (or 14). We think the reason that causes this phenomenon is because of the κ value. When $\kappa = 1$, the bus indexes 11–14 are with little difference. To sum up, the reason that causes this difference of the average absolute phase error is complex, and thus the F-DDIA performance depends on a series of parameters (i.e., κ, σ_{ob}, etc.).

4.6. Comparison on Execution Time. In our experiments, we find that the proposed approach is faster than other

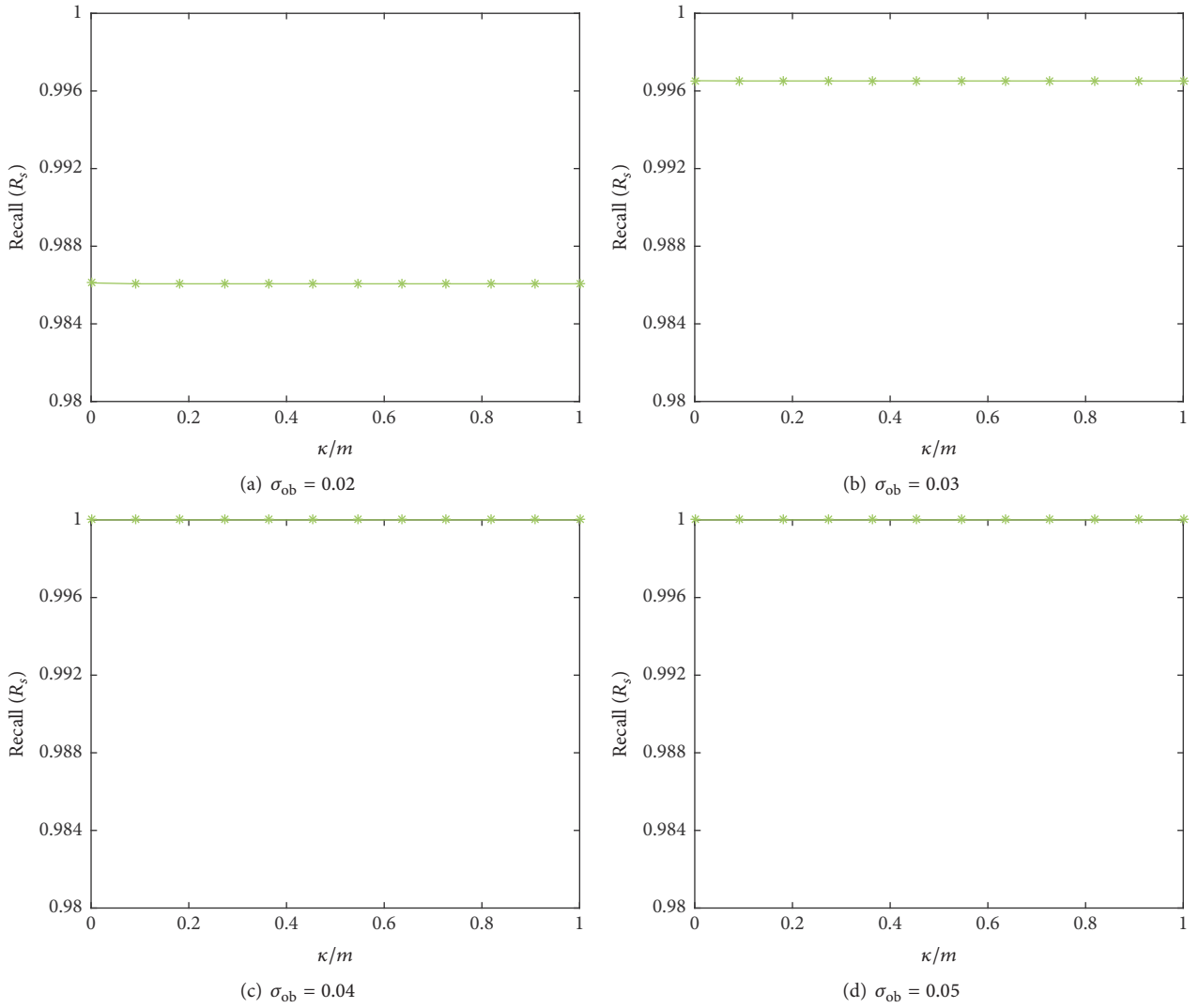

Figure 5: Recall of *secure* samples for the IEEE 14-bus system.

Table 4: $\zeta_{\theta_i[k]}$ for IEEE 14-bus system when $\sigma_{ob} = 0.05$.

