SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

April 2026, Vol. 69, Iss. 4, 142203:1–142203:17
https://doi.org/10.1007/s11432-025-4645-y

A dynamic encryption scheme for detecting FDI attacks in cyber-physical systems

Tongxiang LI[1], Bo CHEN[2,3*], Weiguo SHENG[4] & Wen-An ZHANG[2,3]

[1]*School of Automation and Electrical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China*
[2]*College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China*
[3]*Zhejiang Key Laboratory of Intelligent Perception and Control for Complex Systems, Hangzhou 310023, China*
[4]*School of Information Science and Technology, Hangzhou Normal University, Hangzhou 310030, China*

Received 27 March 2025/Revised 6 July 2025/Accepted 28 August 2025/Published online 20 January 2026

Abstract This paper presents an attack detection approach using data encryption for the cyber-physical systems with the purpose of (i) detecting the false data injection (FDI) attacks, and (ii) significantly improving the attack detection rate without sacrificing any system performance. To this end, a dynamic encryption scheme (DES) is designed to realize the encrypted transmission of measurement output, in which the secret key is updated dynamically using historical measurement outputs. Then, the effectiveness of DES in detecting FDI attacks is discussed through the theoretical analysis. It is certified that, the attack detection approach proposed in this paper can increase the anomaly detector's attack detection rate by accumulating the effects of FDI attacks on the estimation residual. When there are no attacks, the DES can restore the original measurement output. In this case, such a scheme would not deteriorate the system performance. Moreover, the proposed attack detection approach can also be used to detect replay attacks. Finally, the effectiveness of the proposed attack detection approach is demonstrated on an IEEE 6 bus power system.

Keywords cyber-physical systems, attack detection, FDI attacks, replay attacks, dynamic encryption

Citation Li T X, Chen B, Sheng W G, et al. A dynamic encryption scheme for detecting FDI attacks in cyber-physical systems. Sci China Inf Sci, 2026, 69(4): 142203, https://doi.org/10.1007/s11432-025-4645-y

1 Introduction

Cyber-physical systems (CPSs) are a kind of complex system integrating calculation, communication and control, aiming at realizing the expected performance of physical processes [1]. Nowadays, CPSs can be found in applications in different critical infrastructures [2]. However, due to the open communication network, CPSs are susceptible to malicious cyber-attacks, e.g., denial-of-service (DoS) attacks [3–7] and deception attacks [8–15], which would cause great negative impact on the system security [16–18]. It should be pointed out that the deception attacks, including false data injection (FDI) attacks [8–11] and replay attacks [12–15], are carefully designed, which can bypass the anomaly detection mechanism and seriously reduce the system performance. Thus, from the perspective of defenders, it is crucial to investigate the issue of attack detection to identify potential attacks and guarantee the security of CPSs.

1.1 Related work

In recent years, the issue of attack design and attack detection has received a lot of attention. Generally, the residual-based χ^2 detector, which distinguishes anomalies by using the statistical characteristics of system data, is applied to detect anomalies in control systems [19]. To highlight the security risks associated with CPSs, some recent studies have investigated attack strategies from the perspective of the attacker. There are two basic models to design FDI attacks: the stochastic attack strategy [20–22] and the deterministic attack strategy [8–11, 23–25]. Stochastic FDI attacks are crafted as random sequences to ensure that the residual sequences before and after the attacks have the same random distribution. These attacks can potentially avoid detection by anomaly detectors, but real-time system data are required to generate attack signals. Conversely, if the system model is known, deterministic attacks can be created offline. For the χ^2 detector, a deterministic FDI attack is developed in [8, 9], which only modifies the measurement outputs. In [11], this strategy is further developed to incorporate attacks on

* Corresponding author (email: bchen@zjut.edu.cn)

the control signals and measurement outputs. The deterministic FDI attacks were also investigated in [10] against state estimation and closed-loop control with network delays. In [23], a stealthy innovation-based attack scheme is developed that utilizes past and present estimation residuals to degrade the estimation performance while remaining stealthy to the anomaly detector. A finite-horizon strictly stealthy FDI attack, which innovation-based detectors cannot detect, is designed in [24]. In [25], a complete stealthy FDI attack is presented that completely nullifies its effect on estimation residuals (described in Definition 2 below), making it invisible to residual-based detectors.

From the viewpoint of the defender, various methods have been developed to expose FDI and replay attacks. In [26], the multiplicative watermarking scheme is given to increase the attack detection rate (ADR) by actively modifying measurement outputs and control inputs. In [27], an improved multiplicative watermarking scheme is introduced. An optimal linear encryption-based detection technique is presented in [28] to guarantee that the χ^2 detector can detect stochastic linear FDI attacks. To expose random linear FDI attacks in the problem of remote state estimation, watermarking-based methods are developed under χ^2 detector [29], in which the pseudo-random numbers are used as watermarks to encrypt and decrypt measurement outputs. In [30], by utilizing past and present measurement outputs, a summation (SUM) detector is proposed to expose FDI attacks. Furthermore, a noisy-control scheme was developed in [13] to detect replay attacks at the expense of control performance by introducing Gaussian noise into the control input. In [31], this strategy is improved even further by addressing the trade-off between detection efficiency and control performance. A stochastic game technique is presented in [32] to reduce the control performance loss under the noisy-control scheme. In [33], a periodic watermarking strategy is proposed for detecting the so-called discontinuous replay attack in CPSs. Besides, the stochastic coding scheme [14] and output coding scheme [15] are designed and applied in the transmission process of measurement outputs to detect replay attacks without compromising system performance.

