Detecting False Data Injection Attacks Using Canonical Variate Analysis in Power Grid

Chao Pei, Yang Xiao, *Fellow, IEEE*, Wei Liang, *Senior Member, IEEE*, and Xiaojia Han

***Abstract*—With the knowledge of the measurement configuration and the topology structure of a power system, attackers can launch false data injection attacks (FDIAs) without detection by existing bad data detection methods in state estimation. The attacks can also introduce errors to estimated state variables, which are critical to grid reliability and operation stability. Existing protection methods cannot handle dynamic and variable network configurations. In this paper, to effectively defend against FDIAs, we propose a canonical variate analysis based detection method which monitors the variation of statistical detection indicators T^2 and Q about projected canonical variables before and after attacks. Unlike most statistic models that only consider cross-correlation of discrete measurements constrained by Kirchhoff's Law at each independent sampling time, we also consider the auto-correlation of measurements caused by time series characteristics of varying loads. Experiment results on IEEE-14 bus system demonstrate the effectiveness and accuracy of our proposed method based on both synthetically generated data and real-world electricity data from the New York independent system operator.**

Manuscript received February 29, 2020; revised May 7, 2020; accepted July 11, 2020. Date of publication July 14, 2020; date of current version July 7, 2021. The work of Chao Pei and Wei Liang was supported in part by the National Natural Science Foundation of China under Grant 61673371, in part by the International Partnership Program of the Chinese Academy of Sciences under Grant 173321KYSB20180020, and in part by the Liaoning Provincial Natural Science Foundation of China under Grant 2019-YQ-09. This research was carried out while the first author is visiting Department of Computer Science, The University of Alabama, Tuscaloosa, USA, supported by the China Scholarship Council. Recommended for acceptance by Dr. Mohsen Guizani. *(Corresponding authors: Yang Xiao and Wei Liang.)*

Chao Pei is with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China, with the Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China, with the Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China (e-mail: cpei@eng.ua.edu).

Chao Pei is with the University of Chinese Academy of Sciences, Beijing 100049, China, and also with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA (e-mail: cpei@eng.ua.edu).

Yang Xiao is with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290 USA (e-mail: yangxiao@cs.ua.edu).

Wei Liang is with the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China (e-mail: weiliang@sia.cn).

Wei Liang is with the Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China, and also with the Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China (e-mail: weiliang@sia.cn).

Xiaojia Han is with the Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China, with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China (e-mail: hanxiaojia@sia.cn).

Xiaojia Han is with the Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: hanxiaojia@sia.cn).

Digital Object Identifier 10.1109/TNSE.2020.3009299

***Index Terms*—Adversarial attack and defense, artificial intelligence security, attack detection, canonical variate analysis, cyber security, false data injection attack (FDIA), smart grid, state estimation.**

I. Introduction

THE ever increasing demand for reliable, sustainable, and economical electricity services necessitate near real-time monitoring and control in power system operations [1], [2]. Traditionally, key infrastructures such as power systems have been built on local area networks (LANs) and are not connected to the Internet. But in recent years, power systems are gradually exposed to the Internet, facing increasingly severe threats and challenges to cyber security, such as network attacks, viruses, and so on. The security protection of power systems is mainly to prevent unauthorized or accidental access, tampering, and destruction of industrial systems [3], [4].

Measurements that a control center received are obtained from Remote Terminal Units (RTUs), and go over a wide-area network in a supervisory control and data acquisition (SCADA) system. The commonly used protocols in power systems include Distributed Network Protocol (DNP) 3, Modbus, Profibus, etc. The protocols themselves are in clear text. Meanwhile, most devices such as RTUs are still embedded devices, and therefore they do not possess enough computational power to perform encryption and authentication [5], [6]. As a result, the integrity and availability of transmitted measurements are vulnerable to malicious attackers in power systems. The collapse of power grids occurs from time to time due to cyber attacks. For example, in June 2010, Iran's nuclear power plant was attacked by Stuxnet virus [7]. On Dec. 23, 2015, Ukraine was attacked by BlackEnergy, resulting in massive power outages in hundreds of thousands of households [8]. This is the first case of a direct power supply interruption caused by a cyber attack [8]. Recently in the afternoon of March 7, 2019, power blackouts in most parts of Venezuela lasted for more than 24 hours, resulting in the paralysis of major transportation systems and the failure of infrastructure [9].

False data injection attacks are targeted to the integrity and availability of received measurements and have attracted numerous attack and defense studies when they are firstly proposed in 2009 [10]. By systematically tampering with some meter measurements, the main characteristic of FDIAs is that they are able to circumvent traditional detection methods which are based on measurement residual testing [11]. The real-time operation states of power systems are obtained from

the accurate state estimation, which is also the basis for bad data detections. The main function of bad data detection aims at maintaining stability and security of power systems. Generally, bad measurements in power grids can be caused by topology errors or untended measurement abnormalities due to meter failures or malicious attacks [12]. However, by systematically tampering with some meter measurements, attackers in FDIAs are able to circumvent traditional detection methods which are based on measurement residual testing. Meanwhile, FDIAs can lead the state estimator to produce manipulated estimation results. It has been proved in [10] that if the number of compromised measurements κ satisfies the condition $\kappa > m - n + 1$, it is guaranteed that there exists a false data injection attack unobservable by the residual based bad data detection, where m represents the number of measurements and n is the number of state variables in the power system. The attack vectors of FDIAs are not simply random perturbations of meter measurements, but are rather carefully crafted with the information of Jacobian matrix H, which is related to the topological structure and line parameters information of power system. Several methods of constructing successful FDIAs can be found in [10], [13], [14]. Recently, a class of sparse undetectable attacks is also designed to worsen the state estimation performance even in the presence of sensor failures [38].

Existing FDIA mitigation studies can be broadly classified into two major types: protection based defense and detection based defense. The protection based defend methods are aimed at protecting power grids from attackers in advance by offline ways, while the detection based defend methods are aimed at detecting and identifying FDIAs during the state estimation process. The protection based defense mainly focuses on increasing the measurement redundancy, improving the accuracy, and bad data detection ability of state estimation. Adopted measures are either by performing additional encryption algorithms, or by adding isolation devices in power systems. How to find the minimum set of measurements that need to be protected effectively for the purpose of minimizing protection cost is analyzed in [15]. Meanwhile, advanced measurement equipments such as phasor measurement units (PMUs) which can directly obtain estimated state variables are used to defend FDIAs in [16], [17]. The main idea of deploying these advanced PMUs is to guarantee the observability of power systems such that attacks can be observed when they are launched. However, when the topology of a power system is changed, the deployment results need to be reconsidered because the observability is also changed. Therefore, these PMUs deployment protection mechanisms are usually insufficient for variable system structures and configurations. Even though these protection based methods are effective and necessary, it is not enough to completely protect power grids against attacks.

Detection based defend methods identify FDIAs by developing anomaly detection mechanisms based on analyzing measurements before and after attacks. Traditional used bad data detection methods include the $J(x)$ detector and the largest normalized residual based (LNR) detector. Both of them can effectively detect bad data caused by random noise. A Bayesian framework, which can capture the prior information about states of the power system, is proposed in [18] for bad data detection. In [19], FDIAs are detected by monitoring the measurement variance and state changes of two sequential data collection slots. Further, the state change vector is estimated and compared with a predefined threshold. A short-term state forecasting-aided based detection method is proposed in [20] to detect FDIA. This method is based on the fact that predicted measurements are regarded as expected measurements. The authors in [21] present a cosine similarity matching approach to detect FDIAs by comparing estimated results of a Kalman filter and measurements from PMUs. A Robust Principle Component Analysis (RPCA) method is proposed in [12] for FDIA identification based on low rank and sparse decomposition for the attacked measurement matrix. In order to achieve a better balance between computation efficiency and FDIA detection accuracy in the matrix separation problem of RPCA, a fast go-decomposition approach is proposed in [39]. Also, a realtime detection method in [5] leverages gathered information from load forecasts, generation schedules, and real-time measurements from PMUs, which are independent with SCADA measurements to detect anomalies. Several data driven methods such as support vector machine (SVM) [22], [23], improved extreme learning machine (ELM) [24], deep learning-based method [25], and statistical unsupervised method [26] are presented. They all based on the fact that normal measurements and attacked measurements can be statistically distinguished for the reason that normal measurements are governed by physical laws, such as Kirchhoff's law, whereas these attacked measurements are not. Particularly, a real-time principle component analysis (PCA) based detection method is recently proposed in [37] based on extracted information about correlations of collected measurements. The PCA based method can provide accurate and sensitive response towards FDIAs. However, these proposed methods only consider the cross-correlation of discrete measurements constrained by Kirchhoff's Law at each independent sampling time, the auto-correlation of measurements for a period of sampling time is ignored. But the auto-correlation of measurements does exist, because the weather condition and loads in smart grids are changing with the normal operation of the system.

This paper investigates an alternative method to FDIA detection using data-based approach. From the perspective of multivariate statistical process monitoring (MSPM), a canonical variate analysis based detection method which monitors the variation of statistical indicators about canonical variables is proposed. Unlike most existing statistical models that only consider cross-correlation (i.e., across smart meters) of discrete measurements constrained by Kirchhoff's Law at each independent sampling time, the auto-correlation of measurements is also taken into account in our proposed method. The auto-correlation of measurements among consecutive time slots is based on the obvious time series characteristics of changing load and weather condition. Based on real-time received measurements, canonical state variables can be directly obtained by projection matrixes learned from normal historical measurements. Large deviations

between the real-time statistical detection indicators and thresholds are treated as indications of attacked measurements. Then an alarm is raised in the control center to alert FDIAs.

