symmetry

MDPI

*Article*

# A Novel Detection and Identification Mechanism for Malicious Injection Attacks in Power Systems

**Hongfeng Zhang [1], Xinyu Wang [2,3,*], Lan Ban [1] and Molin Sun [1]**

[1] School of Intelligent Manufacturing, Tianjin College, University of Science and Technology Beijing, Tianjin 301830, China; zhanghongfeng0851@163.com (H.Z.); sunmolin0851@163.com (M.S.)
[2] School of Electronic and Electrical Engineering, Yanshan University, Qinhuangdao 066004, China
[3] Jiangsu Collaborative Innovation Center for Smart Distribution Network, Nanjing 210000, China
* Correspondence: wxyzmya@ysu.edu.cn

**Abstract:** The integration of advanced sensor technology and control technology has gradually improved the operational efficiency of traditional power systems. Due to the undetectability of these attacks using traditional chi-square detection techniques, the state estimation of power systems is vulnerable to cyber–physical attacks, For this reason, this paper presents a novel detection and identification framework for detecting malicious attacks in power systems from the perspective of cyber–physical symmetry. To consider the undetectability of cyber–physical attacks, a physical dynamics detection model using the unknown input observers (UIOs) and cosine similarity theorem is proposed. Through the design of UIO parameters, the influence of attacks on state estimation can be eliminated. A cosine similarity value-based detection criterion is proposed to replace the traditional detection threshold. To further cut down the effects caused by malicious attacks, an observer combination-based attack identification framework is established. Finally, simulations are given to demonstrate that the proposed security method can detect and identify the injected malicious attacks quickly and effectively.

**Keywords:** power system; unknown input observer; cosine similarity theorem; attack identification; false data injection attack



**Citation:** Zhang, H.; Wang, X.; Ban, L.; Sun, M. A Novel Detection and Identification Mechanism for Malicious Injection Attacks in Power Systems. *Symmetry* **2023**, *15*, 2104. https://doi.org/10.3390/sym15122104

Academic Editors: Bo Zhang, Pei Liu, Yuanai Xie, Haiyan Zhao and Weifeng Zhong

Received: 19 October 2023
Revised: 20 November 2023
Accepted: 21 November 2023
Published: 23 November 2023



## 1. Introduction

As the foundation of the national economy, power systems play an irreplaceable role. In particular, the integration of advanced intelligent sensor fusion technology and robust control technology has significantly improved the operational efficiency of smart grids [1,2]. Meanwhile, two kinds of uncertainties are symmetry and superimposition, which pose new security challenges. In recent years, cyber-physical attacks have proven able to tamper with the operational status of power generation systems [3,4]. By injecting a bank of false data, the above attacks can deceive traditional detection methods. Thus, the above attacks bring a tremendous security risk to power systems [5]. For instance, unknown cyber–physical attacks occurred in Delta Montrose Electric Power Association (DMEA), USA, 2021; and a malicious ransomware attack event occurred in Taiwan, 2022. For this reason, developing an effective detection and identification mechanism is crucial to ensure the normal running of power systems.

Different from traditional network attacks, such as Denial of service attacks (Dos) and random attacks, the emerging cyber-physical attacks can fool the detection mechanism of power systems. In particular, false data injection (FDI) attacks, as one type of cyber-physical attack, were constructed firstly by Liu [6]. Furthermore, Yang et al. tested the covert characteristic of FDI attacks on different IEEE standard bus systems [7]. Based on this, a novel FDI attack that can disable a programmable logic controller without triggering an alarm was constructed [8,9]. In summary, how we may quickly detect and identify FDI

attacks is a security issue that the current power grid needs to solve from the perspective of cyber–physical symmetry.

## 2. Related Works

Responding to the security risk posed by FDI attacks, various detection methods have been studied. Existing detection methods for FDI attacks can be divided into two categories: data-driven methods and model-based methods [10]. By analyzing the characteristics of the history data in power systems, various data-driven detection methods are studied [11–18]. In [11], Raj et al., proposed a novel machine learning detection method for malicious attacks based on threshold and aggregation. By analyzing the physical behavior, an anomaly detection architecture using a neural network was constructed [12]. In [13], a threat model against malicious attacks was established based on data analysis. A novel anomaly detection strategy against malicious attacks was proposed in [14]. In [15], an innovative detection algorithm against malicious attacks based on the Markov chain was proposed. The proposed method can shed light on the security of smart grids. In [16], a detection framework including generalized cumulative sum and relaxed generalized cumulative sum was proposed. A linear regression-based three-network topology independent detection technique for FDI attacks was developed [17]. In [18], a gradient-boosting theft detector was proposed to improve detection performance against malicious attacks. By analyzing the characteristics of power system transmission data, the above data-driven detection methods can effectively judge abnormal data. Meanwhile, due to a lack of consideration for the actual physical dynamic changes in the power system, the performance of these detection methods can be limited. Furthermore, one can find that the effectiveness of the above methods depends on the assumption that the attack sequence is known. Based on this, model-based detection methods are proposed in [19–25]. In [19], a novel physical dynamic detection architecture against malicious attacks was developed. To address the problem of state estimation under cyber attacks, an adaptive cubature Kalman filter was designed [20]. Considering the undetectability of FDI attacks using traditional chi-square detection techniques, an unscented Kalman filter-based detection model was proposed [21]. In [22], Liu considered a two-step false data injection attack strategy and developed an extended Kalman filter-based detection model. Taking multichannel cyber-attacks into account, Xiahou et al. developed a decentralized detection model [23]. In [24], a hybrid dynamic-state estimation method for FDI attacks was developed based on a physical dynamics model. A novel detection model from the perspective of state estimation was constructed to identify the injected FDI attacks in smart grids [25]. In a word, the above model-based detection methods can effectively respond to injected cyber attacks. As shown in Table 1, the pros and cons of the above detection works are summarized. From the perspective of cyber–physical symmetry, the following issues need to be addressed:

- How to address the limitation caused by precomputed threshold;
- How to cut down the effect of model error and external disturbance.

**Table 1.** Analysis of detection methods against FDI attacks.

| Category | Approach | Advantages | Disadvantages |
| --- | --- | --- | --- |
| Data-driven methods | [11–18] | (1) System models are not required<br>(2) Known attack detection is fast | (1) Lots of historical data needed<br>(2) Selection of detection threshold<br>(3) Effect of model error and external disturbance |
| Model-based methods | [19–25] | (1) No training required<br>(2) Reduced memory need | |

Motivated by the above challenges, this paper develops a detection and identification architecture for use against FDI attacks. Taking the changes in the voltage of physical systems into account, a discrete power grid model is constructed. Considering the covert characteristics of FDI attacks, a novel detection model is developed based on the designed unknown input observers (UIOs) and cosine similarity theorem. Uusing the principle of cosine similarity matching, the proposed detection criterion can address the limitation caused by the precomputed detection threshold. Furthermore, an observer combination-based attack identification framework is proposed, using which the influence caused by FDI attacks can be cut down. The main works of this paper are summarized as follows.

1. A novel detection model is developed based on the UIOs and cosine similarity theorem. By designing the UIOs to handle the effect of model error and external disturbance, the accuracy of state estimation can be improved. By using the principle of cosine similarity matching, the limitation caused by the precomputed detection threshold can be addressed.
2. An observer combination-based attack identification framework is proposed. By introducing the observer combination strategy, the influence caused by the injected FDI attacks can quickly be identified and eliminated.

The outline of this paper is given as follows. In Section 3, we construct a discrete power grid model by considering the three-phase sinusoidal voltage equations. In Section 4, a detection and identification framework is proposed. Section 5 presents the effectiveness of the proposed detection and identification framework. Finally, the conclusion and discussion are given in Section 6.

## 3. Problem Description

### *3.1. Three-Phase Voltage-Based Power State Model*

According to the work in [26], one can find that attackers try to destroy the stability of power systems by tampering with state variables. As shown in Figure 1, there exist three generators and six buses in the power system. In this paper, we only consider the change in voltage signal caused by FDI attacks. Therefore, a voltage signal-based grid state model is given as follows [26].

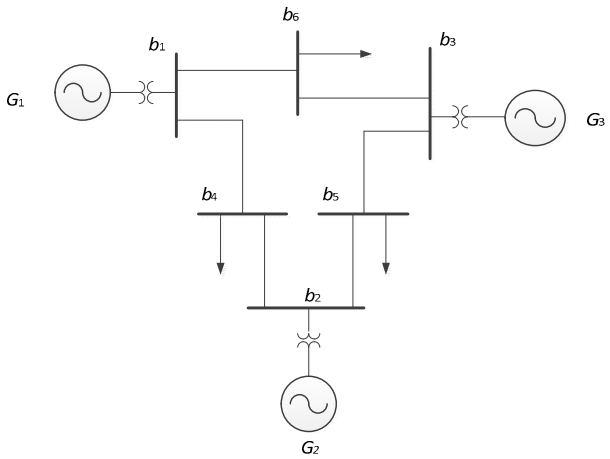


**Figure 1.** IEEE 6-bus power system.

$$V_1(k) = A_\iota \cos(wk + \theta) \tag{1}$$

$$V_2(t) = A_\iota \cos\left(wt + \theta - \frac{2}{3}\pi\right) \tag{2}$$

$$V_3(k) = A_\iota \cos\left(wk + \theta - \frac{4}{3}\pi\right) \tag{3}$$

By using Equations (1)–(3), one can obtain

$$V(t) = A_\iota * \cos(wt)\cos(\theta) - A_\iota * \sin(wt)\sin(\theta) \tag{4}$$

Since this paper only studies the change in three-phase voltage, it is assumed that $\omega t$ is relatively constant over time. Then, the state–space grid model in Equation (4) can be rewritten as

$$x_{k+1} = Ax_k + Bu_k + \Delta_k \tag{5}$$

$$y_k = Cx_k \tag{6}$$

where $x_k = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$, $x_1 = A_\iota \times \cos\theta$, and $x_2 = A_\iota \times \sin\theta$ are state variables at epoch $k$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $C = \begin{pmatrix} \cos\omega k & -\sin\omega k \end{pmatrix}$.