Bus index	$\zeta_{\theta_i[k]}$ ($\kappa = 1$)		$\zeta_{\theta_i[k]}$ ($\kappa = 11$)	
	Residual-based	Our detector	Residual-based	Our detector
2	0.2388	0.0770	0.5855	0.1387
3	0.7420	0.3612	1.2131	0.1374
4	0.3274	0.0430	0.9988	0.2117
5	0.4026	0.1076	0.9927	0.2187
6	1.0687	0.3478	2.0091	0.7807
7	0.6875	0.0493	1.6335	0.4186
8	0.6875	0.0493	1.6335	0.4186
9	0.8095	0.0443	1.8286	0.4823
10	0.9589	0.1817	1.7800	0.7220
11	0.9869	0.2925	1.8136	1.6674
12	1.2767	0.4810	4.2793	0.2886
13	1.1997	0.4977	2.1074	1.1148
14	1.0374	0.2894	1.6876	0.3817

works. The residual-based fault detector takes around 50 min (0.043 s per sample), while the proposed approach only takes 12 min (0.011 s per sample). The 12 min of our approach includes load dynamics fitting and Douglas-Rachford iterations process. The main computation burden for our proposed approach is to proceed Douglas-Rachford iterations for basis pursuit process. In general, we do not consider the state estimation process. This is why our proposed approach is faster than the other one.

5. Conclusions

The paper examines the problem of detecting data injection attacks in smart grid networks. We propose a detection framework named F-DDIA, which can recover the initial system state, as well as the real measurement readings. Due to the sparse nature of data injection attacks, ℓ_1 minimization technique (including Douglas-Rachford) can be applied. The validation of the proposed detecting algorithm is validated using load data from NYISO. Our detector works well in both linear and nonlinear systems.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported in part by National High Technology Research and Development Program of China (no. 2015AA016008).

References

[1] R. Langner, "Stuxnet: dissecting a cyberwarfare weapon," *IEEE Security and Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, article 13, 2011.

[3] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 7, pp. 1460–1470, 2014.

[4] J. Wang, L. C. K. Hui, S. M. Yiu, X. Cui, E. K. Wang, and J. Fang, "A survey on the cyber attacks against non-linear state estimation in smart grids," in *Proceedings of the Australasian Conference on Information Security and Privacy (ACISP '16)*, Springer, Berlin, Germany, 2016.

[5] J. Wang, L. C. K. Hui, S. M. Yiu, E. K. Wang, and J. Fang, "A survey on cyber attacks against nonlinear state estimation in power systems of ubiquitous cities," *Pervasive and Mobile Computing*, vol. 39, pp. 52–64, 2017.

[6] 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, no. 2, pp. 612–621, 2014.

[7] Q. Hu, D. Fooladivanda, Y. H. Chang, and C. J. Tomlin, "Secure state estimation and control for cyber security of the nonlinear power systems," 2016, https://arxiv.org/abs/1603.06894.

[8] D. Han, Y. Mo, and L. Xie, "Robust state estimation against sparse integrity attacks," 2016, https://arxiv.org/abs/1601.04180.

[9] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[10] J. Wang, W. Tu, L. C. K. Hui, S. M. Yiu, and E. K. Wang, "Detecting time synchronization attacks in cyber-physical systems with machine learning techniques," in *Proceedings of the In 37th IEEE International Conference on Distributed Computing Systems (ICDCS '17)*, 2017.

[11] 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.

[12] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting False Data Injection Attacks in AC State Estimation," *IEEE Transactions on Smart Grid*, 2015.

[13] L. Demanet and X. Zhang, "Eventual linear convergence of the Douglas-Rachford iteration for basis pursuit," *Mathematics of Computation*, vol. 85, no. 297, pp. 209–238, 2016.

[14] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*, CRC Press, Boca Raton, Fla, USA, 2004.

[15] J. Wang, L. C. K. Hui, and S. M. Yiu, "Data framing attacks against nonlinear state estimation in smart grid," in *Proceedings of the IEEE Global Communications Conference Workshop (GLOBECOM '15)*, 2015.

[16] C. R. Goodall, *13 Computation Using The QR Decomposition*, Handbook of Statistics, 1993.

[17] E. J. Candes and T. Tao, "Decoding by linear programming," *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[18] J. Zhang, C. Zhao, D. Zhao, and W. Gao, "Image compressive sensing recovery using adaptively learned sparsifying basis via L0 minimization," *Signal Processing*, vol. 103, pp. 114–126, 2014.

[19] Y. Sharon, J. Wright, and Y. Ma, "Computation and relaxation of conditions for equivalence between l_1 and l_0 minimization," Tech. Rep. UILU-ENG-07-2208, University of Illinois, Urbana-Champaign, Illinois, Ill, USA, 2007.

[20] S. K. M. Kodsi and C. A. Canizares, "Modeling and simulation of ieee 14 bus system with facts controllers," Tech. Rep., University of Waterloo, Waterloo, Ontario, Canada, 2003.

[21] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.