The contributions of our paper are presented as follows. We propose a canonical variate analysis based detection method for the first time to rapidly identify FDIAs, which are undetected by traditional residual based detection methods. Our proposed method considers both cross-correlation and auto-correlation of received meter measurements among consecutive time slots. To verify the performance of the proposed detection method, both synthetically generated data and real-world electricity data are used and evaluated. The synthetically generated data is obtained with loads of uniform variation while the real-world electricity data is obtained with non-stationary loads from the New York independent system operator (NYISO) [27]. The proposed detection method can be integrated as an additional function of bad data detection in the control center.

The rest of the paper is organized as follows. The state estimation and FDIAs models are presented in Section II. The canonical variate analysis based detection method is proposed in Section III. The experimental results and performance evaluations are given in Section IV. Finally, Section V concludes the paper and discusses future works.

II. State Estimation and Attack Model

Power systems exhibit nonlinear characteristics in real-world utilities and usually the required state variables are obtained by performing alternating current (AC) state estimation. It has been verified that it is feasible to launch FDIAs in the AC non-linear state estimation process although the condition of constructing FDIA in AC state estimation is more complicated compared with the case of DC situation. However, due to the reason that DC state estimation model is particularly simple for analyzing, we here adopt the commonly used state estimation model as many papers.

State variables such as voltage phase angles of all buses in a power system are very important for making real-time decisions to subsequent operations, such as automatic generation control (AGC) and optimal power flow analysis (OPF). Although advanced measurement equipment such as phasor measurement units (PMUs) can directly measure these phase angles, these measurement equipment is very expensive in practice and it is not feasible to deploy enough PMUs to measure and secure all state variables. In power system state estimation, the unknown state vector x is obtained through redundant known measurement vector z and the topology structure information of power system. A common DC approximation model is usually used in practice to simplify the analysis. It is based on the assumptions as follows: (1) amplitudes of all bus voltages are close to unity, (2) line resistances are negligible, and (3) angle differences between any two buses are sufficiently small. The approximation model is written as follows:

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{x} = [x_1, x_2, \cdots x_n]^T$ is the state vector which includes voltage phase angles of all n buses in a power grid. $\boldsymbol{e} = [e_1, e_2, \cdots e_m]^T$ is the m-dimensional measurement error vector which is commonly assumed to be a zero mean Gaussian random variable with a known covariance $\boldsymbol{R}$. $\boldsymbol{H}$ is a $m \times n$ Jacobian matrix which is related to the power grid topology. $\boldsymbol{z} = [z_1, z_2, \cdots z_m]^T$ is the measurement vector which is obtained from m smart meters. This measurement vector includes active power injection measurements P_i of n buses and $m - n$ active power flow measurements P_{ij} of branches in a power grid, where $i \in [1, n]$, $j \in [1, n]$, $n < m$. The active power injection measurements P_i are defined as the differences of the active power generations and the active power demands at each bus in the power grid, while the active power flow measurement P_{ij} for a transmission branch is the active power between two adjacent buses i and j.

In a power system, the state vector $\boldsymbol{x}$ is usually obtained by performing a state estimation procedure using received redundant measurements and the topology information of the system. The weighted least squares estimation, which minimizes the cost function of $[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]^T \boldsymbol{R}^{-1} [\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]$, is widely used to obtain the estimated state vector $\hat{\boldsymbol{x}}$. By using weighted least squares estimation, the result of estimated state vector $\hat{\boldsymbol{x}}$ is shown in Equation (2). The operation symbol $\arg\min_{\boldsymbol{x}}(f(\boldsymbol{x}))$ simply returns the value of variable $\boldsymbol{x}$ which minimizes the cost function $f(\boldsymbol{x})$ over the set of candidates for $\boldsymbol{x}$.

$$\begin{aligned} \hat{\boldsymbol{x}} &= \arg\min_{\boldsymbol{x}} [\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]^T \boldsymbol{R}^{-1} [\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}] \\ &= (\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{z}. \end{aligned} \tag{2}$$

Based on the estimated state vector $\hat{\boldsymbol{x}}$ and the Equation (1), the estimated measurement vector $\hat{\boldsymbol{z}}$ is calculated as follows.

$$\hat{\boldsymbol{z}} = \boldsymbol{H}\hat{\boldsymbol{x}} = \boldsymbol{H}(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{z}. \tag{3}$$

Let $\boldsymbol{K} = \boldsymbol{H}(\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1}$, we can obtain that $\hat{\boldsymbol{z}} = \boldsymbol{K}\boldsymbol{z}$. Furthermore, the matrix $\boldsymbol{K}$ has the properties that $\boldsymbol{K} \cdot \boldsymbol{K} \cdot \boldsymbol{K} \cdots \boldsymbol{K} = \boldsymbol{K}$ and $\boldsymbol{K} \cdot \boldsymbol{H} = \boldsymbol{H}$ [28]. Therefore, the measurement residual $\boldsymbol{r}$ can be represented as follows:

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{z} - \hat{\boldsymbol{z}} = (\boldsymbol{I} - \boldsymbol{K})(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}) \\ &= (\boldsymbol{I} - \boldsymbol{K})\boldsymbol{e} = \boldsymbol{S}\boldsymbol{e}. \end{aligned} \tag{4}$$

Then, we can obtain that $\boldsymbol{S} = \boldsymbol{r}/\boldsymbol{e}$. $\boldsymbol{S}$ is called the residual sensitivity matrix, which represents the ratio of the measurement residual $\boldsymbol{r}$ to the measurement error $\boldsymbol{e}$. $\boldsymbol{I}$ is the identity matrix with dimensions of $m \times m$.

As for bad data, the commonly used detection and identification methods are the $J(\boldsymbol{x})$ detector and the largest normalized residual (LNR) detector, which are obtained by processing the measurement residual $\boldsymbol{r}$. The detection indicator of the $J(\boldsymbol{x})$ detector is defined in Equation (5).

$$J(\boldsymbol{x}) = \boldsymbol{r}^T \boldsymbol{R}^{-1} \boldsymbol{r}. \tag{5}$$

The objective function $J(\boldsymbol{x})$ approximately follows the χ^2 distribution with $m - n$ degrees of freedom [28]. The bad data is detected when the detection indicator exceeds the threshold τ for a given detection confidence with probability p, where

$p = Pr(J(\boldsymbol{x}) \leq \chi^2_{(m-n),p})$. We use $D_{J(\boldsymbol{x})}(z)$ to represent the $J(\boldsymbol{x})$ detector shown in Equation (6). The value 1 of $D_{J(\boldsymbol{x})}(z)$ indicates that FDIAs are launched, while the value 0 indicates that there is no attack.

$$D_{J(\boldsymbol{x})}(z) = \begin{cases} 1, & if\ J(\boldsymbol{x}) \geq \tau, \\ 0, & otherwise, \end{cases} \tag{6}$$

where $\tau = \chi^2_{(m-n),p}$.

The detection indicator of the LNR detector is defined in Equation (7). The $\boldsymbol{r}_i^*$ represents the normalized value of the measurement residual for measurement i.

$$\boldsymbol{r}_i^* = \frac{|\boldsymbol{r}_i|}{\sqrt{\boldsymbol{\Omega}_{ii}}}, \tag{7}$$

The $\boldsymbol{r}_i$ is the i-th residual in measurement residual $\boldsymbol{r}$. $\boldsymbol{\Omega}$ is the covariance of the measurement residual $\boldsymbol{r}$ and it can be calculated as $\boldsymbol{\Omega} = E[\boldsymbol{r} \cdot \boldsymbol{r}^T] = \boldsymbol{S}E[\boldsymbol{e} \cdot \boldsymbol{e}^T]\boldsymbol{S}^T = \boldsymbol{RS}$ with dimensions of $m \times m$. The symbol $E[\cdot]$ is the operation of expectation. $\boldsymbol{\Omega}_{ii}$ is the i-th diagonal entry in the measurement residual covariance matrix $\boldsymbol{\Omega}$ corresponding to $\boldsymbol{r}_i$. When the largest element of $\boldsymbol{r}_i^*$ exceeds the predefined threshold ζ, it implies that the i-th measurement z_i in the measurement vector z has been attacked by attackers. The detection result of the LNR detector can be expressed in Equation (8).

$$D_{LNR}(\boldsymbol{z}) = \begin{cases} 1, & if \max \boldsymbol{r}_i^* \geq \zeta, i \in [1, m] \\ 0, & otherwise. \end{cases} \tag{8}$$

where the value 1 of the $D_{LNR}(z)$ also indicates that FDIAs are launched, while the value 0 indicates that there is no attack.