### *3.2. Problem Formulation*

As usual, the commonly used detection methods are based on chi-square detectors. By comparing the measurement residual and measurement output, the operator can determine if there are any abnormalities in the power system. As shown in [27], the detection criterion for abnormal data is given as follows.

$$\begin{cases} \lambda_k = \|\mathbf{z}_k - C\hat{x}_k\| < \tau & \text{No attack} \\ \lambda_k = \|\mathbf{z}_k - C\hat{x}_k\| \geq \tau & \text{Attack} \end{cases} \tag{7}$$

According to the work in [28], the $\tau$ is set to three times the size of the noise, which can then meet a false alarm rate of less than 5%. According to the work in [16], the detection process of chi-square detector is given as follows. Firstly, the residual $\lambda_k = \|\mathbf{z}_k - C\hat{x}_k\|$ can be obtained using the Kalman filter. Secondly, the computation of precomputed threshold $\tau$ is based on noise. Thirdly, operators can judge the injected attacks by comparing the residual and precomputed threshold. The above computation process can be seen in [27].

Making full use of the above detection mechanism, hackers can design a special set of attack sequence $f_k^a$, which needs to satisfy the following constraint.

$$\begin{aligned} \hbar_k^f &= \left\| z_k^f - C\hat{x}_k^a \right\| \\ &= \left\| (z_k + y_k^f) - (C\hat{\boldsymbol{x}}_k + \ell) \right\| \\ &= \left\| (z_k - C\hat{x}_k) + (y_k^f - \ell) \right\| \\ &\leq \|z_k - C\hat{x}_k\| + \left\| y_k^f - \ell \right\| \end{aligned} \tag{8}$$

From Equation (8), one can find that the output residual $\hbar_k^f$ has not changed if $y_k^f = \ell$. In other words, the output residual $\hbar_k^f$ under FDI attacks still is smaller than the precomputed detection threshold $\tau$. It is obvious that the injected FDI attacks can fool the above detection methods by using the chi-square detector. Based on this, the malicious attackers can successfully tamper with the operating status of the power system without triggering system alerts.

To sum up, the emergence of attacks poses a huge challenge to existing security mechanisms in power systems. To address this problem, this paper develops a novel detection and identification method based on actual physical state changes. Taking a large-scale power grid, the lth state–space grid model can be described as follows:

$$\begin{cases} x_{k+1}^l = A^l x_k^l + B^l u_k^l + \Delta^l{}_k + F^l f_k^l \\ y_k^l = C^l x_k^l \end{cases} \tag{9}$$

where $f_k^l$ denotes the attack vector in Equation (8). The above parameter definitions are given in Nomenclature. Based on the established problem, a novel detection and identification framework including three steps is proposed from the perspective of cyber–physical symmetry.

## 4. Detection and Identification Mechanism for FDI Attacks

In this section, a detection and identification framework for FDI attacks is established. Taking the model error and external disturbance into account, we design UIOs to obtain the physical dynamics accurately. Then, a novel detection criterion based on a cosine matching theorem is proposed to address the limitations of the precomputed detection threshold. Finally, an observer combination-based identification method against multiple FDI attacks is developed. The detailed steps are given as follows from the perspective of cyber–physical symmetry.

Step 1: To obtain the accurate state estimation, a bank of UIOs is established to deal with the model error and external disturbance.

Step 2: A cosine theorem-based detection criterion is proposed to replace the precomputed detection threshold.

Step 3: An observer combination-based method is developed to identify the injected FDI attacks.

### *4.1. UIO-Based State Estimation*

Without the injected FDI attacks, the lth state–space grid model can be rewritten as

$$\begin{cases} x_{k+1}^l = A^l x_{kk}^l + B^l u_k^l + \Delta^l{}_k \\ y_k^l = C^l x_k^l \end{cases} \tag{10}$$

To ensure the accuracy of state estimation, this paper designs the UIO to handle model error and external disturbance. Before designing the proposed UIO, some assumptions and lemmas are given as follows.

**Assumption 1** ([29]). *$\left(C^l, \widetilde{A}^l\right)$ is a detectable, where $\widetilde{A}^l = \left(I - K^l C^l\right)$.*

**Lemma 1** ([30]). *Consider a discrete system as*

$$\begin{cases} \xi_{\kappa+1} = \mathrm{P}\xi_\kappa + \mathrm{Z}v_\kappa \\ \psi_\kappa = \mathrm{X}\xi_\kappa + \Theta v_\kappa \end{cases} \tag{11}$$

*which meets the following robust performance:*

$$\left\| G_{yu}(s) \right\|_- > \eta \tag{12}$$

*where $G_{yu}(s) = \mathrm{X}(\sigma\mathrm{I} - \mathrm{P})^{-1}\mathrm{Z} + \Theta$, if there exist positive definite matrices $\aleph$ and $\Im$ meeting the following constraint*

$$\begin{bmatrix} \mathrm{P} & \mathrm{Z} \\ \mathrm{I} & 0 \end{bmatrix}^T \partial \begin{bmatrix} \mathrm{P} & \mathrm{Z} \\ \mathrm{I} & 0 \end{bmatrix} + \begin{bmatrix} \mathrm{X} & \Theta \\ \mathrm{I} & 0 \end{bmatrix}^T \wp \begin{bmatrix} \mathrm{X} & \Theta \\ \mathrm{I} & 0 \end{bmatrix} < 0 \tag{13}$$

*where*

$$\partial = \begin{bmatrix} -I & 0 \\ 0 & \lambda^2 I \end{bmatrix}, \wp = \begin{bmatrix} -\aleph & \Im \\ \Im & \aleph - 2\cos(\omega_1)\Im \end{bmatrix}$$

*for the low-frequency range $|\omega| \le \omega_1$.*

**Lemma 2** ([31]). *If there exists a matrix* $\Omega$ *meeting the following constraint:*

$$\begin{bmatrix} -\Omega & \Omega(\sum -\omega I) \\ \Omega(\sum -\omega I) & -\ell^2\Omega \end{bmatrix} < 0, \tag{14}$$

*then the eigenvalues of matrix* $\sum$ *belong to the circular region* $\mho(\omega, \ell)$.

Based on this, the proposed UIO can be established as follows.

$$\begin{cases} z_{k+1}^l = H^l z_k^l + G^l B^l u_k^l + U^l y_k^l \\ \hat{x}_k^l = z_k^l + K^l y_k^l \end{cases} \tag{15}$$

Defining estimation error as

$$e_k^l = x_k^l - \hat{x}_k^l, \tag{16}$$

From Equations (9) and (15), one can obtain

$$\begin{aligned} e_{k+1}^l &= x_{k+1}^l - \hat{x}_{k+1}^l \\ &= \left(I^l - K^l C^l\right) x_{k+1}^l - z_{k+1}^l \\ &= \left(A^l - K^l C^l A^l - U_1^l C^l\right) e_k^l + \left(A^l - K^l C^l A^l - U_1^l C^l - H_1^l\right) z_k^l \\ &+ \left[\left(A^l - K^l C^l A^l - K_1^L C^l\right) K^l - U_2^l\right] y_k^l + \left[\left(I^l - K^l C^l\right) B^l - G^l B^l\right] u_k^l \\ &+ \left(I^l - K^l C^l\right) \Delta^l + \left(I^l - K^l C^l\right) F^l f^l \end{aligned} \tag{17}$$

Based on the design of UIO, one can obtain

$$\begin{cases} \left(I^l - K^l C^l\right) B^l - G^l B^l = 0 \\ A^l - K^l C^l A^l - U_1^l C^l - H_1^l = 0 \\ U_2^l = H^l K^l \end{cases} \tag{18}$$

Then, Equation (17) can be rewritten as follows:

$$e_{k+1}^l = H^l e_k^l + G^l \Delta^l + G^l F^l f_k \tag{19}$$

Without FDI attacks, the state residual can be obtained:

$$r_k^l = r_k^l - C^l x_k^l = C^l e_k^l = H^l e_k^l + G^l \Delta^l \tag{20}$$

By designing the UIOs, the effect of model error and external disturbance can be reduced. The corresponding UIOs-based state estimation framework is presented in Figure 2. Then, the following theorem is given to ensure the stability of UIOs.

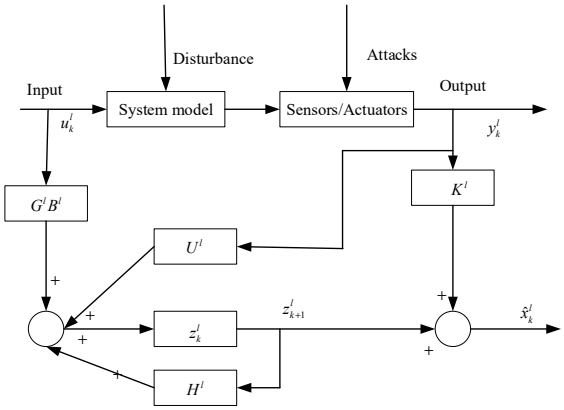


**Figure 2.** UIO-based state estimation framework.

**Theorem 1.** *Under* ***Lemma 1*** *and* ***Lemma 2****, there exist matrices* $P_1^l = \left(P_1^l\right)^T > 0$, $P_2^l = \left(P_2^l\right)^T > 0$, $P_3^l$ *and* $P_4^l$ *meeting the following constraint.*

$$\begin{bmatrix} -\kappa_1 P_5^l - \kappa_1\left(P_3^l\right)^T & \rho_1 & \rho_2 \\ * & \rho_3 & \rho_4 \\ * & * & -(I-\vartheta)I \end{bmatrix} < 0 \tag{21}$$

$$\begin{bmatrix} -(1-\sigma)P_1^l & 0 & \left(C^l\right)^T \\ * & -(\mu-\vartheta)I & 0 \\ * & * & -\mu I \end{bmatrix} < 0 \tag{22}$$

$$\begin{bmatrix} P_2^l - \kappa_3 P_3^l - \kappa_3\left(P_3^l\right)^T & \rho_5 \\ * & -\ell_2^l \end{bmatrix} < 0 \tag{23}$$

where

$$\begin{gathered} \rho_1 = \left(P_3^l\right)^T + P_1^l + \kappa_1 P_3^l \widetilde{A}^l - \kappa_1 P_4^l C^l \\ \rho_2 = \kappa_1 P_3^l\left(I^l - K^l C^l\right)F^l \\ \rho_3 = -\sigma P_1^l + P_3^l \widetilde{A}^l - P_4^l C^l + \left(\widetilde{A}^l\right)^T\left(P_3^l\right)^T - \left(C^l\right)^T\left(P_4^l\right)^T \\ \rho_4 = P_3^l\left(I^l - K^l C^l\right) \\ \rho_5 = -\kappa_3 P_3^l \widetilde{A}^l + \kappa_3 P_4^l C^l + \kappa_3 w P_3^l \end{gathered}$$

and $\kappa_3$, $\kappa_3$, $\kappa_3$ and $w_1 > 0$.