1.2 Motivation and contributions

Generally, to satisfy CPS security requirements in engineering applications, an efficient attack detection approach should aim to achieve the following three objectives from the perspective of the defender.

- The ADR of the anomaly detector will improve significantly when the attack occurs.
- When there are no attacks, the attack detection approach will not damage the system performance.
- The attack detection approach should be capable of detecting various deception attacks, such as FDI and replay attacks.

To the best of our knowledge, it is still a challenging problem to realize the three aforementioned goals at the same time. For example, in the noisy-control-based approaches [13,31–33], the Gaussian noise superimposed on the control input will degrade system performance. The coding-based approaches [14, 15] only consider the problem of replay attack detection. Furthermore, it is another challenging problem to design an attack detection method that can effectively detect both FDI and replay attacks. The watermarking-based detection approaches [26, 27] and encryption-based detection approaches [28] can be used to detect FDI attacks, but are unable to detect replay attacks. In [29], the authors claim that the proposed method is also effective against replay attacks by using the time stamp of data packets. However, the time stamp can be easily modified by malicious attackers to render the detection method invalid. In summary, due to different natures of various cyber-attack types, it is difficult to develop a detection approach that can reliably detect both FDI and replay attacks.

Motivated by the above discussions, this paper will investigate the attack detection problem with the goal of detecting FDI and replay attacks to satisfy the three security requirements in CPSs. To this end, the dynamic encryption scheme (DES) is designed to encrypt and decrypt the measurement output. When the attack occurs, the DES can assist the anomaly detector in detecting both FDI and replay attacks. The main contributions of this paper can be summarized as follows.

(1) By utilizing the historical measurement outputs to dynamically update the secret key, a novel DES is first proposed to significantly improve the ADR of anomaly detector compared with the noisy-control-based approaches [13,31–33] and the coding-based approaches [14,15].

(2) Different from the noisy-control-based approaches [13, 31–33], the proposed DES can recover the original measurement output without degrading the system performance when there are no attacks.

(3) It is certified that through theoretical analysis and simulations, the designed attack detection approach can accumulate the attacked estimation residual to assist the anomaly detector in detecting both FDI attacks (Theorem 1, Theorem 2, Corollary 1) and replay attacks (Theorem 3).

Notations: $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote the n-dimensional Euclidean space and the set of all $n \times m$ real matrices, respectively. X^{T} represents the transpose of X. I_n is the identity matrix with n dimension. $diag\{\cdot\}$ stands for a

block diagonal matrix. $\|\cdot\|$ refers to the Euclidean norm of vectors or matrices. $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian distribution, with a mean of μ and a variance of σ^2.

2 Problem formulation

Consider the CPSs represented by the linear discrete-time system as follows:

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + w_k, \\ y_k = Cx_k + v_k, \end{cases} \tag{1}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ represent the system state, control input, and measurement output, respectively. The process noise $w_k \sim \mathcal{N}(0, Q)$ and measurement noise $v_k \sim \mathcal{N}(0, R)$ are the independent zero-mean Gaussian white noise. The initial state x_0 follows $x_0 \sim \mathcal{N}(0, \Sigma)$ and is independent of both w_k and v_k. It is assumed that (A, B) and (C, A) are completely controllable and observable matrix pairs, respectively.

The Kalman filter below is utilized to obtain the state estimation $\hat{x}_k$ from y_k [13]:

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_k + Bu_k, \\ P_{k+1|k} &= AP_kA^{\mathrm{T}} + Q, \\ L_k &= P_{k|k-1}C^{\mathrm{T}}(CP_{k|k-1}C^{\mathrm{T}} + R)^{-1}, \\ \hat{x}_k &= \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1}), \\ P_k &= P_{k|k-1} - L_kCP_{k|k-1}, \end{aligned} \tag{2}$$

where $\hat{x}_{0|-1} = \bar{x}_0, P_{0|-1} = \Sigma$. It can be seen from [13] that the Kalman filter gain L_k will converge to a constant value in a few steps, and then we define

$$P \triangleq \lim_{k\to\infty} P_{k|k-1},\ L \triangleq PC^{\mathrm{T}}(CPC^{\mathrm{T}} + R)^{-1}. \tag{3}$$

Further, the steady state form of the Kalman filter is given as follows:

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} + Bu_k, \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L(y_k - C\hat{x}_{k|k-1}), \end{aligned} \tag{4}$$

where $z_k = y_k - C\hat{x}_{k|k-1}$ is zero-mean Gaussian distributed which has a covariance of $P_z = CPC^{\mathrm{T}} + R$ [10], which is called the estimation residual. The controller $u_k = K\hat{x}_k$ is designed to ensure $\rho(A + BK) < 1$. Define $e_k \triangleq x_k - \hat{x}_k$ as the estimation error. Based on (1) and (4), e_k and z_k are derived as follows:

$$\begin{aligned} e_{k+1} &= \Phi e_k + (I - LC)w_k - Lv_{k+1}, \\ z_{k+1} &= CAe_k + Cw_k + v_{k+1}, \end{aligned} \tag{5}$$

where $\Phi = (A - LCA)$.