The main feature of FDIAs is that attack vectors constructed by attackers are able to circumvent traditional bad data detection methods and can impose significant bias to estimated state variables. In the process of state estimation in the power system, FDIAs can be launched by eavesdropping and interfering communication links between smart meters and the control center, directly compromising some specific smart sensors, or hacking the control center and the database. It has been proved in [10] that the attack vector $\boldsymbol{a} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots \boldsymbol{a}_m]^T$ is called FDIAs if and only if the attack vector $\boldsymbol{a}$ can be represented as the linear combination of columns in the Jacobian matrix $\boldsymbol{H}$. In other words,

$$\boldsymbol{a} = \boldsymbol{Hc}, \tag{9}$$

where $\boldsymbol{c}$ is the injected deviation of the state vector $\boldsymbol{x}$. Therefore, the received attacked measurement vector $\boldsymbol{z_a}$ in the control center with an attack vector $\boldsymbol{a}$ can be expressed as follows:

$$\boldsymbol{z_a} = \boldsymbol{z} + \boldsymbol{a} = \boldsymbol{H}(\boldsymbol{x} + \boldsymbol{c}) + \boldsymbol{e} = \boldsymbol{Hx_a} + \boldsymbol{e}, \tag{10}$$

where $\boldsymbol{x_a}$ is the estimated state vector after attacks and it corresponds to the attacked measurement vector $\boldsymbol{z_a}$. The measurement residual $\boldsymbol{r_a}$ under FDIAs is presented as shown in Equation (11).

$$\begin{aligned} \|\boldsymbol{r_a}\|_2 &= \|\boldsymbol{z_a} - \boldsymbol{Hx_a}\|_2 = \|\boldsymbol{z} + \boldsymbol{a} - \boldsymbol{H}(\boldsymbol{x} + \boldsymbol{c})\|_2 \\ &= \|\boldsymbol{z} - \boldsymbol{Hx}\|_2. \end{aligned} \tag{11}$$

where the symbol $\|\cdot\|_2$ represents the operation of l_2 norm. Equation (11) shows that the attack vector $\boldsymbol{a}$ does not change the measurement residual $\boldsymbol{r}$ such that it can remain undetected by existing bad data detection methods in the state estimation. But the attack vector stealthy alters these estimated state variables.

III. Proposed Detection Methodology

It is well-known that state variables such as bus phase angles reflect the process operation statuses of power systems. Usually, it is quite difficult to directly measure voltage phase angles by smart meters. However, the canonical variate analysis (CVA) can obtain changes of state variables directly from the received redundant measurement vector of consecutive time slots. In other words, CVA can determine canonical variables with the greatest correlation by the measurement data of past and future [29]. Since estimated state variables after FDIAs are injected with deviations, directly monitoring the change of state variables can effectively detect FDIAs by the CVA. Meanwhile, the CVA is a data-based multivariate statistical process monitoring (MSPM) tool based on state-space model, which is suitable for dynamic process monitoring [30]–[32]. It considers not only the cross-correlation of measurements constrained by Kirchhoff's Law, but also the auto-correlation of measurements with time evolution. The auto-correlation of measurements is based on the fact that loads in power systems vary with the weather and temperature.

Our proposed CVA based detection method includes two stages. The first stage is a historical measurement-based training process. The second stage is a real-time detection process for FDIAs. The basic idea of the proposed method is shown in Fig. 1. The outline of the method is listed as follows.

- During the training stage, the main aim is to identify the maximum correlation between the past and future measurements and obtain parameters which are used for the real-time detection stage. There are five main steps that need to be processed as follows.
  - Step 1: Collected N consecutive measurement vectors are firstly normalized and the past and future Hankel matrixes are obtained.
  - Step 2: A scaled Hankel matrix is constructed.
  - Step 3: Singular value decomposition is performed and projection matrixes are obtained.
  - Step 4: Canonical state variables and detection statistics are computed.
  - Step 5: Based on these detection statistics, detection thresholds are determined for a given significance level.
- During detection stage, there are four steps to identify FDIAs in real-time.
  - Step 1: Real-time measurements are normalized by using means and variances and the past observation vector is constructed. The means and variances are obtained of the training stage.

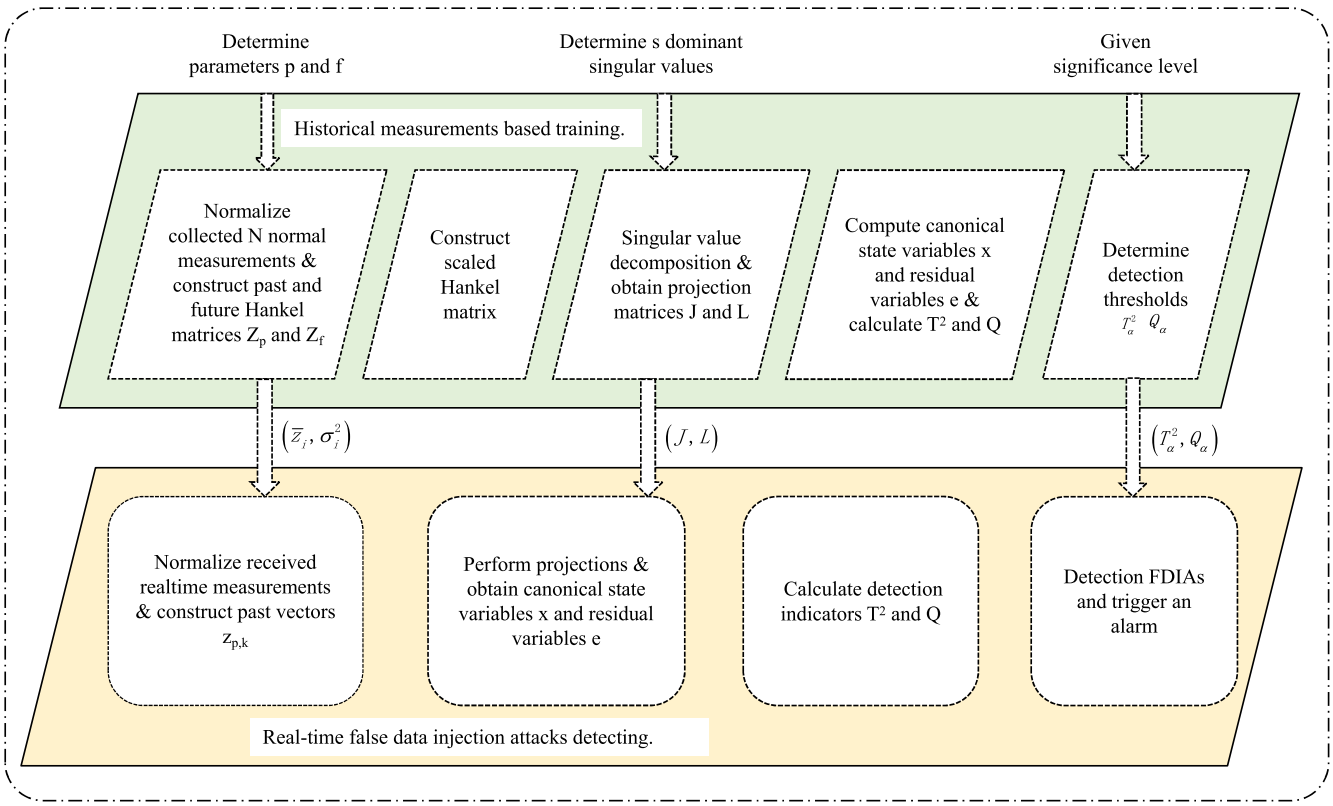


Fig. 1. The flowchart of the proposed CVA based detecting method against FDIAs.

- Step 2: Normalized real-time measurements are projected based on the projection matrixes to obtain state variables and residual variables.
- Step 3: Detection statistics T^2 and Q are calculated.
- Step 4: With the obtained detection thresholds in the training stage, alarms are triggered when FDIAs occur.

The specific process of each module in the flowchart is described and discussed in the subsequent subsections.

A. Training Stage, Step 1: Normalizing Measurements and Constructing Past and Future Hankel Matrixes

Assume that there are consecutive N collected measurement vectors under normal operating conditions for CVA training. We use $z_k^o = [z_{k,1}, z_{k,2}, \cdots z_{k,m}]^T$ to represent the original m-dimensional measurement vector at time k. All the measurements of consecutive N time slots can form a original measurement matrix Z^o, which can be represented as follows:

$$Z^o = \begin{bmatrix} z_{1,1} & z_{2,1} & \cdots & z_{k,1} & \cdots & z_{N,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{k,2} & \cdots & z_{N,2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ z_{1,j} & z_{2,j} & \ddots & z_{k,j} & \ddots & z_{N,j} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ z_{1,m} & z_{2,m} & \cdots & z_{k,m} & \cdots & z_{N,m} \end{bmatrix}, \tag{12}$$

where $k \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, m\}$. For simplicity, the matrix Z^o is represented as a vector group consisting of N column vectors, i.e., $Z^o = [z_1^o, z_2^o, \cdots z_k^o, \cdots z_N^o]$, where $k \in \{1, 2, \ldots, N\}$. To avoid the disturbance of measurements with larger absolute values, each variable in historical normal measurement vectors is normalized to zero mean and unit variance. Based on these N normal measurement vectors, the mean value $\bar{z}_j$ of the j-th variable in each measurement vector z_k^o and its corresponding variance σ_j^2 are obtained as follows:

$$\bar{z}_j = \frac{1}{N}\sum_{k=1}^{N} z_{k,j}, \quad \sigma_j^2 = \frac{1}{N-1}\sum_{k=1}^{N} (z_{k,j} - \bar{z}_j)^2,$$

where $j \in \{1, 2, \ldots, m\}$. Then, the measurement vector at time k is normalized as $z_k = [\frac{z_{k,1}-\bar{z}_1}{\sigma_1^2}, \frac{z_{k,2}-\bar{z}_2}{\sigma_2^2}, \cdots \frac{z_{k,m}-\bar{z}_m}{\sigma_m^2}]^T$. Finally, the original measurement matrix Z^o are represented as a new matrix $Z = [z_1, z_2, \cdots z_N]$.