**Proof.** Without FDI attacks, we can obtain

$$\begin{cases} e_{k+1}^l = H^l e_k^l + G^l \Delta^l \\ r_k^l = C^l e_k^l \end{cases} \tag{24}$$

The Lyapunov function is selected as

$$V_k^l = \left(e_k^l\right)^T \Psi^l e_k^l \tag{25}$$

Based on the work in [32], one can obtain the $L_\infty$ performance as

$$index\left\|r_k^l\right\| < \sqrt{\mu(1-\sigma)\sigma V_0^l + \mu(1-\sigma+\mu-\vartheta)\left\|\Delta^l\right\|_\infty^2} \tag{26}$$

$$V_{k+1} < \sigma V_k + (1-\sigma)\left(\Delta^l\right)^T \Delta^l \tag{27}$$

$$\left(r_k^l\right)^T r_k^l < \mu(1-\sigma)V_k + \mu(\mu-\vartheta)\left(\Delta^l\right)^T \Delta^l \tag{28}$$

Based on Equation (25), one can obtain

$$V_{k+1}^l = \left(e_k^l\right)^T\left(P_1^l\left(\widetilde{A}^l - U_1^l C^l\right) + \left(\widetilde{A}^l - U_1^l C^l\right)^T P_1^l\right)e_k^l + 2\left(e_k^l\right)^T P_1^l\left(I^l - K^l C^l\right)\Delta^l \tag{29}$$

From Equations (25)–(29), one can obtain

$$\begin{bmatrix} e_k \\ \Delta^l \end{bmatrix}^T \Gamma \begin{bmatrix} e_k \\ \Delta^l \end{bmatrix} < 0 \tag{30}$$

where

$$\Gamma = \begin{bmatrix} P_1^l\left(\widetilde{A}^l - U_1^l C^l\right) + \left(\widetilde{A}^l - U_1^l C^l\right)^T P_1^l - \sigma P_1^l & P_1^l\left(I^l - K^l C^l\right) \\ P_1^l\left(I^l - K^l C^l\right) & -(1-\sigma)I \end{bmatrix} \tag{31}$$

Since matrix $\Gamma$ is not a LMI, Equation (31) can be rewritten as

$$\chi_1 + \phi_1^T \varphi_2^T + \phi_1 \varphi_2 < 0 \tag{32}$$

where

$$\phi_1 = \begin{bmatrix} \widetilde{A}^l - U_1^l C^l & I^l - K^l C^l \end{bmatrix}$$

$$\varphi_2 = \begin{bmatrix} \left(P_1^l\right)^T & 0 \end{bmatrix}^T$$

$$\chi = \begin{bmatrix} \sigma P_1^l & 0 \\ 0 & -(1-\sigma)I \end{bmatrix}$$

Then, Equation (32) can be equivalent to

$$\begin{bmatrix} -\kappa_1 P_3^l - \left(\kappa_1 P_3^l\right)^T & -\phi_2^T + \varphi_2^T + \kappa_1 P_3^l \phi_1 \\ * & \chi_1 + \phi_2 \phi_1 + \phi_1^T \phi_2^T \end{bmatrix} < 0 \tag{33}$$

Multiplying the left and right sides of Equation (30) by $\begin{bmatrix} \phi_1 & I \end{bmatrix}$, and its transposition, we can obtain Equation (21).

Based on Equation (27), one can obtain

$$\begin{aligned} V_k &< \sigma^k V_k(0) + (1-\sigma)\sum_{i=0}^{k-1} \sigma^i \left\|\Delta^l\right\|^2 \\ &< \sigma^k V_k(0) + \left\|\Delta^l\right\|^2 \end{aligned} \tag{34}$$

Taking Equation (34) into Equation (28), one can obtain

$$\begin{bmatrix} e_k \\ \Delta^l \end{bmatrix}^T \begin{bmatrix} \varepsilon^{-1}\left(C^l\right)^T C^l - (1-\sigma)P_1^l & 0 \\ 0 & -(1-\sigma)I \end{bmatrix} \begin{bmatrix} e_k \\ \Delta^l \end{bmatrix} < 0 \tag{35}$$

By using the Schur complement lemma [33], Equation (22) can be obtained based on Equation (35).

Based on **Lemma 2**, one can obtain

$$\begin{bmatrix} -P_2^l & P_2^l\left(\widetilde{A}^l - U_1^l C^l - \omega I\right) \\ * & -\ell^2 P_2^l \end{bmatrix} < 0 \tag{36}$$

By using the Schur complement lemma [28], one can obtain

$$\left(\widetilde{A}^l - U_1^l C^l - \omega I\right)^T P_2^l \left(\widetilde{A}^l - U_1^l C^l - \omega I\right) - \ell^2 P_2^l < 0 \tag{37}$$

Equation (37) can be equivalent to

$$\begin{bmatrix} P_2^l - \kappa_3 P_3^l - \left(\kappa_1 P_3^l\right)^T & -\kappa_3 P_3^l\left(\widetilde{A}^l - U_1^l C^l - \omega I\right) \\ * & \ell^2 P_2^l \end{bmatrix} < 0 \tag{38}$$

Multiplying the left and right sides of in Equation (38) by $\begin{bmatrix} \left(\widetilde{A}^l - U_1^l C^l - \omega I\right) & I \end{bmatrix}$ and its transposition, we can obtain Equation (23). □

**Remark 1.** *In general, the shortcomings of model-based detection methods, such as model error and external disturbance, will affect the accuracy of state estimation. For this reason, this paper proposes to design the UIOs to attenuate the effects caused by model error and external disturbance.*

### *4.2. Cosine Similarity Theorem-Based Detection Method*

In this section, a cosine similarity theorem-based detection criterion is developed. Cosine similarity is a method of calculating correlation, which maps individual indicator data to a vector space and calculates the cosine value of the angle between two vectors as a measure of similarity between two variables. In other words, the value of cosine similarity tends to be 0 if two values are similar. If two values are not similar, the value of cosine similarity tends to be 1.

Using the above cosine similarity theorem, a cosine similarity value (CSV)-based attack detection criterion is proposed as follows.

$$\begin{cases} \lim\limits_{k\to\infty} \cos(|x_k - \hat{x}_k|) = \Xi_k = 0 & Normal \\ \lim\limits_{k\to\infty} \cos(|x_k - \hat{x}_k|) = \Xi_k = 1 & Abnormal \end{cases} \tag{39}$$

Based on the designed UIOs, a cosine similarity theorem-based attack detection process is developed as follows.

Step 1: Establish the proposed voltage signal-based grid state model in Equations (2) and (3).

Step 2: Design a bank of the proposed UIOs in Equation (15) to obtain the estimation state.

Step 3: Compute the cosine similarity value based on the cosine similarity theorem.

Step 4: Apply the detection criterion in Equation (39) to detect the injected FDI attacks.

Based on the proposed detection process, the corresponding pseudo-code of the attack detection method is given in Algorithm 1.

**Algorithm 1**: Cosine similarity theorem-based detection algorithm against FDI attacks

1. $GCM \rightarrow$ Grid state model;
2. $UIOs \rightarrow$ Unknown input observers;
3. $ES \rightarrow Estimation$ state;
4. $CSV \rightarrow$ *cosine similarity value*
5. $UIOs \leftarrow GCM$;
6. $ES \leftarrow UIOs$;
7. $CSV \leftarrow ES$
8. $IF\ 0 \leftarrow CSV$, No attacks
9. $ELSEIF$ attacks

### *4.3. Observer Combination-Based Identification Method*

To ensure the stable operation of power systems, an observer combination-based attack identification framework is proposed to minimize the influence of FDI attacks, as shown in Figure 3.

By applying the method of bisection, the proposed identification method can detect and identify multiple attacked nodes. The detailed identification processes are given as follows.

Step 1: Assuming all $y_1 \cdots y_N$ outputs are driven by one bank of UIOs, the CSV detector is designed to detect attack nodes.

Step 2: If there exist attack nodes, all $y_1 \cdots y_N$ outputs are divided into two parts, which are driven by two banks of UIOs.

Step 3: If there exist attack nodes in $y_1 \cdots y_{N/2}$, all $y_1 \cdots y_{N/2}$ outputs are divided into two parts, which are driven by two banks of UIOs.

Step 4: If there exist attack nodes in $y_{N/2} \cdots y_N$, all $y_{N/2} \cdots y_N$ outputs are divided into two parts, which are driven by two banks of UIOs.

Step 5: This process is repeated iteratively for each half of the $y_1 \cdots y_{N/2}$ or $y_{N/2} \cdots y_N$, if there exist attack nodes.