Generally, the χ^2 detector is used to identify anomalies in CPSs, which is defined as follows [19]:

$$g(z_k) = z_k^{\mathrm{T}} P_z^{-1} z_k. \tag{6}$$

Given that z_k is a zero-mean Gaussian distribution, then $g(z_k)$ follows χ^2 distributed with p degrees of freedom [19]. At every time, the χ^2 anomaly detector judges whether $g(z_k)$ is greater than η, in which η is selected to maintain a false alarm rate (FAR). If $g(z_k) > \eta$, the χ^2 anomaly detector will trigger an alarm. Otherwise, if $g(z_k) \leqslant \eta$, the system is considered normal. Define $\theta \triangleq Pr(g(z_k) > \eta)$ as the FAR, which will be a constant value in steady state.

As shown in Figure 1, the FDI attack taken into consideration in the existing studies is deemed to possess the following abilities [8–10]: (1) the attacker can access to the system parameters A, B, C, K and L; (2) the attacker is capable of injecting false data into the measurement output through communication channel. The compromised measurement output is modeled as follows:

$$y'_k = y_k + \Gamma a_k = Cx'_k + v_k + \Gamma a_k, \tag{7}$$

where a_k is the attack signals, $\Gamma = diag\{\gamma_1, \ldots, \gamma_p\}$ is the attack selection matrix. Here, $\gamma_j = 1$ means that the attacker injects false data into the jth communication channel, otherwise $\gamma_j = 0$.

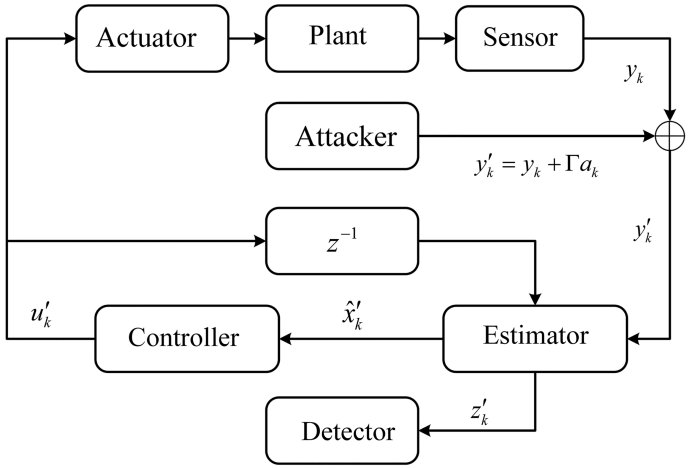

Figure 1 The block diagram of the CPSs under FDI attacks.

Let x'_k, y'_k and $\hat{x}'_k$ be the attacked system state, measurement output, and state estimation, respectively. Then, under the FDI attack (7), system models (1)–(4) are rewritten as follows:

$$\begin{aligned} x'_{k+1} &= Ax'_k + Bu'_k + w_k, \\ y'_k &= Cx'_k + \Gamma a_k + v_k, \\ \hat{x}'_{k+1|k} &= A\hat{x}'_k + Bu'_k, \\ \hat{x}'_{k+1} &= \hat{x}'_{k+1|k} + L(y'_{k+1} - C\hat{x}'_{k+1|k}), \\ u'_k &= K\hat{x}'_k, \end{aligned} \tag{8}$$

where $z'_k \triangleq y'_k - C\hat{x}'_{k|k-1}$ is the estimation residual under FDI attacks. In general, assume that the attacker begins at time 0 and $\hat{x}'_{-1} = \hat{x}_{-1}$.

Let $e'_k = x'_k - \hat{x}'_k$ be the attacked estimation error. To analyze the difference between the normal system (1)–(4) and the attacked system (8), we define $\Delta e_k \triangleq e'_k - e_k$ and $\Delta z_k \triangleq z'_k - z_k$, respectively. Furthermore, based on (5) and (8), Δe_k and Δz_k are obtained as follows:

$$\begin{aligned} \Delta e_{k+1} &= \Phi \Delta e_k - L\Gamma a_{k+1}, \\ \Delta z_{k+1} &= CA\Delta e_k + \Gamma a_{k+1}, \end{aligned} \tag{9}$$

where $\Delta e_{-1} = 0$ because FDI attack begins at time 0. Thus, one has $\Delta e_0 = -L\Gamma a_0$ and $\Delta z_0 = \Gamma a_0$.