At each time slot k, the measurement vector z_k is expanded by adding p past measurement vectors and f future measurement vectors in order to take into account time correlations. The generated augmented past observation vector $z_{p,k}$ and the future observation vector $z_{f,k}$ are represented as follows:

$$z_{p,k} = [z_{k-1}^T, z_{k-2}^T, \cdots z_{k-p}^T]^T, \tag{13}$$

$$z_{f,k} = [z_k^T, z_{k+1}^T, \cdots z_{k+f-1}^T]^T. \tag{14}$$

Note that the past observation vectors are formed into a $(mp) \times 1$ vector $(z_{p,k})$ while the future observation vectors are formed into a $(mf) \times 1$ vector $(z_{f,k})$. The above process of constructing the past and future observation vectors at each time slot k is shown in Fig. 2. The consecutive past p normal

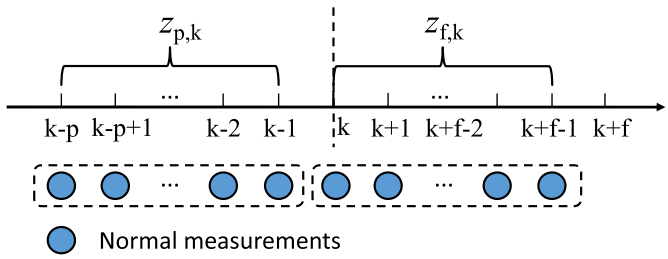


Fig. 2. The process of constructing the past and future observation vectors at each time slot k.

measurements before time slot k forms the past observation vector $z_{p,k}$ and the consecutive past f normal measurements after time slot k forms the future observation vector $z_{f,k}$.

Since there are N known historical measurement vectors which are used to perform the training, the value range of time k is $k \in \{1, 2, \ldots, N\}$. However, based on the constructed past observation vector $z_{p,k}$ in Equation (13), the last element is z_{k-p}^T in which the smallest subscript $k - p > 0$. Therefore, the time k needs to satisfy $k > p$, $k \in \mathbb{N}^*$. The symbol $\mathbb{N}^*$ represents the set of positive integers. Also, due to the last element in the constructed future observation vector $z_{f,k}$ in Equation (14) is z_{k+f-1}^T, the biggest subscript $k + f - 1$ needs to satisfy the condition $k + f - 1 \leq N$, $k \in \mathbb{N}^*$. As a result, the value range of time k is $p < k \leq N - f + 1$, $k \in \mathbb{N}^*$. Based on the description above, these future measurement vectors $z_{f,k}$ at each time k are also known in the training process. During the process of the training, meter measurements in a moving window of width $p + f$ are used at each time to construct past and future observation vectors, which are $z_{p,k}$ and $z_{f,k}$. In the constructed past and future observation vectors, the values of p and f are determined by autocorrelation analysis of the $z'_k = \sum_{j=1}^m z_{k,j}^2$, which is the square sum of each measurement at each time slot k. The autocorrelation function is calculated as follows:

$$\hat{\rho}(i) = \frac{\sum_{k=1}^{N-i} \left(\sum_{j=1}^m z_{k,j}^2 \sum_{j=1}^m z_{k+i,j}^2 \right)}{\sum_{k=1}^{N} \left(\sum_{j=1}^m z_{k,j}^2 \right)^2} \tag{15}$$

where parameter i is the time lag for observed measurement vectors and $i \in \{0, 1, 2, \ldots, N-1\}$. Based on the Equation (15), we can obtain these measurement-based autocorrelation function values about the number of time lags. The values of p and f are determined by the following equation:

$$p = f = \max\{i | \hat{\rho}(i) \geq \delta\}, \tag{16}$$

where the parameter δ is a given confidence bound.

The past and future Hankel matrixes Z_p and Z_f are formed by arranging the whole N observed measurement vectors in columns. The specific forms of the two matrixes are shown in Equation (17) and Equation (18).

$$Z_p = [z_{p,p+1}, z_{p,p+2}, \cdots z_{p,p+M}], \tag{17}$$

$$Z_f = [z_{f,p+1}, z_{f,p+2}, \cdots z_{f,p+M}], \tag{18}$$

where M is the number of columns of Hankel matrixes and it satisfies $M = N - p - f + 1$. In Equation (13) and Equation (14), since the range about subscript k in the past observation vector $z_{p,k}$ is $p < k \leq N - f + 1$, $k \in \mathbb{N}^*$, the first element of the past Hankel matrixes Z_p is $z_{p,p+1}$. Similarly, the first element of the future Hankel matrixes Z_f is $z_{f,p+1}$. The dimensions of the past Hankel matrix Z_p are $(mp) \times M$ and the dimensions of the future Hankel matrix Z_f are $(mf) \times M$. The above description is the first step of the historical measurements based training in Fig. 1. For the following real-time detection stage, the constructed past vector $z_{p,k}$ is also obtained in the same way as presented above.

B. Training Stage, Step 2: Constructing Scaled Hankel Matrix $\mathbf{\Phi}$

Based on the past and future Hankel matrixes Z_p and Z_f, the sample measurement-based covariance and cross-covariance matrixes $\mathbf{\Sigma}_{ff}$, $\mathbf{\Sigma}_{pp}$ and $\mathbf{\Sigma}_{fp}$ are defined as follows:

$$\mathbf{\Sigma}_{ff} = \frac{1}{M-1} Z_f Z_f^T,$$

$$\mathbf{\Sigma}_{pp} = \frac{1}{M-1} Z_p Z_p^T,$$

$$\mathbf{\Sigma}_{fp} = \frac{1}{M-1} Z_f Z_p^T.$$

The $\mathbf{\Sigma}_{ff}$ and $\mathbf{\Sigma}_{pp}$ represent the autocorrelation of the past and future Hankel matrixes Z_p and Z_f with time lags separately, while $\mathbf{\Sigma}_{fp}$ represents the cross-correlation between the Z_p and Z_f.

The process of constructing scaled Hankel matrix corresponds to the second step of the historical measurements based training in Fig. 1. The scaled Hankel matrix $\mathbf{\Phi}$ is defined as follows:

$$\mathbf{\Phi} = \mathbf{\Sigma}_{ff}^{-1/2} \mathbf{\Sigma}_{fp} \mathbf{\Sigma}_{pp}^{-1/2}, \tag{19}$$

where the dimensions of $\mathbf{\Phi}$ are $(mf) \times (mp)$, the negative power operation of a matrix means the inverse matrix after the power operation of the original matrix, and any matrix B satisfying $B^2 = A$ is known as the square root matrix of A, denoted as $A^{1/2}$ [34]. The purpose of the canonical variate analysis is to find the linear combinations that maximize the correlation between the past and future Hankel matrixes. As proved in [30], it can be realized by performing singular value decomposition (SVD) of the obtained scaled Hankel matrix $\mathbf{\Phi}$ and the correlation is the corresponding singular value of $\mathbf{\Phi}$.

C. Training Stage, Step 3: Singular Value Decomposition and Projection Matrixes

Let r denote the rank of the scaled Hankel matrix $\mathbf{\Phi}$, i.e., $rank(\mathbf{\Phi}) = r$. Based on the Theorem 5.1 (SVD) in [34], by performing SVD of $\mathbf{\Phi}$, the following equation can be obtained:

$$\mathbf{\Phi} = UDV^T, \tag{20}$$

where U and V are the left and right singular column vectors, which can be represented as $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{(mf)}]$ and $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{(mp)}]$, respectively. Each $\boldsymbol{u}_i$ in U represents a column vector with the dimension of (mf), $i \in [1, 2, \ldots, mf]$. Similarly, each $\boldsymbol{v}_j$ in V represents a column vector with the dimension of (mp), $j \in [1, 2, \ldots, mp]$. D is the obtained diagonal matrix in which each of its elements represents the correlation between the corresponding column vectors of U and V. The diagnal matrix D can be represented as:

$$D = \begin{bmatrix} diag(\lambda_1, \lambda_2, \ldots, \lambda_r) & \mathbf{0}_{r\times(mp-r)} \\ \mathbf{0}_{(mf-r)\times r} & \mathbf{0}_{(mf-r)\times(mp-r)} \end{bmatrix},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are the singular values and they satisfy $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. The symbol $\mathbf{0}$ represents zero matrix, in which all the elements are zeros.

The number of nonzero singular values is r. In order to realize the dimension reduction, only the first s singular values are used. These s singular values are called dominant singular values [35]. The value of s is determined by the following equation:

$$s = \max\{i | \lambda_i \geq \varphi, 0 < \varphi < \lambda_1, i \in \{1, 2, \cdots r\}\}, \quad (21)$$

where φ is a given threshold.