Step 6: Repeating the above Steps 2–4, the attacked nodes can be detected and identified.

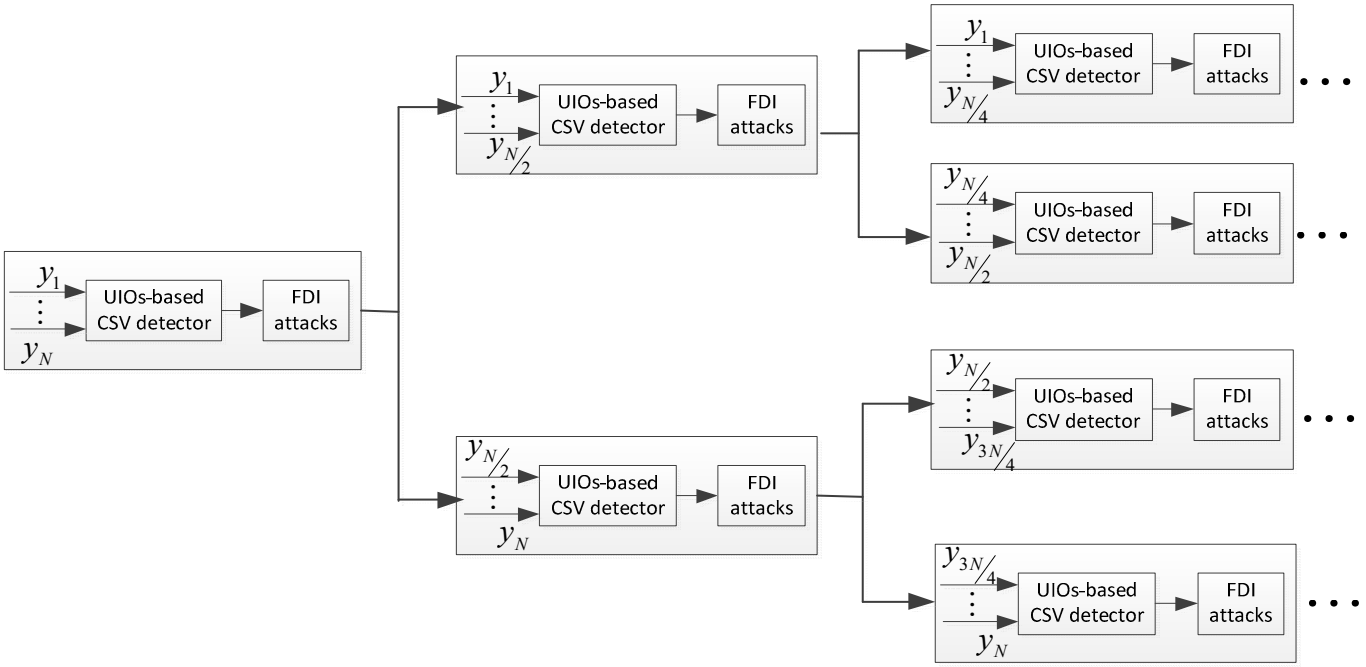


**Figure 3.** Observer combination-based identification framework.

Based on the above the detection and identification framework, operators can quickly respond to the attacked nodes. Then, the corresponding pseudo-code of attack identification method is given in Algorithm 2.

**Algorithm 2**: Observer combination-based attack identification

1. $y_1 \cdots y_N \rightarrow All$ outputs
2. $RIFEH \rightarrow repeated\ iteratively\ for\ each\ half\ of\ the\ attacked\ ouputs$
3. $UIOs \rightarrow$ Unknown input observers;
4. $CSV \rightarrow cosine\ similarity\ value$
5. $UIOs \leftarrow y_1 \cdots y_N$
6. $IF\ 1 \leftarrow CSV$, attacks;
7. $RIFEH \leftarrow y_1 \cdots y_{N/2}$ and $RIFEH \leftarrow y_{N/2} \cdots y_N$
8. Else no attacks;
9. Repeat steps 5–7;
10. Output: the attacked nodes.

## 5. Results

In this section, experiments are carried out to verify the effectiveness of the proposed detection and identification framework on IEEE 6-bus, 39-bus (as shown in Figure 4), and 118-bus power systems. Partial experimental parameters are set as follows: a frequency of 60 Hz and an amplitude of 0.5 V, $\left\|\Delta^l{}_k\right\| \leq 0.05$. In addition, the model matrix parameters and attack sequence can be seen in [26]. In the following, case 1 is used to test the detection of one FDI attack on an IEEE 6-bus power system. Case 2 is used to verify the effectiveness of multiple FDI attacks on an IEEE 39-bus power system. Case 3 is used to verify the effectiveness of multiple FDI attacks on an IEEE 118-bus power system. The simulation environment is simulated using MATLAB 2020b software on a Lenovo Y7000 computer (i7-10875H CPU,16GB, RTX2060).

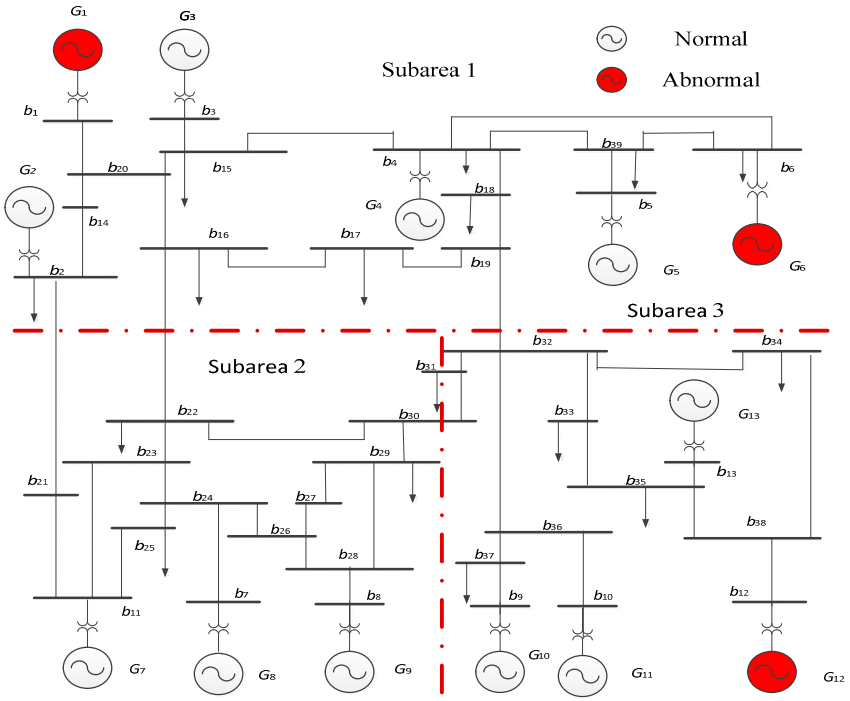


**Figure 4.** IEEE 39-bus power system.

### *5.1. Case 1: Detection of One FDI Attack on an IEEE 6-Bus Power System*

It is assumed that the first generator is an injected FDI attack by hacker at $t$ = 160 s. By tampering with sensor data, the hacker can fool the detection technique using a Chi-square detector. Based on this, the proposed detection algorithm is applied to obtain the estimated state under FDI attacks, as shown in Figure 5.

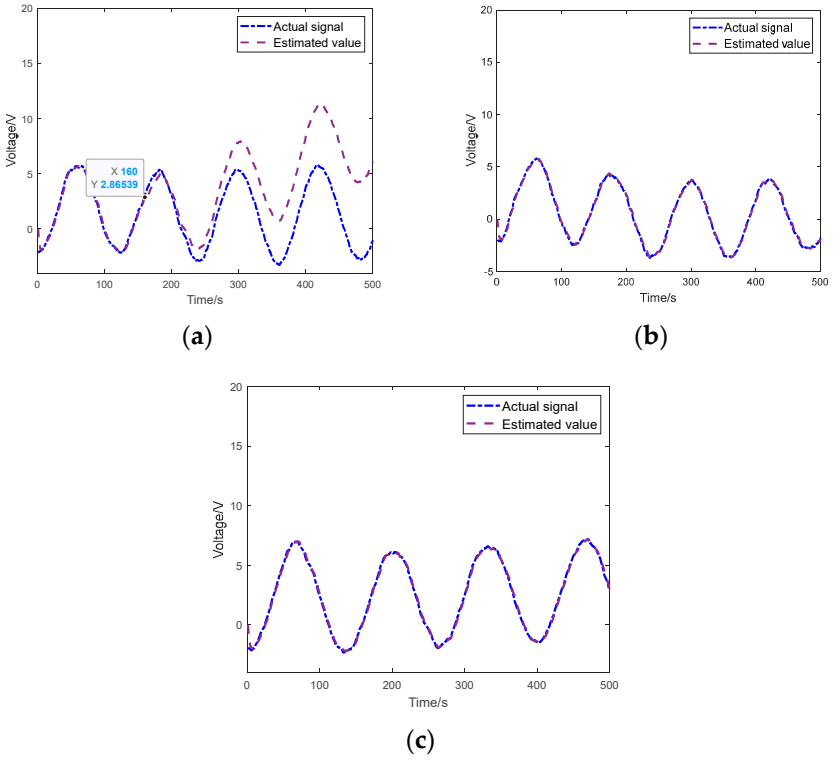


**Figure 5.** The change in voltage under FDI attacks. (**a**) The first generator; (**b**) The second generator; (**c**) The third generator.

As shown in Figure 5, it is obvious that only the state estimation of the first generator has been changed. Compared to first generator, there does not exist the change for other generators. In other words, the hacker has successfully injected false data and changed the voltage state of the first generator. To detect the injected FDI attacks, a traditional chi-square detection technique and the proposed CSV-based attack detection criterion are applied. As shown in Figures 6 and 7, the corresponding detection results for the above two methods are given.