In the existing literature [8–10], the purpose of the FDI attack is to cause the estimation error e'_k divergent and have a sufficiently small impact on estimation residual z'_k. Since $z'_k = z_k + \Delta z_k$, if $\|\Delta z_k\|$ is small enough, then the χ^2 anomaly detector will not be able to reliably identify z'_k and z_k. The definition of FDI attack in [8–10] is given as follows.

Definition 1 ([8–10])**.** The FDI attacks are stealthy if there is a constant α such that the error system (9) under FDI attacks satisfies

$$\lim_{k\to\infty} \|\Delta e_k\| = \infty, \|\Delta z_k\| \leqslant \alpha. \tag{10}$$

Lemma 1 ([8])**.** The existence condition of the FDI attack sequence in Definition 1 is that the system matrix A has at least one eigenvalue λ greater than 1, and the corresponding eigenvector v satisfies the conditions as follows:

(i) $Cv \in span(\Gamma)$;

(ii) $v \in span(V)$, where V is the controllability matrix pair $(A - LCA, L\Gamma)$.

Note that, although α is small enough in Definition 1, the FDI attack still brings some influence on the χ^2 detector. To remove the influences on the χ^2 detector, completely stealthy FDI attack is developed in [25], which can guarantee $\lim_{k\to\infty} \|\Delta z_k\| = 0$. The definition of completely stealthy FDI attack is described as follows.

Definition 2 ([25])**.** The FDI attacks have complete stealthiness if there is a constant α such that the error system (9) under FDI attacks satisfies

$$\lim_{k\to\infty} \|\Delta e_k\| = \infty, \|\Delta z_k\| \leqslant \alpha, \lim_{k\to\infty} \|\Delta z_k\| = 0. \tag{11}$$

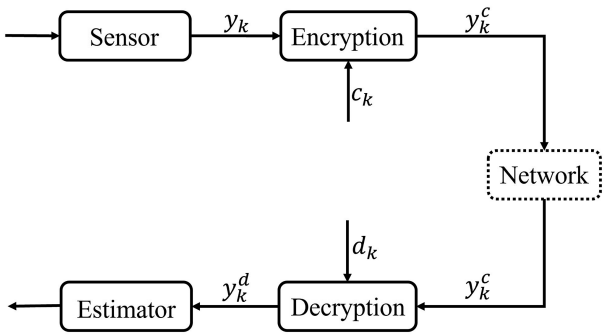

Figure 2 The DES without attacks.

Note that, the FDI attack in Definition 1 can make the estimation error divergent and remain undetected by χ^2 detector. Then, from the attackers' perspective, a completely stealthy FDI attack in Definition 2 can fully eliminate its impact on the estimation residual, that is, $\lim_{k\to\infty}\|\Delta z_k\| = 0$. Therefore, residual-based detectors, such as the χ^2 detector and SUM detector [30], fail to detect these completely stealthy FDI attacks because $\lim_{k\to\infty}\|\Delta z_k\| = 0$. Therefore, it is crucial to develop an efficient detection approach for these completely stealthy FDI attacks.

To sum up, this paper aims to achieve the following main goals.

- The first goal is to develop the DES to improve the ADR of the anomaly detector.
- The second goal is to ensure that the proposed DES does not damage the system performance when there are no attacks.

Remark 1. Compared with the FDI attacks in Definition 1, the completely stealthy FDI attacks in Definition 2 are more difficult to detect due to $\lim_{k\to\infty}\|\Delta z_k\| = 0$. In [25], a specific FDI attack was designed to achieve the complete stealthiness described in Definition 2, which will be used as an example to verify the effectiveness of our proposed attack detection approach.

3 A dynamic encryption scheme

In this section, the DES is developed to realize encrypted transmission of measurement output, in which the historical measurement outputs are used to dynamically update the secret key. As shown in Figure 2, the measurement output y_k is first encrypted using the secret key c_k. Before the encrypted y_k^c is fed into the estimator, the decryption scheme is applied to restore the measurement output. Concretely, the following encryption process is designed:

$$\begin{cases} y_k^c = y_k - c_k, \\ c_{k+1} = y_k + s_k, c_0 = s_0, \end{cases} \tag{12}$$

where y_k^c is the encrypted measurement output, c_k is a pre-designed secret key, and $s_k \sim \mathcal{N}(0, S)$ is produced by a random number generator. Then, y_k^c will be sent to the estimator. To restore y_k from y_k^c, the following decryption process is designed:

$$\begin{cases} y_k^d = y_k^c + d_k, \\ d_{k+1} = y_k^d + s_k, d_0 = s_0, \end{cases} \tag{13}$$

where y_k^d is the decrypted value of y_k, and d_k is a pre-designed secret key.

Note that the random number generator uses the same random seeds in (12) and (13) to ensure the consistency of s_k in the encryption and decryption process. Then, without attacks, it follows from (12) and (13) that

$$y_k^d = \begin{cases} y_0^c = y_0 - c_0 + d_0 = y_0, & \text{if } k = 0, \\ y_k^c + d_k = y_k - c_k + d_k, & \text{otherwise.} \end{cases} \tag{14}$$

According to (14), the following proposition is given.