By utilizing the $V_s = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s]$, projection matrixes J and L to achieve matrix transformation are obtained as follows.

$$J = V_s^T \boldsymbol{\Sigma}_{pp}^{-1/2}, \quad (22)$$

$$L = (I - V_s V_s^T)\boldsymbol{\Sigma}_{pp}^{-1/2}. \quad (23)$$

The dimensions of projection matrix J are $s \times (mp)$ and the dimensions of projection matrix L are $(mp) \times (mp)$. The above content is the third step of the historical measurements based training in Fig. 1. These two projection matrixes J and L are then retained and used during the detection stage.

Since generating the projection matrixes in the training process is off-line, the computation complexity of those cross covariance matrixes and the SVD is not an issue. During the online detection process, the projection matrixes are directly used to compute the canonical state variables by using the received measurements.

D. Training Stage, Step 4: Computing Canonical State Variables and Detection Indicators

Canonical variables can be estimated from the past measurement vector $\boldsymbol{z}_{p,k}$. A canonical state variable $\boldsymbol{x}$ corresponds to these voltage phase angles in the power grid. A residual variable $\boldsymbol{w}$ stands for the small noise during the projection process. The canonical state variable $\boldsymbol{x}$ and the residual variable $\boldsymbol{w}$ at time slot k are calculated as follows:

$$\boldsymbol{x}_k = J\boldsymbol{z}_{p,k}, \quad (24)$$

$$\boldsymbol{w}_k = L\boldsymbol{z}_{p,k}. \quad (25)$$

The dimensions of $\boldsymbol{x}_k$ are $s \times 1$. Therefore, through the transformation of projection matrix J, the mp-dimensional past measurement vector $\boldsymbol{z}_{p,k}$ can be reduced to the s-dimensional $\boldsymbol{x}_k$, which has realized the dimension reduction. Similarly, the dimensions of $\boldsymbol{w}_k$ are $(mp) \times 1$. At each time slot, all elements in the canonical state variable $\boldsymbol{x}$ forms a state subspace and all elements in the residual variable $\boldsymbol{w}$ forms a residual subspace.

Finally, the Hotelling T^2 statistic and the Q statistic are both used as attack detection indicators to monitor variations of state variables and residual variables at each time slot. The Hotelling T^2 statistic is usually constructed to test whether the mean value of a random variable is equal to the expected value. The detection indicator Q is used to monitor the variations of the sum of square errors in the residual subspace, which is the complement of state space at each time slot. Q statistic can capture unobservable variable changes in the T^2 statistic and they are complementary.

Usually, for the canonical state variable $\boldsymbol{x}$ of M samples with normal distribution, the T^2 statistic is calculated as follows:

$$\boldsymbol{T}_k^2 = (\boldsymbol{x}_k - \bar{\boldsymbol{x}})^T \boldsymbol{\Psi} (\boldsymbol{x}_k - \bar{\boldsymbol{x}}), \quad (26)$$

where the symbol $\bar{\boldsymbol{x}}$ is the expected value of the canonical state variable $\boldsymbol{x}$ and $\bar{\boldsymbol{x}} = \frac{1}{M}\sum_{k=1}^{M} \boldsymbol{x}_k$. $\boldsymbol{\Psi}$ is the covariance of canonical state variable $\boldsymbol{x}$. It can be deduced that the following equation holds:

$$\begin{aligned} \boldsymbol{\Psi} &= \frac{1}{M-1}\sum_{k=1}^{M} \boldsymbol{x}_k \boldsymbol{x}_k^T \\ &= \frac{1}{M-1}\sum_{k=1}^{M} V_s^T \boldsymbol{\Sigma}_{pp}^{-1/2} \boldsymbol{z}_{p,k} \boldsymbol{z}_{p,k}^T \boldsymbol{\Sigma}_{pp}^{-1/2} V_s \\ &= V_s^T \boldsymbol{\Sigma}_{pp}^{-1/2} \left(\sum_{k=1}^{M} \frac{1}{M-1} \boldsymbol{z}_{p,k} \boldsymbol{z}_{p,k}^T \right) \boldsymbol{\Sigma}_{pp}^{-1/2} V_s \\ &= V_s^T \boldsymbol{\Sigma}_{pp}^{-1/2} \boldsymbol{\Sigma}_{pp} \boldsymbol{\Sigma}_{pp}^{-1/2} V_s = V_s^T V_s = I. \end{aligned} \quad (27)$$

Equation (27) means that the covariance of $\boldsymbol{x}$ is an identity matrix. At the same time, since the generated past and future observation vectors are normalized, the expected value $\bar{\boldsymbol{x}}$ of is zero. Therefore, the T_k^2 detection indicator is simplified as follows:

$$\boldsymbol{T}_k^2 = \boldsymbol{x}_k^T \boldsymbol{x}_k. \quad (28)$$

To monitor the variations of the sum of square errors in the residual subspace, which is the complement of state space at each time slot, the detection indicator Q is defined as follows:

$$\boldsymbol{Q}_k = \boldsymbol{w}_k^T \boldsymbol{w}_k. \quad (29)$$

Once the detection indicators of state variables are obtained from received real-time measurement vector, they are compared with detection thresholds to detect if false data injection attacks occur during the online detection stage. The above process described the fourth step of the historical measurements based training in Fig. 1.

E. Training Stage, Step 5: Determining Detection Thresholds

As for detection thresholds in the training stage, we assume that the process noise and measurement noise in the power grid follow Gaussian distributions. Therefore, the state variables and measurements are also Gaussian distributions. The detection thresholds are determined by the given confidence level of the Gaussian distribution. For a given significance level α, the detection thresholds for T^2 statistic and Q statistic are represented as T_α^2 and Q_α such that $P(T^2 < T_\alpha^2) = \alpha$ and $P(Q < Q_\alpha) = \alpha$. Based on the Theorem 2.16 in [36], the detection thresholds T_α^2 and Q_α are defined as follows:

$$T_\alpha^2 = \frac{s(M-1)(M+1)}{M(M-s)} F_\alpha(s, M-s), \tag{30}$$

$$Q_\alpha = \gamma_1 \left[\frac{c_\alpha (2\gamma_2 h_0^2)^{1/2}}{\gamma_1} + 1 + \frac{\gamma_2 h_0 (h_0 - 1)}{\gamma_1^2} \right]^{1/h_0}, \tag{31}$$

where $F_\alpha(s, M-s)$ is the upper percentile of F-distribution with s and $M-s$ degrees of freedom given α significance level. The F-distribution is a probability density function, which is defined as a ratio of the variances about two normally distributed random variables. For more details about the F-distribution, please refer to [36]. c_α is the normal deviate corresponding to the upper $1-\alpha$ percentile. $\gamma_i = \sum_{j=s+1}^{r} \sigma_j^{2i}$ for $i = 1, 2, 3$ and γ is the function of the eigenvalue about the covariance of past measurement vector. $h_0 = 1 - 2\gamma_1\gamma_3/(3\gamma_2^2)$. The σ_j is the j-th eigenvalue about the covariance of past measurement vector. The above description is the fifth step of the historical measurements based training in Fig. 1. These obtained detection thresholds are then retained.

F. Detection Stage: Real-Time FDIAs Detection

Step 1: Each element in the received real-time measurement vector $z_k^\dagger = [z_{k,1}, z_{k,2}, \cdots z_{k,m}]^T$ is firstly normalized by the retained means $\bar{z}_j$ and variances σ_j^2, which are in the first step of the training stage and $j \in \{1, 2, \ldots, m\}$. After that, the past observation vector is construct in the same manner of Equation (13).

Step 2: The two retained projection matrixes J and L are directly used to compute the canonical state variables x_k and the residual variables w_k. This process is the same with Equations (24) and (25).

Step 3: The constructed detection indicators T^2 and Q are also obtained in the same way as presented in Equations (28) and (29). The detection thresholds retained in the training stage are directly used in the real-time detection stage.

Step 4: If the detection indicator T_k^2 is larger than the detection threshold T_α^2 or the detection indicator Q_k is larger than the detection threshold Q_α at time slot k, the FDIAs to the measurement vector $z_k^\dagger$ are detected.

IV. Performance Evaluation

To verify the performance of our proposed method, we evaluate it using both synthetically generated data and real-world electricity data. We first show that traditional measurement residual-based $J(\mathrm{x})$ detector and LNR detector cannot detect FDIAs when the estimated state variables are injected with bias. Then, the effectiveness and accuracy of the proposed CVA-based method are illustrated and presented.

A. Simulation Setup

The test system for validating the performance of proposed detection method is based on the IEEE 14-bus system. There are 14 buses and 20 branches with a total load demand of 259MW. We consider the full measurement system, in which these SCADA measurements include 14 bus injected active power measurements, 20 active power flow measurements of branches at the from-end and, 20 active power flow measurements of branches at the to-end. The from-end and to-end of a power transmission branch mean the power flow sending and receiving ends, respectively. Therefore, there are 54 SCADA measurements at one specific time slot. For simulating the continuous system operation process of the power system, load demands at every load bus may dynamically change and generation in the system also varies in order to keep the balance of power.

In this paper, to verify the detection performance of our proposed CVA-based method, we consider two different cases, which are based on the synthetically generated load data in Case 1 and the real-world electricity load data in Case 2, respectively. In Case 1, in order to generate the synthetically generated data, loads on each bus are supposed to be uniformly distributed between 80% and 120% of its base load for 1000 consecutive time slots. In Case 2, for the purpose of generating real-world electricity measurement data, loads used are adopted from the New York independent system operator (NYISO) [27]. These load data are online load flow profiles for 11 regions recorded every five minutes. The loads data used are between January 1, 2016 and January 7, 2016. We link each load bus of the test system with the normalized load data of one region of NYISO.