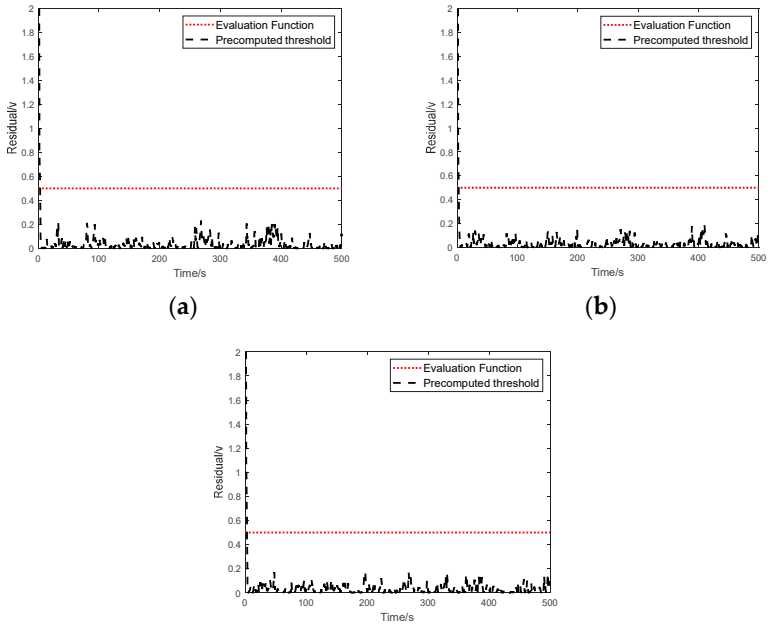


**Figure 6.** The detection results under the chi-square detection technique. (**a**) First generator; (**b**) second generator; (**c**) third generator.

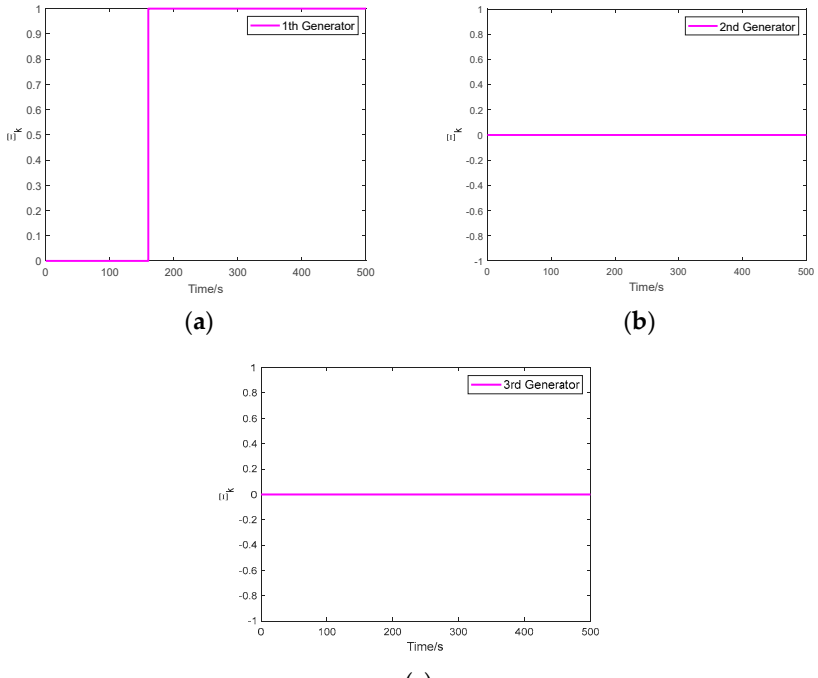


**Figure 7.** The detection results under the CSV-based attack detection technique. (**a**) First generator; (**b**) second generator; (**c**) third generator.

Obviously, one can find that all the detection residuals do not exceed the precomputed threshold, as shown in Figure 6. Namely, the designed FDI attack can deceive the existing chi-square detection technique methods without triggering alerts. Meanwhile, the proposed CSV-based attack detection criterion can judge that the CSV of first generator is 1 at $t$ = 160 s, as shown in Figure 7a. In other words, the proposed detection algorithm can successfully detect the change in voltage on the first generator caused by FDI attack. In summary, the developed CSV-based detection Algorithm 1 can effectively detect the injected FDI attacks in the power system.

### *5.2. Case 2: Detection and Identification of Multiple FDI Attacks on an IEEE 39-Bus Power System*

In this section, we consider the detection and identification of multiple FDI attacks on an IEEE 39-bus power system. As shown in Figure 4, the IEEE 39-bus power system consists of three grid subareas. Of note, the method for dividing power grid areas can be seen in [34]. In addition, the calculation of extended model parameters can be seen in [28]. It is assumed that hacker can inject multiple FDI attacks, such as the first generator bus at $t \in (100\text{ s} - 160\text{ s})$, sixth generator bus at $t = 260$ s, and thirteenth generator bus at $t = 380$ s. By designing a bank of UIOs for three subareas, the corresponding change in the estimated state can be obtained, as shown in Figure 8. Obviously, one can find that there exist changes on the first and third grid subareas. To identify the attacked generator buses, the proposed Algorithm 2 is applied. Then, the corresponding detection and identification results against FDI attacks are obtained, as shown in Figures 9 and 10.

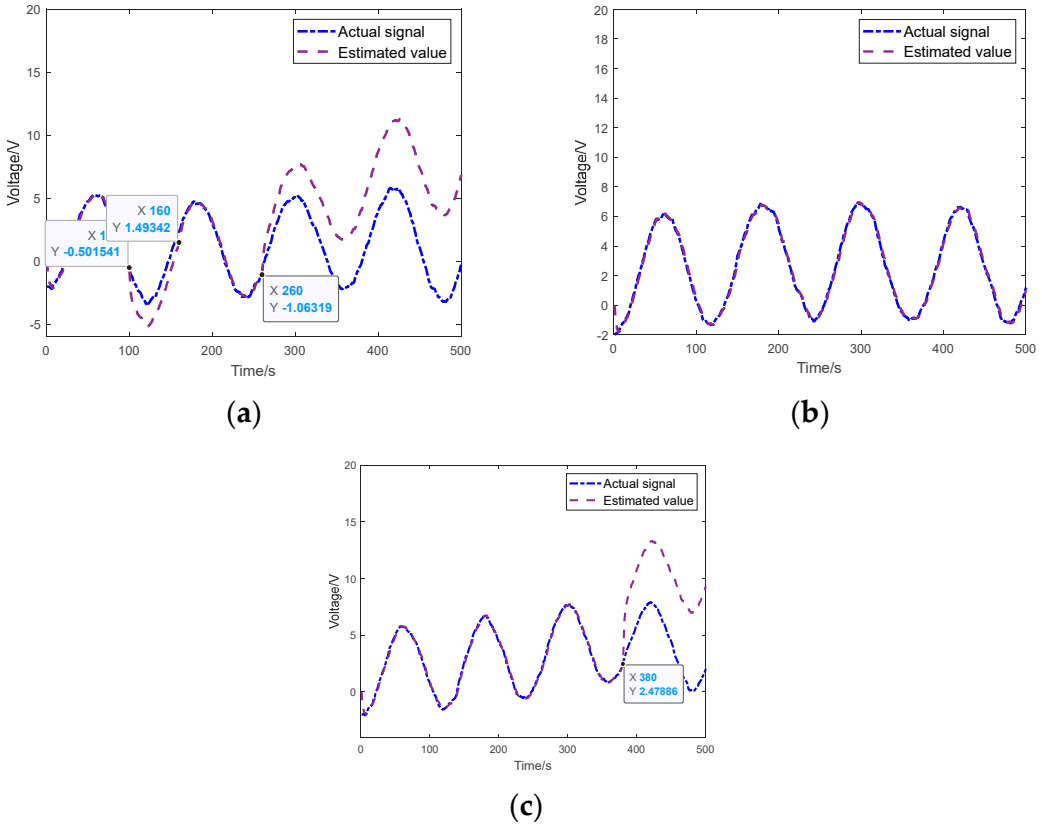


**Figure 8.** The detection results of three subareas under the CSV-based attack detection technique. (**a**) First subarea; (**b**) second subarea; (**c**) third subarea.