Proposition 1. When there is no attacks, the y_k is restored from the decrypted measurement output y_k^d, i.e.,

$$y_k^d = y_k, d_{k+1} = c_{k+1}, k = 0, 1, 2, \ldots. \tag{15}$$

Proof. This proposition can be proven by mathematical induction.

Initial step. It can be seen from (14) that $y_0^d = y_0$ holds.

Inductive step. Suppose $y_k^d = y_k$ and $d_{k+1} = c_{k+1}$ hold for $k > 0$. Then, it can be obtained from (14) that

$$\begin{aligned} y_{k+1}^d &= y_{k+1} - c_{k+1} + d_{k+1} = y_{k+1} - y_k - s_k + y_k^d + s_k = y_{k+1}, \\ d_{k+2} &= y_{k+1}^d + s_{k+1} = y_{k+1} + s_{k+1} = c_{k+2}. \end{aligned}$$

According to mathematical induction, it can be concluded that Eq. (15) holds. This completes the proof.

In Proposition 1, y_k can be restored from y_k^d in the absence of attacks, i.e., $y_k^d = y_k$, which implies that the proposed DES (12) and (13) does not damage the system performance.

Remark 2. In the DES (12) and (13), the random signal s_k must remain consistent throughout both encryption and decryption. Since a given seed generates a unique random sequence [34], allowing us to use the same seed for both processes to ensure s_k remains consistent. In practical applications, seeds can be stored in advance on the on-board chip in the encryptor and decryptor to ensure seed synchronization. In fact, this solution is commonly used in cryptography [35]. Since encryption and decryption operations need to be performed at the sensor and the estimator, respectively, it is required that the sensor and the estimator have simple computing capabilities. Besides, considering the issue of data transmission delay, the timestamp of data packets can be used to ensure the synchronization of the encryption and decryption processes.

Remark 3. The most important feature of the proposed DES (12) and (13) is establishing a dynamic relationship between y_k, y_k^c, y_k^d and s_k. Particularly, when an FDI attack occurs, this dynamic relationship changes, aiding the anomaly detector (such as the χ^2 detector or SUM detector) in identifying FDI attacks. Therefore, different from the data encryption algorithms [36] in the information science (the purpose of these algorithms is to ensure data security by designing complex encryption functions), the main purpose of the DES (12) and (13) is to detect cyber-attacks. The next section will analyze the effectiveness of the proposed attack detection approach under FDI and replay attacks.

Remark 4. The DES (12) and (13) is a low computational complexity algorithm, which can be used for real-time closed-loop feedback control in CPSs. Specifically, from the perspective of computation, there are some linear operations designed in (12) and (13) that require very little computation. From the perspective of communication, the ciphertext y_k^c and measurement output y_k have the same dimension, and it does not require additional bits compared to the transmission of y_k. Besides, it should be pointed out that the proposed DES (12) and (13) does not use any system parameters except y_k, which means that this scheme is decoupled from system dynamics. Therefore, the proposed encryption-based attack detection approach is simple, practical and easy to extend.

4 Effectiveness analysis of the proposed scheme

4.1 Analysis of detecting FDI attacks

In this section, the effectiveness of DES (12) and (13) will be given in detecting FDI and replay attacks, respectively. As shown in Figure 3, under the DES, the FDI attack model (7) can be rewritten as follows:

$$y_k^a = y_k^c + \Gamma a_k. \tag{16}$$

Then, the following decryption process is given:

$$\begin{cases} y_k^d = y_k^a + d_k, \\ d_{k+1} = y_k^d + s_k. \end{cases} \tag{17}$$

Based on the (12), (16) and (17), it can be obtained by iterating from time 0 to k that

$$\begin{cases} y_k^d = y_k^a + d_k = y_k + \sum\limits_{i=0}^{k} \Gamma a_i, \\ d_{k+1} = y_k + \sum\limits_{i=0}^{k} \Gamma a_i + s_k. \end{cases} \tag{18}$$

Furthermore, according to (1) and (4), the dynamics of attacked system model under DES (12) and (13) is rewritten as follows:

$$x_{k+1}^a = Ax_k^a + Bu_k^a + w_k,$$

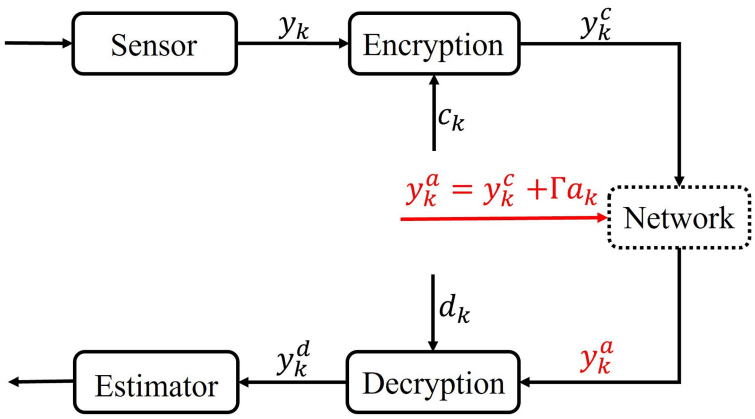

Figure 3 (Color online) The DES under FDI attacks.