Afterwards, the measurement vectors at each time slot of both experimental cases are collected and obtained by running DC power flow analysis procedures from the MATPOWER Toolbox [33]. At the same time, in order to mimic the effect of measurement random errors of meters, the Gaussian white noises are added to the measurement vectors with the signal noise ratio (SNR) of 20 dB. The redundancy of measurements under our simulation condition is 3.86, which is computed as the ratio of the number of measurements to the number of state variables. For generating the false data injection attacks data, we adopt the targeted FDIAs [10], in which attackers intend to inject specific errors into the estimation of certain chosen state variables. The targeted FDIAs in which only a certain number of state variables are polluted in one specific region has greater potential harm to power system. We have simulated 5% incremental FDIAs on system state θ_2, which is the phase angle of bus 2. All measurements related to bus 2 are replaced with falsified measurements, which are generated by the means in Section II. All experiments are conducted in Matlab R2017a

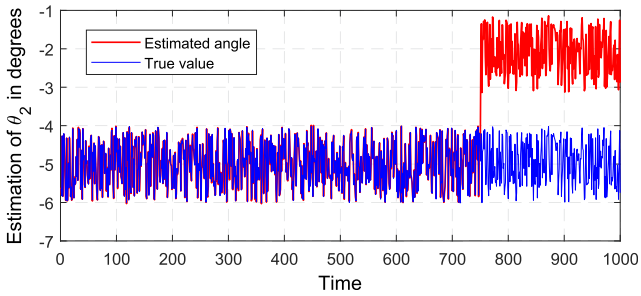


Fig. 3. Estimation result of θ_2 by weighted least square under FDIA.

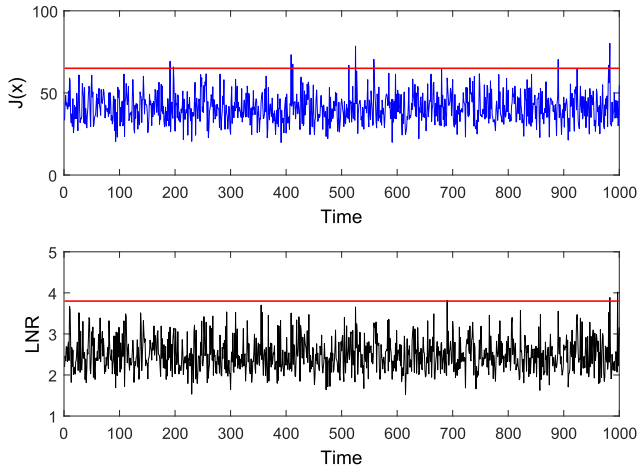


Fig. 4. Detection performance of $J(x)$ and LNR detectors under FDIA.

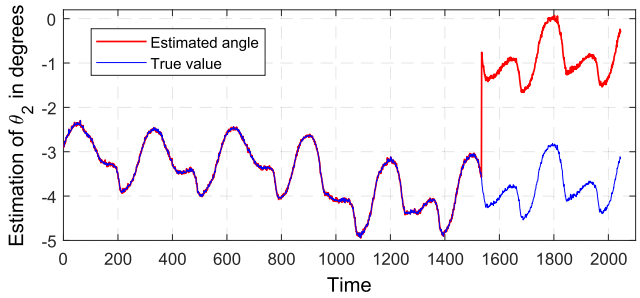


Fig. 5. Estimation result of θ_2 by weighted least square under FDIA.

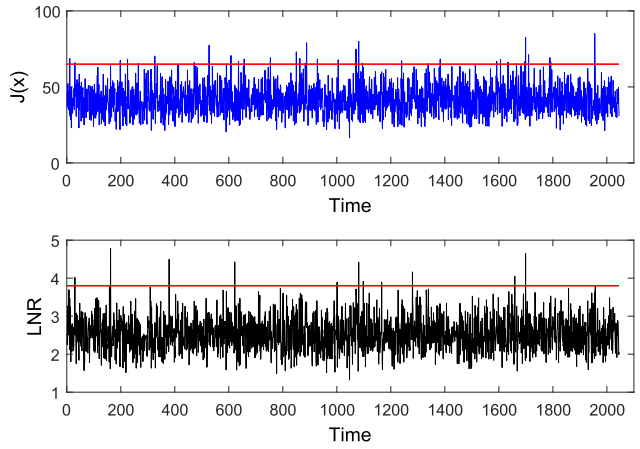


Fig. 6. Detection performance of $J(x)$ and LNR detectors under FDIA.

environment with Lenovo laptop of 2.20 GHz Intel Core i5 processor.

B. Detection Performance

The detection performance of our proposed canonical variate analysis based method is evaluated by checking whether it is able to identify FDIAs in obtained datasets. For the first scenery, the FDIA is launched at time T=750 and it lasts until time T=1000. For the second scenery, there are 2045 measurement vectors in total, the FDIA is launched at time T=1535, and it lasts until time T=2045. In both cases, the strength of FDIAs is simulated of 5 percent increment of its original values, i.e., the manipulated state variable of θ_2 is 5% bigger than its true value. Since the indicator of $J(x)$ detector follows χ^2 distribution and the degrees of freedom is $m - n = 40$, we have the significance level $\alpha = 0.01$ and the threshold of $J(x)$ detector is set as 63.69. The detection threshold of LNR detector is set as 3.8, which is learnt from historical measurement statistics.

It can be seen from Fig. 3 in Case 1, the estimated phase angles of bus 2 closely follows the true values when there is no attacks. When the FDIA occurs, the estimated state variable θ_2 deviates directly from its true value and the injected bias is about 2.87 degrees. However, the FDIA during the attack period is undetected by commonly used $J(x)$ detector and LNR detector, in which the detection statistics are almost smaller than thresholds as shown in Fig. 4. This indicates that the received false measurement vectors are regarded as normal in control center. Moreover, the estimated incorrect state variables are used as inputs for subsequent EMS application modules. Then the control center produces wrong control commands and decisions.

As for the real-world electricity data of one week generated by loads from NYISO in Case 2, Figs. 5 and 6 also show the similar results as in Case 1. When the FDIA is launched at time T=1535, the estimated state variable θ_2 deviates suddenly from its true value, but the detectors do not alarm any abnormalities. However, the traditional measurement residual-based $J(x)$ detector and the LNR detector cannot detect FDIAs when the estimated state variables are injected with bias.

The proposed method is firstly trained based on normal historical measurement vectors to generate the projection matrixes and detection thresholds. Note that the dynamics of power grids vary with the changes of loads.

Therefore, the obtained measurement vectors are temporally correlated during the consecutive observation period. For simplicity, to consider more realistic load changes in power grid, we only present the detail analysis procedure of the proposed detection method in Case 2. We use the first 1000 normal measurement vectors to perform the offline training. At the beginning, the number of past and future measurement vectors p and f are determined by analyzing time correlations of the square sum of all normalized measurement vectors [32]. As shown in Fig. 7, there are total 1000 training data of

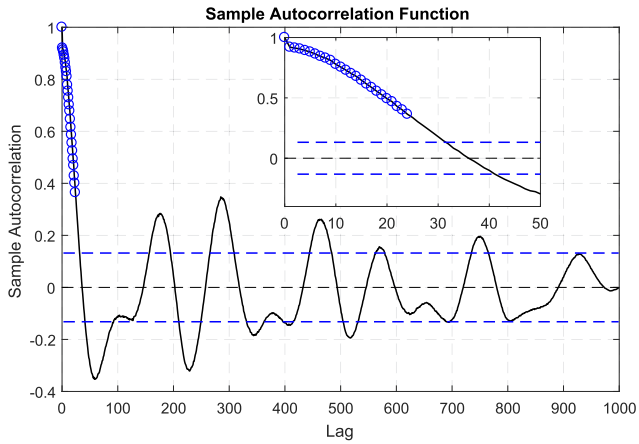


Fig. 7. Sample autocorrelation function of the number of time lags for observed measurement vectors.

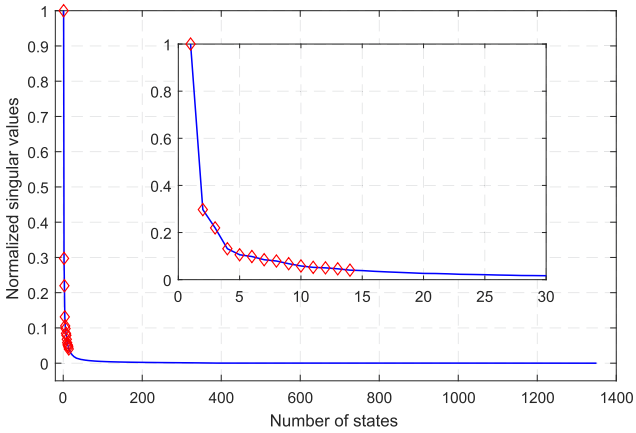


Fig. 8. Normalized singular values from the scaled Hankel matrix.

observed measurement vectors. We observe that the sample autocorrelation decreases rapidly with the number of time lags and fluctuates up and down near zero. For this case study, when the time lags are greater than 25, the autocorrelations fluctuate up and down basically within a confidence bound of $\pm 15\%$. Therefore, for the given confidence bound, the maximum time lag is set as 25 and then parameters p and f are set to be 25 for this study. Also, the selection about the number of time lags can affect the dimensions of constructed covariance matrixes and scaled Hankel matrix. Consequently, the calculation speed of the training algorithm and the number of canonical state variables can also be affected.