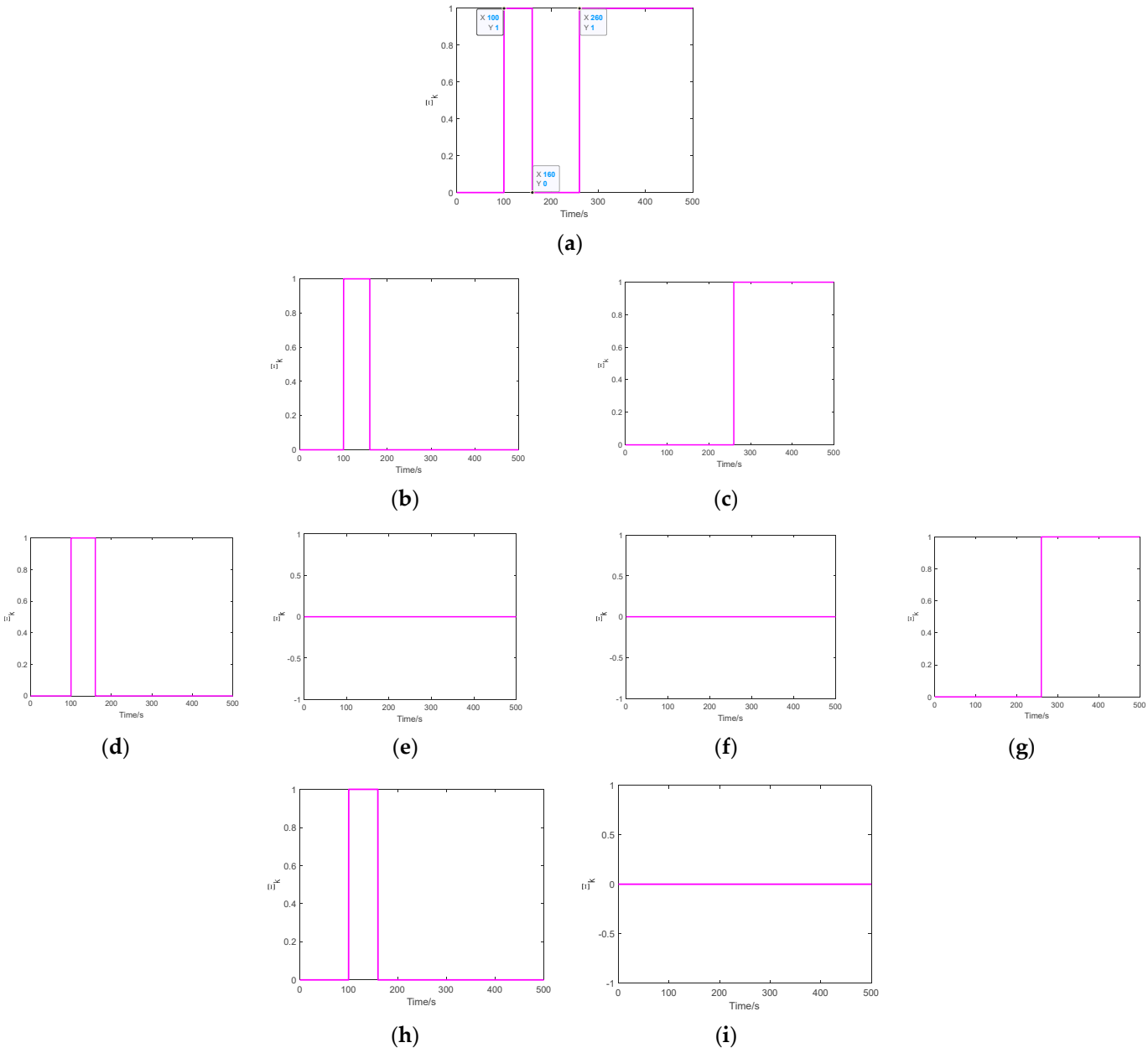


**Figure 9.** Detection and identification results against multiple FDI attacks in the first subarea. (**a**) 1–6 generator buses; (**b**) 1–3 generator buses; (**c**) 4–6 generator buses; (**d**) 1–2 generator buses; (**e**) 3 generator bus; (**f**) 4–5 generator buses; (**g**) 6 generator buses; (**h**) 1 generator bus; (**i**) 2 generator bus.

As shown in Figure 9a, the corresponding detection and identification results indicate that there may exist one or multiple FDI attacks in the first grid subarea. By applying the proposed Algorithm 2, one can find that there may exist one or multiple FDI attacks on the first to third generator buses, as shown in Figure 9b. Similarly, there may exist one or multiple FDI attacks on the fourth to sixth generator buses, as shown in Figure 9c. By using the detection and identification algorithm, we can obtain that there exists one FDI attack in the first and sixth generator bus, as shown in Figure 9d–i. Using a similar process as above, one can find that there may exist attack in the 10th–13th generator buses, as shown in Figure 10a. Through further detection and identification, we can find that the injected FDI attack is on the 12th or 13th generator bus, as shown in Figure 10c. Namely, there does not exist injected FDI attack on the 10th or 11st generator bus, as shown in Figure 10b.

Based on the proposed Algorithm 2, we can detect and identify that the injected FDI attack is on the 13th generator bus, as shown in Figure 10d,e.

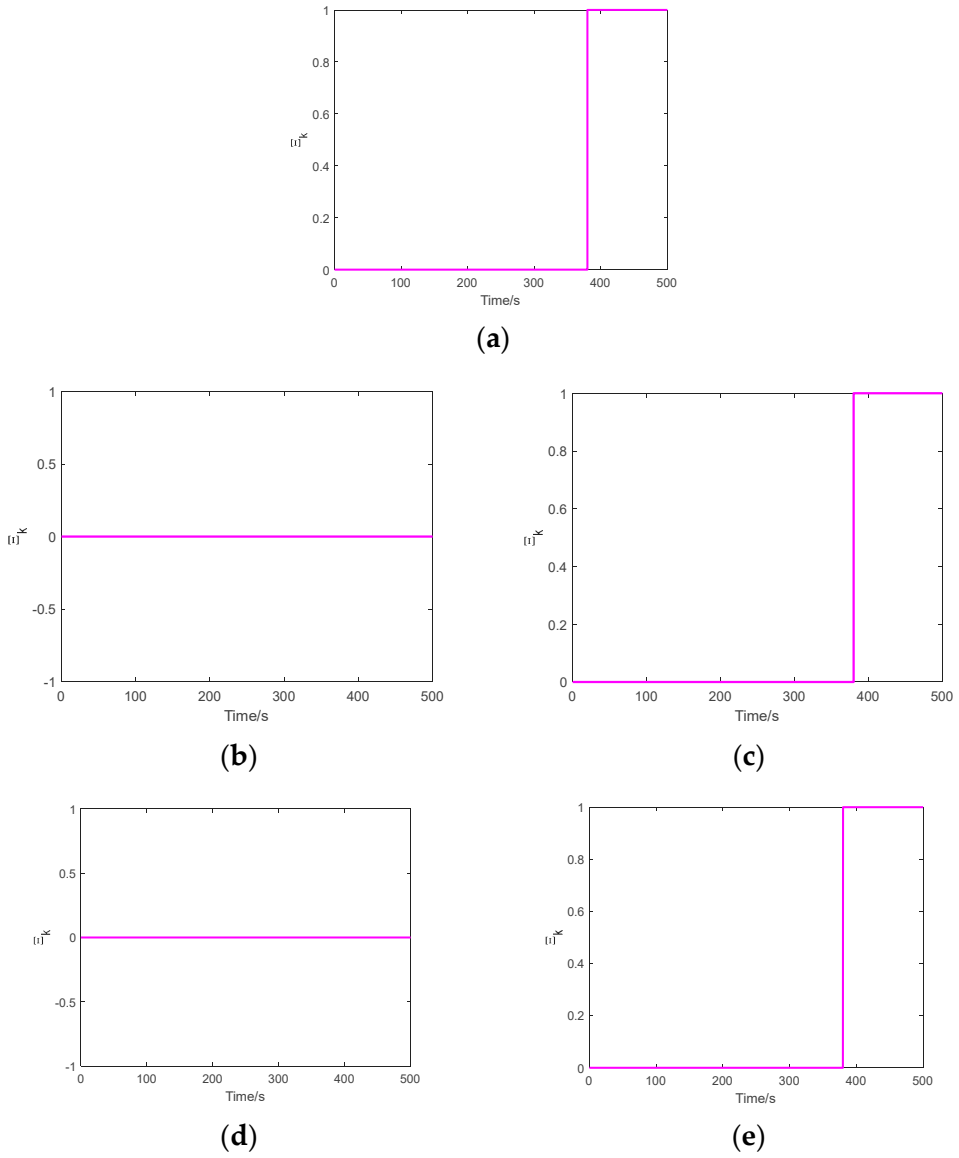


**Figure 10.** The detection and identification results against multiple FDI attacks in the third subarea. (**a**) 10–13 generator buses; (**b**) 10–11 generator buses; (**c**) 12–13 generator buses; (**d**) 12 generator bus; (**e**) 13 generator bus.

### *5.3. Case 3: Detection and Identification of Multiple FDI Attacks on IEEE 118-Bus Power System*

In this case, the effectiveness of the proposed method on a large-scale 118-bus system (Figure 11) is demonstrated. It is assumed that hacker can inject multiple FDI attacks, such as the 8th generator bus at $t = 120$ s and 14th generator bus at $t = 180$ s. Applying the proposed UIO-based state estimation method, one can find that there exist one or multiple FDI attacks in the second grid subarea, as shown in Figure 12. By using the proposed Algorithm 2, the corresponding detection and identification results against FDI attacks are obtained, as shown in Figure 13.