$$\begin{aligned}\hat{x}_{k+1|k}^a &= A\hat{x}_k^a + Bu_k^a,\\ \hat{x}_{k+1}^a &= \hat{x}_{k+1|k}^a + L(y_{k+1}^d - C\hat{x}_{k+1|k}^a),\\ u_k^a &= K\hat{x}_k^a,\end{aligned} \tag{19}$$

where $\hat{x}_k^a$ and $z_k^a = y_k^d - C\hat{x}_{k|k-1}^a$ are the attacked state estimation and residual, respectively. Define $g(z_k^a) = (z_k^a)^{\mathrm{T}} P_z^{-1} z_k^a$ and $\mathbb{P}_k^a \triangleq Pr(g(z_k^a) > \eta)$. Compared with z_k', under the DES (12) and (13), the estimation residual z_k^a is changed after the FDI attack occurs. Accordingly, we need to obtain the dynamics of estimation residual z_k^a to analyze the effectiveness of the DES (12) and (13).

Let $e_k^a = x_k^a - \hat{x}_k^a$ be the attacked estimation error. Then, e_k^a and z_k^a are obtained as follows:

$$\begin{aligned}e_{k+1}^a &= \Phi e_k^a - L\sum_{i=0}^{k+1}\Gamma a_i + (I - LC)w_k - Lv_{k+1},\\ z_{k+1}^a &= CAe_k^a + \sum_{i=0}^{k+1}\Gamma a_i + Cw_k + v_{k+1}.\end{aligned} \tag{20}$$

Furthermore, we define $\Delta e_k^a = e_k^a - e_k$ and $\Delta z_k^a = z_k^a - z_k$, respectively. Based on (5) and (20), the dynamics of Δe_k^a and Δz_k^a are obtained as follows:

$$\begin{aligned}\Delta e_{k+1}^a &= \Phi\Delta e_k^a - L\sum_{i=0}^{k+1}\Gamma a_i,\\ \Delta z_{k+1}^a &= CA\Delta e_k^a + \sum_{i=0}^{k+1}\Gamma a_i,\end{aligned} \tag{21}$$

where $\Delta e_{-1}^a = \Delta e_{-1} = 0$, $\Delta e_0^a = \Delta e_0 = -L\Gamma a_0$ and $\Delta z_0^a = \Delta z_0 = \Gamma a_0$.

Proposition 2. The dynamics of Δe_k^a and Δz_k^a satisfy the following condition:

$$\Delta e_k^a = \sum_{i=0}^{k}\Delta e_i, \quad \Delta z_k^a = \sum_{i=0}^{k}\Delta z_i. \tag{22}$$

Proof. Combining (21) with the fact that $\Delta e_{-1}^a = \Delta e_{-1} = 0$, we have

$$\begin{aligned}\Delta e_{k+1}^a - \Delta e_k^a &= \Phi\Delta e_k^a - L\sum_{i=0}^{k+1}\Gamma a_i - \Phi\Delta e_{k-1}^a + L\sum_{i=0}^{k}\Gamma a_i\\ &= \Phi\left(\Delta e_k^a - \Delta e_{k-1}^a\right) - L\Gamma a_{k+1} = \Phi^{k+1}\Delta e_0^a - L\sum_{i=0}^{k+1}\left[\Phi^i\Gamma a_{k+1-i}\right],\end{aligned}$$

$$\begin{aligned}\Delta z_{k+1}^a - \Delta z_k^a &= CA\Delta e_k^a + \sum_{i=0}^{k+1}\Gamma a_i - CA\Delta e_{k-1}^a - \sum_{i=0}^{k}\Gamma a_i \\ &= CA\left(\Delta e_k^a - \Delta e_{k-1}^a\right) + \Gamma a_{k+1} = CA\left[\Phi^k \Delta e_0^a - L\sum_{i=0}^{k}\left(\Phi^i \Gamma a_{k-i}\right)\right] + \Gamma a_{k+1}. \end{aligned} \tag{23}$$

On the other hand, Eq. (9) is derived by iteration as follows:

$$\begin{aligned}\Delta e_{k+1} &= \Phi^{k+1}\Delta e_0 - L\sum_{i=0}^{k+1}\left(\Phi^i \Gamma a_{k+1-i}\right), \\ \Delta z_{k+1} &= CA\left[\Phi^k \Delta e_0 - L\sum_{i=0}^{k}\left(\Phi^i \Gamma a_{k-i}\right)\right] + \Gamma a_{k+1}. \end{aligned} \tag{24}$$

Then, due to $\Delta e_0^a = \Delta e_0$, substituting (24) into (23) leads to

$$\begin{aligned}\Delta e_{k+1}^a - \Delta e_k^a &= \Delta e_{k+1}, \\ \Delta z_{k+1}^a - \Delta z_k^a &= \Delta z_{k+1}. \end{aligned} \tag{25}$$

Furthermore, based on (25), Eq. (22) can be obtained, which implies that the proof is complete.