Then, the length of the past and future observation vectors for processing detection is 1350 and the number of columns of Hankel matrixes is $M = 951$. According to the proposed detection method described, the singular value decomposition is then performed on the scaled Hankel matrix in order to estimate the number of state variables. The most popular method for calculating the optimal number of state variables is based on the dominant singular values in the diagonal matrix $\boldsymbol{D}$. Usually, since there are noise in the measurements, the diagonal matrix is typically full rank. It can be seen from Fig. 8 in which the whole non-zero elements of matrix $\boldsymbol{D}$ are presented.

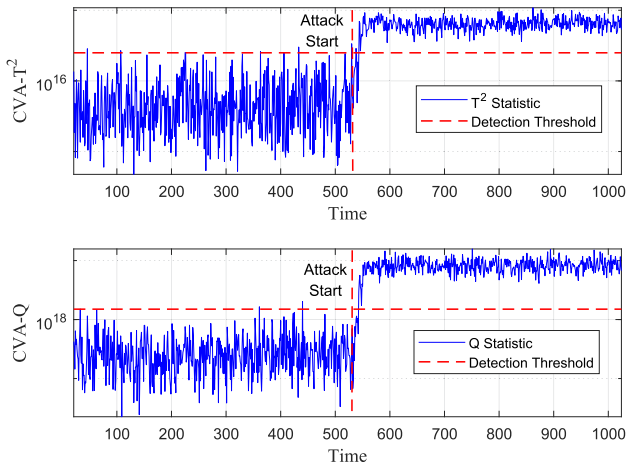


Fig. 9. Detection performance against FDIA towards voltage phase angle of bus 2 in Case 2.

Clearly, the first 14 dominant singular values have contained the most characteristics of observed measurements. The remaining singular values are very small and approximately equal to zero. In IEEE 14-bus power system, the whole state variables are phase angles of buses and this means that there are 14 state variables. Therefore, it is satisfactory that the reserved number of dominant singular values is set as $s = 14$. At the same time, the remaining singular values are the residual variables in the subspace. The CVA model is then trained to obtain the projection matrixes and the detection thresholds for a given significance level.

Fig. 9 shows the detection performance results of our proposed detection method against FDIAs targeted to voltage phase angle of bus 2. The last 1045 measurement vectors are used for FDIAs detection. The attacks are launched at T=535. Before the FDIAs starts, both $\boldsymbol{T}^2$ and $\boldsymbol{Q}$ statistics are fluctuated with time, but they are almost below the detection thresholds. After the FDIAs occur, only after a very short time delay, the $\boldsymbol{T}^2$ and $\boldsymbol{Q}$ statistic indicators exceed their thresholds. In other words, the FDIAs can be effectively detected. Also, the missed detection rate, which is defined as the ratio of the number of undetected samples under attacks to the total number of attacked samples, is used to describe the detection accuracy. The missed detection rate of $\boldsymbol{T}^2$ statistic is 0.73% and the missed detection rate of $\boldsymbol{Q}$ statistic is 1.28%. Since the generated past and future observation vectors are constructed by the measurements in a moving window, the correlation of these measurements is not significantly changed when FDIAs have just occurred. However, as the window slides backward, the correlation changes greatly and the detection index obviously exceeds the threshold. Meanwhile, we can conclude that the canonical state variables and the residual variables are both sensitive to FDIAs. It is because the cross-correlations and auto-correlations of observed measurements are both affected after attacks, and the influence can be effectively captured through canonical variate analysis.

The detection performance of Case 1 is shown in Fig. 10. We use 500 normal historical measurement vectors for the training, in which the time lags p and f are set as 10. The

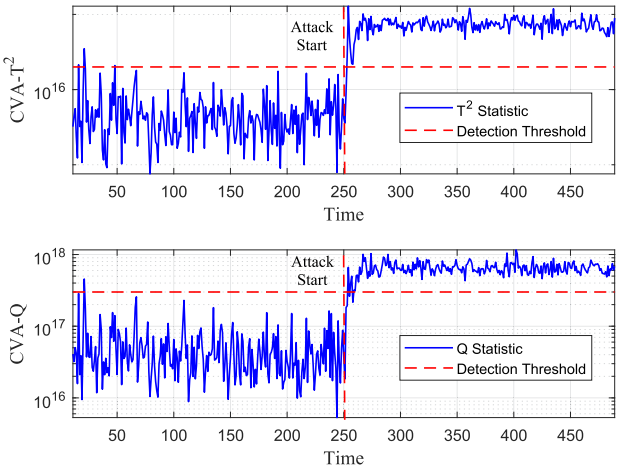


Fig. 10. Detection performance against FDIA towards voltage phase angle of bus 2 in Case 1.

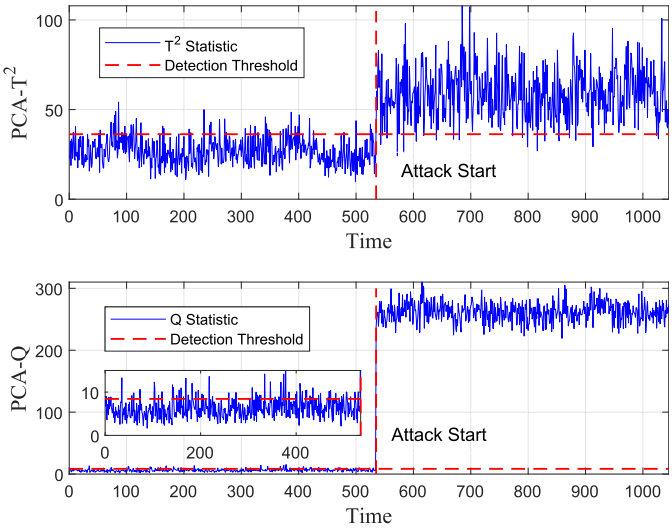


Fig. 11. PCA-based method detection performance towards voltage phase angle of bus 2 in Case 2.

number of columns of Hankel matrixes is $M = 481$. The order of the system which corresponds to the number of canonical state variables for the CVA-basd detection method is also set as 14. The last 500 measurement vectors are used for FDIAs detection. The attacks are launched at $T = 251$. It can be seen that under the steady operation of power system under uniformly varying loads, both T^2 and Q statistical indicators can effectively detect the occurrence of FDIAs. There is basically no detection delay and the missed detection rate is almost zero. Meanwhile, both T^2 and Q statistical indicators are sensitive to the attacks. Usually, due to the reason that if the system variations cannot be captured in state variable subspace, it will be captured by the residual subspace. Therefore, we adopt both T^2 and Q statistical indicators for FDIAs detection.

To compare the performance of the proposed CVA-based method with the PCA-based detection algorithm in [37], the detection result of PCA-based detection algorithm using real-world electricity load data in Case 2 is presented in Fig. 11. In order to meet the rationality of experimental comparison, voltage phase angle of bus 2 is attacked by FDIAs and the settings of launched FDIAs are the same as Case 2. For the PCA-based detection algorithm, the largest L singular values of the measurement matrix is remained such that $(\sum_{i=1}^{L} s_i / \sum_{i=1}^{T} s_i) = 90\%$ is satisfied, where s represents the singular value of measurement matrix and T is the number of total singular values. Based on the Fig. 11, we observe that T^2 and Q statistics of the PCA-based method have more false alarms. Also, although the number of missing detections of PCA-T^2 is large, the number of missing detections of PCA-Q is very rare. It shows that the detection performance of PCA-Q statistics is better than the PCA-T^2. We define the false alarm rate (FAR) and miss detection rate (MDR) as follows.

$$FAR = \frac{N_{hit}}{N_{hit} + N_{miss}},$$

$$MDR = \frac{N_{false}}{N_{false} + N_{correct}},$$

where N_{hit} represents the number of successful detections of attacks, N_{miss} represents the number of miss detections of attacks, N_{false} represents the number of false alarms, and $N_{correct}$ represents the number of correct reports of no attack. To quantify the comparison of experimental results, the performance comparison of three different detection methods about Case 1 and Case 2 is presented in Table I. It shows that the residual-based $J(x)$ detector and LNR detector can hardly detect the occurrence of FDIAs. As for the PCA-based detector, the MDRs for both Case 1 and Case 2 are very low. The FAR for Case 2 of the PCA-based detector is higher than the FAR for Case 1, mainly because the fluctuation of real-world electricity load data is relatively larger than that of uniform load changes in Case 1. However, compared with results of the proposed CVA-based detection method, the FARs of PCA-based detector for both Case 1 and Case 2 are higher. This is because only the cross-correlation constrained by Kirchhoff's Law is considered and the auto-correlation of consecutive measurements in the PCA method is neglected. Therefore, based on the above analysis and comparison, experiments show that our proposed CVA-based detection algorithm has better detection performance.