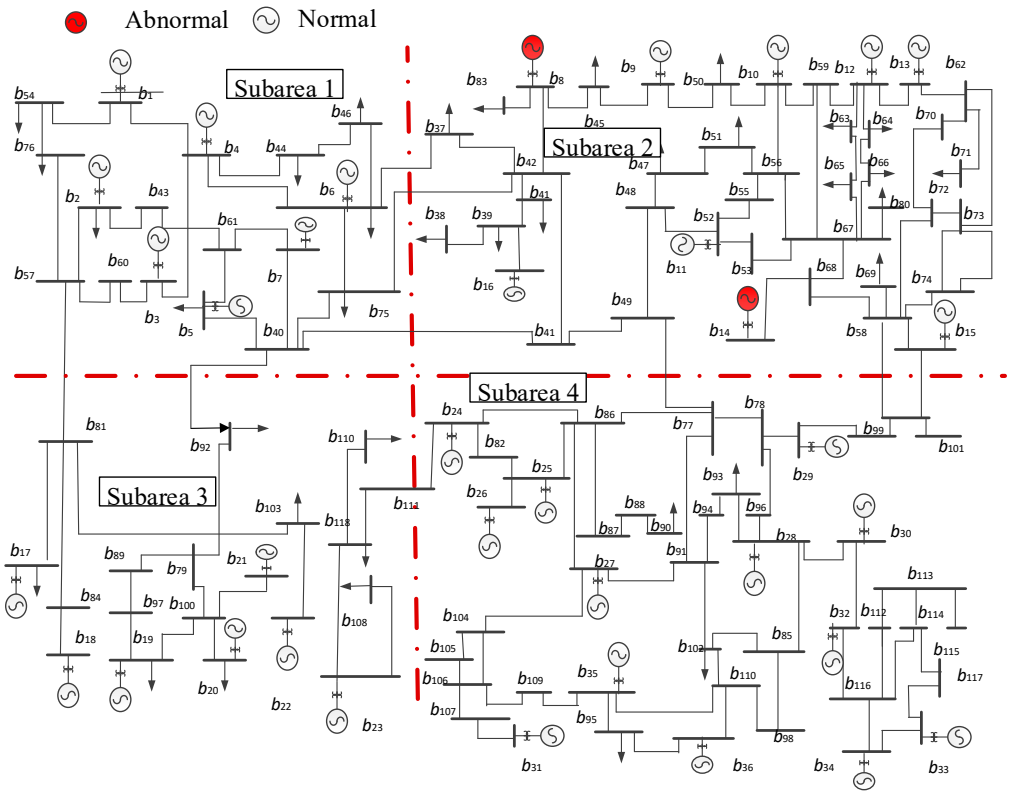


**Figure 11.** IEEE 118-bus power system.

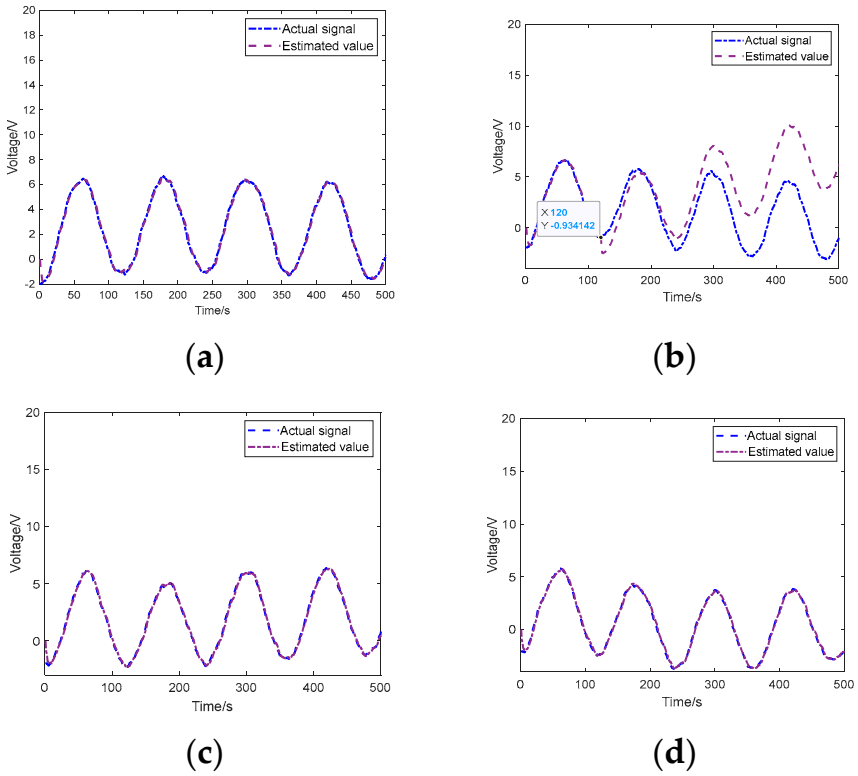


**Figure 12.** The detection results of four subareas under the CSV-based attack detection technique. (**a**) First subarea; (**b**) second subarea; (**c**) third subarea; (**d**) fourth subarea.

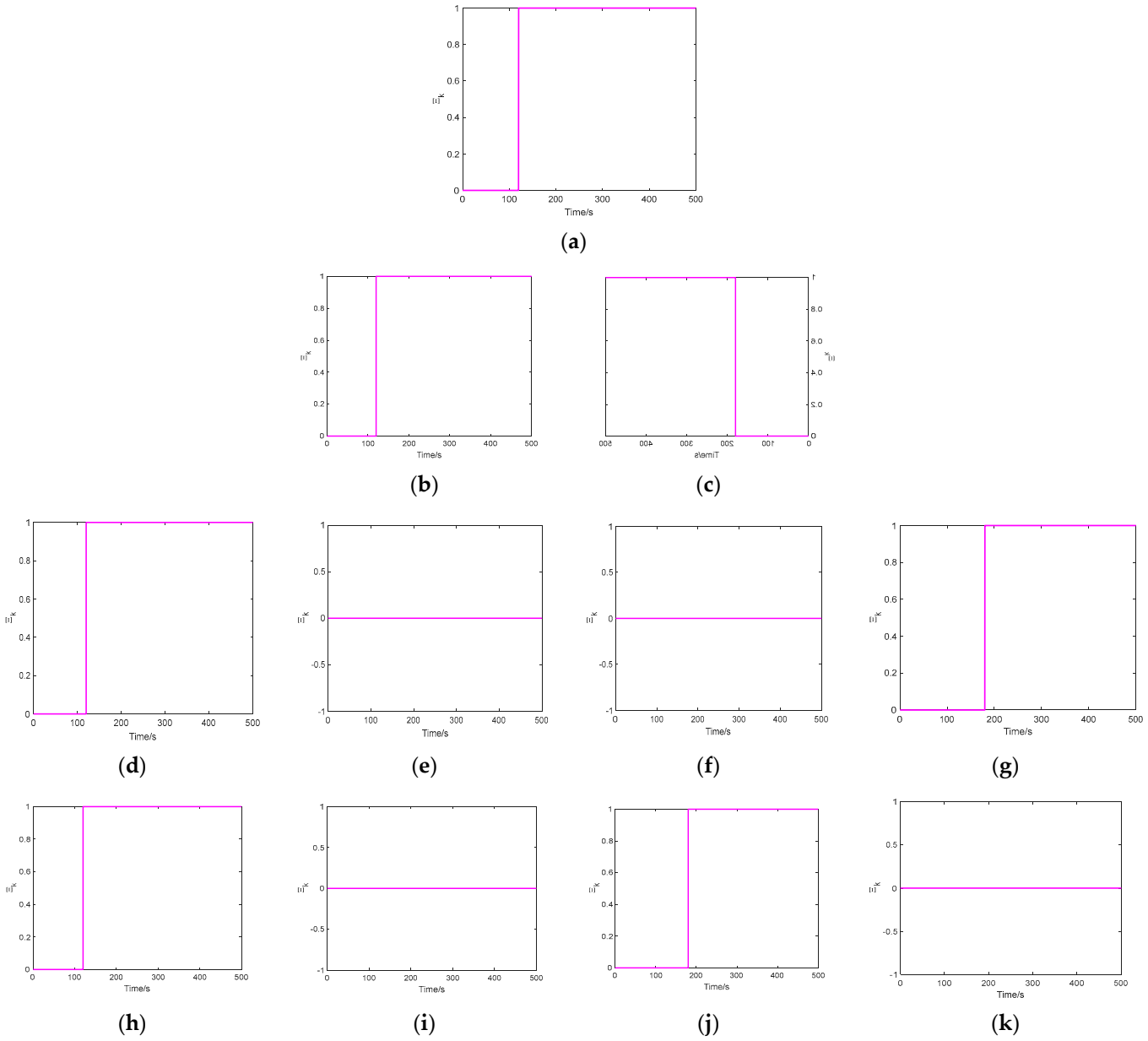


**Figure 13.** The detection and identification results against multiple FDI attacks in the second subarea. (**a**) 8–16 generator buses; (**b**) 8–11 generator buses; (**c**) 12–16 generator buses; (**d**) 8–9 generator buses; (**e**) 10–11 generator buses; (**f**) 12–13 generator buses; (**g**) 14–16 generator buses; (**h**) 8 generator bus; (**i**) 9 generator bus; (**j**) 14 generator bus; (**k**) 15–16 generator buses.

The simulation results indicate that there exist one or multiple FDI attacks on the IEEE 118-bus grid system. Obviously, it is difficult to identify the injected FDI attacks. By applying the proposed method, we can obtain that there may exist FDI attacks in the 8–16th generator buses, as shown in Figure 13a–c. Repeat the Algorithm 2, we can identify the injected FDI attacks on the 8th–9th and 14th–16th generator buses, as shown in Figure 13d–g. Using a similar process as above, the injected FDI attacks on the 8th and 14th generator buses can be identified, as shown in Figure 13h–k. In summary, the simulation results demonstrate that the injected malicious attacks can be detected and identified quickly using the proposed detection and identification algorithm. In addition, the detection and identification time are given, as shown in Table 2. Compared with the data-driven methods, the proposed method can cut down the time of training data. As the power network grows, the corresponding system parameter dimensions also increase.

Thus, the detection time of FDI attacks will increase, as shown in Table 2. Meanwhile, the identification time will increase in comparison with the detection time. In addition, the identification time will be affected by the number of the injected FDI attacks, as shown in Table 2.

**Table 2.** Detection and identification times.

| System | Detection Time (s) | Identification Time (s) |
|---|---|---|
| 6-bus | 0.8 | 1.2 |
| 39-bus | 1.5 | 3.8 |
| 118-bus | 2.2 | 4.5 |

## 6. Conclusions and Discussion

This paper proposes a novel detection and identification mechanism for countering FDI attacks in power systems. A novel detection model using UIOs and a cosine similarity theorem is constructed, using which FDI attacks can be detected. Furthermore, the developed CSV-based attack detection criterion can replace the design of a precomputed threshold. To lessen the impact of attacks on the system, an attack identification method is developed. Through the combination of UIOs, the influence of FDI attacks can minimized quickly. Finally, simulation experiments are carried out to verify the effectiveness of the proposed detection and identification framework on IEEE 6-bus, 39-bus and 118-bus power systems.

It is notable that the proposed method is only applicable to balanced power systems. For unbalanced power systems, operators can predict the state of electricity load using data-driven methods. Future works will further consider this problem by introducing artificial intelligence methods.