Based on Proposition 2, the effectiveness analysis of the DES (12) and (13) in detecting the FDI attack in Definition 1 is given in the following theorem. First, to derive our main results, the following lemma is necessary.

Lemma 2 ([25])**.** For the χ^2 detector with the threshold η, the following statements are satisfied when the system (1) subjects to the FDI attack:

(i) $\mathbb{P}_k' \geqslant \mathbb{P}_k$, if $\|\Delta z_k\| \leqslant \alpha$, $\forall k \in \mathbb{N}[0,\infty)$;

(ii) $\lim_{k\to\infty}\mathbb{P}_k' = \lim_{k\to\infty}\mathbb{P}_k$, if $\lim_{k\to\infty}\|\Delta z_k\| = 0$;

(iii) $\lim_{k\to\infty}\mathbb{P}_k' > \lim_{k\to\infty}\mathbb{P}_k$, if $\lim_{k\to\infty}\|\Delta z_k\| \neq 0$;

(iv) $\lim_{k\to\infty}\mathbb{P}_k' = 1$, if $\lim_{k\to\infty}\|\Delta z_k\| = \infty$,

where $\mathbb{P}_k = Pr(g(z_k) > \eta)$, $\mathbb{P}_k' = Pr(g(z_k') > \eta)$, $g(z_k) = z_k^{\mathrm{T}} P_z^{-1} z_k$, and $g(z_k') = z_k'^{\mathrm{T}} P_z^{-1} z_k'$. $\mathbb{P}_k$ and $\mathbb{P}_k'$ are the FAR and ADR of the χ^2 detector, respectively.

Theorem 1. Considering the system (8) under an arbitrary FDI attack satisfying Definition 1, using the DES (12) and (13), the condition (10) is transformed into

$$\lim_{k\to\infty}\|\Delta e_k^a\| = \infty, \|\Delta z_k^a\| \leqslant \infty, \tag{26}$$

which implies that $\lim_{k\to\infty}\mathbb{P}_k^a = 1$, i.e., the FDI attack in Definition 1 is no longer stealthy under the χ^2 detector.

Proof. Based on the condition (10) in Definition 1, it can be obtained from (22) in Proposition 2 that

$$\begin{aligned}&\lim_{k\to\infty}\|\Delta e_k^a\| = \infty, \\ &\lim_{k\to\infty}\|\Delta z_k^a\| = \lim_{k\to\infty}\left\|\sum_{i=0}^{k}\Delta z_i\right\| \overset{1}{\leqslant} \lim_{k\to\infty}\sum_{i=0}^{k}\|\Delta z_i\| \overset{2}{\leqslant} \lim_{k\to\infty} k\alpha = \infty, \end{aligned} \tag{27}$$

where the triangle inequality is used in step 1 of the proof, and the condition (10) in Definition 1 is used in step 2 of the proof. Therefore, it follows from Lemma 2 that $\lim_{k\to\infty}\mathbb{P}_k^a = 1$ due to $\lim_{k\to\infty}\|\Delta z_k^a\| = \infty$, which implies that the FDI attack in Definition 1 is no longer stealthy under the χ^2 detector. The proof is complete.

Next, the effectiveness analysis of the DES (12) and (13) in detecting the completely stealthy FDI attack in Definition 2 is given in the following theorem.

Theorem 2. Considering the system (8) under an arbitrary FDI attack satisfying Definition 2, using the DES (12) and (13), the condition (11) is transformed into

$$\lim_{k\to\infty}\|\Delta e_k^a\| = \infty, \lim_{k\to\infty}\|\Delta z_k^a\| \leqslant \lim_{k\to\infty}\sum_{i=0}^{k}\|\Delta z_i\| \triangleq \beta, \tag{28}$$

which implies that $\lim_{k\to\infty}\mathbb{P}_k^a > \lim_{k\to\infty}\mathbb{P}_k$, i.e., the FDI attack in Definition 2 is no longer completely stealthy under the χ^2 detector, where β is a calculable constant corresponding to a specific FDI attack.

Proof. Based on the condition (11) in Definition 2, it can be obtained from (22) in Proposition 2 that

$$\begin{aligned}&\lim_{k\to\infty}\|\Delta e_k^a\|=\infty,\\&\lim_{k\to\infty}\|\Delta z_k^a\|=\lim_{k\to\infty}\left\|\sum_{i=0}^{k}\Delta z_i\right\|\overset{1}{\leqslant}\lim_{k\to\infty}\sum_{i=0}^{k}\|\Delta z_i\|\triangleq\beta,\end{aligned}\tag{29}$$

where the triangle inequality is used in step 1 of the proof. Furthermore, due to $\lim_{k\to\infty}\|\Delta z_k\|=0$ in Definition 2, it can be obtained that β is a constant, which can be calculated corresponding to a specific FDI attack. Therefore, it follows from Lemma 2 that $\lim_{k\to\infty}\mathbb{P}_k^a>\lim_{k\to\infty}\mathbb{P}_k$ due to $\lim_{k\to\infty}\|\Delta z_k^a\|\neq 0$, which implies that the FDI attack in Definition 2 is no longer completely stealthy under the χ^2 detector. The proof is now complete.