V. Conclusion and Future Work

Since FDIAs have the capability to successfully bypass conventional residual-based detections and can stealthily cause erroneous state estimation results in smart grid, the subsequent control and operations can be severely affected. However, most existing FDIAs detection methods only consider cross-correlation of discrete measurement vectors at each current sampling time, which are constrained by the Kirchhoff's Law. The auto-correlation of state variables and measurements for a consecutive time slots is usually neglected. In this paper, a canonical variate analysis based new detection method is proposed for detecting FDIAs. The proposed canonical variate analysis based detection method captures both the cross-correlations and the auto-correlations of observed measurements

TABLE I
PERFORMANCE COMPARISON OF THREE DIFFERENT DETECTION METHODS ABOUT CASE 1 AND CASE 2

Detection Method	Detection Statistic	Case 1		Case 2	
		FAR(%)	MDR(%)	FAR(%)	MDR(%)
CVA-based	T^2	0.80	0.20	1.68	0.73
	Q	0.40	0.80	0.74	1.28
PCA-based [37]	T^2	3.20	0.00	11.59	4.31
	Q	8.80	0.00	14.39	0.59
Residual-based	$J(x)$	1.20	99.00	1.76	97.84
	LNR	0.00	99.40	0.72	99.02

through the least number of canonical state variables. Then, FDIAs are effectively detected by checking the statistical consistence of correlations before and after attacks. Extensive results demonstrate the accuracy of the proposed method. In the future work, the detection sensitivity towards different attacked state variables in smart grid can also be analyzed.

REFERENCES

[1] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. L. P. Chen, "A survey of communication/networking in smart grids," *(Elsevier) Future Gener. Comput. Syst.*, vol. 28, no. 2, Feb. 2012, pp. 391–404.

[2] R. Fang, J. Wang, and W. Sun, "Cross-layer control of wireless sensor network for smart distribution grid," *Int. J. Sensor Netw.*, vol. 27, no. 2, 2018, pp. 71–84.

[3] J. Liu, Y. Xiao, S. Li, W. Liang, and C. L. P. Chen, "Cyber security and privacy issues in smart grids," *IEEE Commun. Surv. Tut.*, vol. 14, no. 4, pp. 981–997, Oct.–Dec. 2012.

[4] K. Cheena, T. Amgoth, and G. Shankar, "Emperor penguin optimised self-healing strategy for WSN based smart grids," *Int. J. Sensor Netw.*, vol. 32, no. 2, 2020, pp. 87–95.

[5] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1636–1646, May 2018.

[6] D. Liu, J. Shen, A. Wang, and C. Wang, "Lightweight and practical node clustering authentication protocol for hierarchical wireless sensor networks," *Int. J. Sensor Netw.*, vol. 27, no. 2, 2018, pp. 95–102.

[7] A. Naiara, M. Elias, L. Jesus, J. Eduardo, and A. Armando, "Cyber-security in substation automation systems," *Renewable Sustain. Energy Rev.*, vol. 54, pp. 1552–1562, Feb. 2016.

[8] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, Jul. 2017.

[9] 2019. [Online]. Available: https://en.wikipedia.org/wiki/2019_Venezuelan_blackouts

[10] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, 2009, pp. 21–32.

[11] E. McCary and Y. Xiao, "Malicious device inspection home area network in smart grids," *Int. J. Sensor Netw.*, vol. 25, no. 1, Sep. 2017, pp. 45–62.

[12] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 612–621, Mar. 2014.

[13] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *Proc. Preprints First Workshop Secure Control Syst.*, Apr. 2010, pp. 1–6.

[14] Z. Yu and W. Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.

[15] S. Bi and Y. J. Zhang, "Defending mechanisms against false-data injection attacks in the power system state estimation," in *Proc. IEEE GLOBECOM Workshops*, 2011, pp. 1162–1167.

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

[17] Q. Yang, D. An, R. Min, W. Yu, X. Yang, and W. Zhao, "On optimal PMU placement-based defense against data integrity attacks in smart grid," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 7, pp. 1735–1750, Jul. 2017.

[18] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Limiting false data attacks on power system state estimation," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, 2010, pp. 1–6.

[19] J. Jiang and Y. Qian, "Defense mechanisms against data injection attacks in smart grid networks," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 76–82, Oct. 2017.

[20] J. Zhao, G. Zhang, M. La Scala, Z. Y. Dong, C. Chen, and J. Wang, "Short-term state forecasting-aided method for detection of smart grid general false data injection attacks," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1580–1590, Jul. 2017.

[21] D. B. Rawat and C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1652–1656, Oct. 2015.

[22] P. Lynggaard, "Controlling interferences in smart building IoT networks using machine learning," *Int. J. Sensor Netw.*, vol. 30, no. 1, 2019, pp. 46–55.

[23] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.

[24] L. Q. Yang, Y. C. Li, and Z. J. Li, "Improved-ELM method for detecting false data attack in smart grid," *Int. J. Elect. Power Energy Syst.*, vol. 91, pp. 183–191, Oct. 2017.

[25] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.

[26] M. Mohammadpourfard, A. Sami, and A. R. Seifi, "A statistical unsupervised method against false data injection attacks: A visualization-based approach," *Expert Syst. Appl.*, vol. 84, pp. 242–261, Oct. 2017.

[27] NYISO, "Load data profile," 2020. [Online]. Available: http://www.nyiso.com

[28] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. 1st ed. New York, NY, USA: Marcel Dekker, Mar. 2004.

[29] B. C. Juricek, D. E. Seborg, and W. E. Larimore, "Fault detection using canonical variate analysis," *Ind. Eng. Chemistry Res.*, vol. 43, no. 2, pp. 458–474, Jan. 2004.

[30] P. P. Odiowei and Y. Cao, "Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 36–45, Feb. 2010.

[31] K. E. S. Pilario and Y. Cao, "Canonical variate dissimilarity analysis for process incipient fault detection," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5308–5315, Dec. 2018.

[32] C. Ruiz-Carcel, L. Lao, Y. Cao, and D. Mba, "Canonical variate analysis for performance degradation under faulty conditions," *Control Eng. Practice*, vol. 54, pp. 70–80, Sep. 2016.

[33] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[34] X. Zhang, *Matrix Analysis and Applications*, Cambridge, U.K.: Cambridge Univ. Press, 2017.

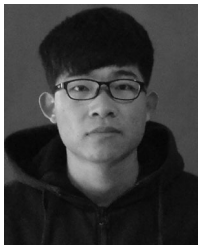
[35] A. Negiz and A. Cinar, "Statistical monitoring of multivariable dynamic processes with state-space models," *AIChE J.*, vol. 43, no. 8, pp. 2002–2020, 1997.

[36] K. Knight, *Mathematical Statistics*. London, U.K.: Chapman & Hall, 2000.

[37] A. S. Musleh, M. Debouza, H. M. Khalid, and A. Al-Durra, "Detection of false data injection attacks in smart grids: A real-time principle component analysis," in *Proc. 45th Annu. Conf. IEEE Ind. Electron. Soc.*, 2019, pp. 2958–2963.

[38] A. Y. Lu and G. H. Yang, "False data injection attacks against state estimation in the presence of sensor failures," *Inf. Sci.*, vol. 508, pp. 92–104, Jan. 2020.

[39] B. Li, T. Ding, C. Huang, J. Zhao, Y. Yang, and Y. Chen, "Detecting false data injection attacks against power system state estimation with fast go-decomposition approach," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2892–2904, May 2019.



Chao Pei is currently pursuing the Ph.D. degree in control theory and control engineering at the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. He is currently visiting the Department of Computer Science of The University of Alabama as a Ph.D. student. His research interests are in cyber physical security of smart grid, power system state estimation, PMU deployment, and signal processing.

Yang Xiao (Fellow, IEEE) is currently a Professor with the Department of Computer Science at the University of Alabama, Tuscaloosa, AL, USA. His current research interests include cyber-physical systems, Internet of Things, security, wireless networks, smart grid, and telemedicine. He has published over 290 journal papers (including over 50 IEEE/ACM transactions papers) and over 200 conferetnce papers. Dr. Xiao was a Voting Member of IEEE 802.11 Working Group from 2001 to 2004, involving IEEE 802.11 (WIFI) standardization work. He is an IEEE Fellow and an IET Fellow. He currently serves as Editor-in-Chief for *Cyber-Physical Systems (Journal)*. He had (s) been an Editorial Board or Associate Editor for 20 international journals, including IEEE TRANSACTIONS ON CYBERNETICS, during 2020 to now, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, during 2014 to 2015, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, during 2007 to 2009, and IEEE Communications Survey and Tutorials, during 2007 to 2014. He served (s) as a Guest Editor for over 20 times for different international journals, including IEEE Network, IEEE Wireless Communications, and ACM/Springer Mobile Networks and Applications (MONET).

Wei Liang (Senior Member, IEEE) is currently a Professor with the Shenyang Institute of Automation, Chinese Academy of Sciences. She received the Ph.D. degree in mechatronic engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2002. As a Primary Participant/a Project Leader, she developed the WIA-PA and WIA-FA standards for industrial wireless networks, which are specified by IEC 62601 and IEC 62948, respectively. Her research interests include industrial wireless sensor networks and wireless body area networks. She received the International Electrotechnical Commission 1906 Award in 2015 as a Distinguished Expert of industrial wireless network technology and standard.

Xiaojia Han is currently pursuing the Ph.D. degree in measurement technology and automatic equipment at the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. Her research interests include sensor and actuator fault detection methods, smart grid, and signal processing. She is currently visiting The University of Alabama.