**Author Contributions:** Conceptualization, H.Z. and L.B.; methodology, H.Z.; writing—original draft preparation, X.W.; writing—review and editing M.S.; visualization, M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the third batch of school-level top-class construction projects of Tianjin College, University of Science and Technology Beijing, Motor and Drive Systems, YLKC202306; by the National Nature Science Foundation of China under 62103357; by the Hebei Natural Science Foundation under F2021203043; and by the Open Research Fund of Jiangsu Collaborative Innovation Center for Smart Distribution Network, Nanjing Institute of Technology, No.XTCX202203.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available on request due to restrictions eg privacy or ethical.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

**Definition of all parameters**

| | |
|---|---|
| $A_t$ | amplitude of the three-phase voltage |
| $\omega t$ | angular frequency |
| $\theta$ | angular phase |
| $V_1(k)$ | voltage signal of the 1st generator |
| $V_2(k)$ | voltage signal of the 2nd generator |
| $V_3(k)$ | voltage signal of the 3rd generator |
| $\Delta_k$ | model error and external disturbance, which is norm-bounded |
| $C$ | observation matrix of appropriate dimensions |
| $B$ | constant matrix of appropriate dimensions |
| $\lambda_k$ | output residual |

| | |
|---|---|
| $z_k$ | measurement output |
| $\tau$ | the precomputed detection threshold |
| $\hbar_k^f$ | output residual under FDI attacks |
| $z_k^f$ | measurement output under FDI attacks |
| $\hat{x}_k^a$ | state estimation under FDI attacks |
| $\ell$ | increment of state caused by the attack |
| $F$ | attack-selected matrix of appropriate dimensions |
| $\widetilde{A}^l$ | constant matrix of appropriate dimensions |
| $K^l$ | constant matrix of appropriate dimensions |
| $z_{k+1}^l$ | state vector of UIO |
| $\hat{x}_k^l$ | estimation value of $x_k^l$ |
| $H^l G^l U^l K^l$ | designed system parameter matrix |
| $P_6^l$ | constant matrix of appropriate dimensions |

## References


1. Oyewole, P.A.; Jayaweera, D. Power System Security with Cyber-Physical Power System Operation. *IEEE Access* **2020**, *8*, 179970–179982. [CrossRef]
2. Wang, C.; Jiang, C.; Wang, J.; Shen, S.; Guo, S.; Zhang, P. Blockchain-aided network resource orchestration in intelligent Internet of Things. *IEEE Internet Things J.* **2023**, *10*, 6151–6163. [CrossRef]
3. Lau, P.; Wang, L.; Liu, Z.; Wei, W.; Ten, C.-W. A Coalitional Cyber-Insurance Design Considering Power System Reliability and Cyber Vulnerability. *IEEE Trans. Power Syst.* **2021**, *36*, 5512–5524. [CrossRef]
4. Ashok, A.; Govindarasu, M.; Wang, J. Cyber-Physical Attack-Resilient Wide-Area Monitoring, Protection, and Control for the Power Grid. *Proc. IEEE* **2017**, *105*, 1389–1407. [CrossRef]
5. Das, L.; Munikoti, S.; Natarajan, B.; Srinivasan, B. Measuring smart grid resilience: Methods challenges and opportunities. *Renew. Sustain. Energy Rev.* **2021**, *130*, 109918. [CrossRef]
6. Liu, S.; Mashayekh, S.; Kundur, D.; Zourntos, T.; Butler-Purry, K. A framework for modeling cyber physical switching attacks in smart grid. *IEEE Trans. Emerg. Top. Comput.* **2014**, *1*, 273–285. [CrossRef]
7. Saini, A.; Bhui, P.; Singh, A.K.; Haq, F.U.; Kotakonda, C. Impact of False Data Injection Attacks in Wide Area Damping Control. In Proceedings of the 2022 22nd National Power Systems Conference (NPSC), New Delhi, India, 17–19 December 2022; pp. 218–223.
8. Xiao, M.; Wu, J.; Long, C.; Li, S. Construction of false sequence attack against PLC based power control system. In Proceedings of the 2016 35th Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016; pp. 10090–10095.
9. Yu, X.; Xue, Y. Smart grids: A cyber-physical systems perspective. *Proc. IEEE* **2016**, *104*, 1058–1070. [CrossRef]
10. Du, D.; Zhu, M.; Li, X.; Fei, M.; Bu, S.; Wu, L.; Li, K. A Review on Cybersecurity Analysis, Attack Detection, and Attack Defense Methods in Cyber-physical Power Systems. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 727–743. [CrossRef]
11. Shukla, R.M.; Sengupta, S. A novel machine learning pipeline to detect malicious anomalies for the Internet of Things. *Internet Things* **2022**, *20*, 100603. [CrossRef]
12. Gaggero, G.B.; Caviglia, R.; Armellin, A.; Rossi, M.; Girdinio, P.; Marchese, M. Detecting Cyberattacks on Electrical Storage Systems through Neural Network Based Anomaly Detection Algorithm. *Sensors* **2022**, *22*, 3933. [CrossRef]
13. Mashima, D.; Cárdenas, A.A. Evaluating Electricity Theft Detectors in Smart Grid Networks. In *Research in Attacks, Intrusions, and Defenses. RAID 2012*; Balzarotti, D., Stolfo, S.J., Cova, M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7462. [CrossRef]
14. Gaggero, G.B.; Rossi, M.; Girdinio, P.; Marchese, M. Detecting System Fault/Cyberattack within a Photovoltaic System Connected to the Grid: A Neural Network-Based Solution. *J. Sens. Actuator Netw.* **2020**, *9*, 20. [CrossRef]
15. Gunduz, H.; Jayaweera, D. Modern power system reliability assessment with cyber-intrusion on heat pump systems. *IET Smart Grid* **2020**, *3*, 561–571. [CrossRef]
16. Zhang, J.; Wang, X. Low-complexity quickest change detection in linear systems with unknown time-varying pre- and post-change distributions. *IEEE Trans. Inf. Theory* **2021**, *67*, 1804–1824. [CrossRef]
17. Nawaz, R.; Akhtar, R.; Shahid, M.A.; Qureshi, I.M.; Mahmood, M.H. Machine learning based false data injection in smart grid. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106819. [CrossRef]
18. Punmiya, R.; Choe, S. Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. *IEEE Trans. Smart Grid* **2019**, *10*, 2326–2329. [CrossRef]
19. Gallo, A.J.; Turan, M.S.; Boem, F.; Parisini, T.; Ferrari-Trecate, G. A Distributed Cyber-Attack Detection Scheme with Application to DC Microgrids. *IEEE Trans. Autom. Control* **2020**, *65*, 3800–3815. [CrossRef]
20. Lv, Y.W.; Yang, G.H. An adaptive cubature Kalman filter for nonlinear systems against randomly occurring injection attacks. *Appl. Math. Comput.* **2022**, *418*, 126834. [CrossRef]

21. Živković, N.; Sarić, A.T. Detection of false data injection attacks using unscented Kalman filter. *J. Mod. Power Syst. Clean Energy* **2018**, *6*, 847–859. [CrossRef]
22. Liu, Y.; Xue, W.; He, S.; Cheng, L. Stealthy False Data Injection Attacks against Extended Kalman Filter Detection in Power Grids. In Proceedings of the 2021 8th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS), Beijing, China, 10–12 December 2021; pp. 459–464.
23. Xiahou, K.; Liu, Y.; Wu, Q.H. Decentralized Detection and Mitigation of Multiple False Data Injection Attacks in Multiarea Power Systems. *IEEE J. Emerg. Sel. Top. Ind. Electron.* **2022**, *3*, 101–112. [CrossRef]
24. Kazemi, Z.; Safavi, A.A.; Naseri, F.; Urbas, L.; Setoodeh, P. A secure hybrid dynamic-state estimation approach for power systems under false data injection attacks. *IEEE Trans. Ind. Inform.* **2020**, *16*, 7275–7286. [CrossRef]
25. Wang, H.; Ruan, J.; Zhou, B.; Li, C.; Wu, Q.; Raza, M.Q.; Cao, G.Z. Dynamic data injection attack detection of cyber physical power systems with uncertainties. *IEEE Trans. Ind. Inform.* **2019**, *15*, 5505–5518. [CrossRef]
26. Manandhar, K.; Cao, X.; Hu, F.; Liu, Y. Detection of faults and attacks including false data injection attack in smart grid using Kalman filter. *IEEE Trans. Control Netw. Syst.* **2014**, *1*, 370–379. [CrossRef]
27. Tan, R.; Nguyen, H.H.; Foo, E.Y.; Yau, D.K.; Kalbarczyk, Z.; Iyer, R.K.; Gooi, H.B. Modeling and mitigating impact of false data injection attacks on automatic generation control. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 1609–1624. [CrossRef]
28. Luo, X.; Zhu, M.; Wang, X.; Guan, X. Detection and isolation of false data injection attack via adaptive Kalman filter bank. *J. Control Decis.* **2022**, 1–13. [CrossRef]
29. Venkateswaran, S.; Kravaris, C. Linear Unknown Input Observers for Sensor Fault Estimation in Nonlinear Systems. *IFAC-PapersOnLine* **2023**, *56*, 61–66. [CrossRef]
30. Pipeleers, G.; Vandenberghe, L. Generalized KYP Lemma with Real Data. *IEEE Trans. Autom. Control* **2011**, *56*, 2942–2946. [CrossRef]
31. Pertew, A.M.; Marquez, H.J.; Zhao, Q. H∞ observer design for lipschitz nonlinear systems. *IEEE Trans. Autom. Control* **2006**, *51*, 1211–1216. [CrossRef]
32. Zhang, Z.H.; Yang, G.H. Event-triggered fault detection for a class of discrete-time linear systems using interval observers. *ISA Trans.* **2017**, *68*, 160–169. [CrossRef]
33. Gerhat, B. Schur complement dominant operator matrices. *J. Funct. Anal.* **2023**, *286*, 110195. [CrossRef]
34. Chow, J.; Kokotovic, P. Time scale modeling of sparse dynamic networks. *IEEE Trans. Autom. Control* **2003**, *30*, 714–722. [CrossRef]