In fact, a specific FDI attack sequence a_k was designed in [25] to achieve the complete stealthiness in Definition 2, which takes the following form:

$$\begin{cases}a_k-a_{k-1}=-\rho_1\lambda^{k+1}a^*, & 1\leqslant k\leqslant\varepsilon,\\ a_k-a_{k-1}=-\rho_2\lambda^{k+1}a^*, & k\geqslant\varepsilon,\end{cases}\tag{30}$$

where λ is a unstable eigenvalue of A, v is a corresponding eigenvector, i.e., $Av=\lambda v$, ρ_1 and ρ_2 are non-zero constants, and the constant vector a^* satisfies $Cv=\Gamma a^*$. Meanwhile, the following Lemma shows the conditions for FDI attack (30) to achieve complete stealthiness.

Lemma 3 ([25])**.** Consider the system (8), the FDI attack (30) achieves complete stealthiness if and only if

$$(\lambda^{\varepsilon+1}-1)\rho_1-\lambda^{\varepsilon+1}\rho_2=0.\tag{31}$$

Under condition (31), the FDI attack (30) can realize the complete stealthiness as described in Definition 2, i.e., $\lim_{k\to\infty}\|\Delta e_k^a\|=\infty,\|\Delta z_k^a\|\leqslant\alpha,\lim_{k\to\infty}\|\Delta z_k^a\|=0$. In this case, we present the following Corollary 1 to give the effectiveness analysis of the DES (12) and (13) in detecting the completely stealthy FDI attack (30). The following concept about w norm is first given to further obtain Corollary 1.

Lemma 4 ([25])**.** Define the matrix w norm in $\mathbb{R}^{n\times n}$ as $\|Y\|_w\triangleq\|(D_\epsilon U)Y(D_\epsilon U)^{-1}\|_\infty$, where $D_\epsilon=diag\{1,\epsilon,\epsilon^2,\ldots,\epsilon^n\}$, $\epsilon<1-p(X)$, $X\in\mathbb{R}^{n\times n}$ is a given matrix, $\rho(X)<1$, and U is an invertible matrix so that $J=UXU^{-1}$ is the Jordan canonical form of X. Then, the following conditions hold:

(i) $\|X\|_w<1$;

(ii) for $\gamma\in\mathbb{R}^n$, its vector w norm $\|\gamma\|_w\triangleq\|D_\epsilon U\gamma\|_\infty$ is compatible with matrix w norm, i.e., $\|Y\gamma\|_w\leqslant\|Y\|_w\|\gamma\|_w$;

(iii) for $\gamma\in\mathbb{R}^n$, the vector w norm $\|\gamma\|_w\triangleq\|D_\epsilon U\gamma\|_\infty$ satisfies $\|\gamma\|\leqslant\sqrt{n}\|(D_\epsilon U)^{-1}\|\|v\|_w$.

Corollary 1. Considering the system (8) subjected to the FDI attack (30) under condition (31), using dynamic encryption scheme (12) and (13), the condition (11) is transformed into

$$\begin{aligned}&\lim_{k\to\infty}\|\Delta e_k^a\|=\infty,\\&\lim_{k\to\infty}\|\Delta z_k^a\|\leqslant\sqrt{n}\,|\rho_1|\left\|(D_\epsilon U)^{-1}\right\|\|CA\|\,\bar\kappa\triangleq\beta,\end{aligned}\tag{32}$$

where $D_\epsilon=diag\{1,\epsilon,\epsilon^2,\ldots,\epsilon^n\}$, $\epsilon<1-\rho(A-LCA)$ and U is the Jordan canonical form of $A-LCA$. $\bar\kappa$ is a positive constant satisfying $\bar\kappa=\kappa_1(\varepsilon)+\bar\kappa_2$, where $\kappa_1(k)$ is defined as $\kappa_1(k)\triangleq\sum_{l=0}^{k}\sum_{i=0}^{l}\|A-LCA\|_w^i\|v\|_w$ and $\bar\kappa_2$ is defined as $\bar\kappa_2\triangleq\sum_{i=0}^{\varepsilon+1}\|A-LCA\|_w^i\|v\|_w/(1-\|A-LCA\|_w)$.

Proof. Please see Appendix A.

At present, we have given the effectiveness analysis of the DES (12) and (13) in detecting FDI attacks (including completely stealthy FDI attacks). Next, we will discuss the effectiveness of the DES (12) and (13) in detecting replay attacks.

4.2 Analysis of detecting replay attacks

Consider the replay attack model studied by [13], which consists of two stages.

(1) The attacker records measurement outputs from time $-\tau$ to -1 as attack signals, in which τ is sufficiently large to capture an adequate number of attack signals.

(2) Consider the replay attack starting at time 0. In this stage, the attacker replaces y_k with $y_{k-\tau}$. Then, under the DES (12) and (13), the replay attack is given as follows:

$$y_k^{rc}=y_{k-\tau}^c=y_{k-\tau}-c_{k-\tau},\tag{33}$$

where y_k^{rc} is the replay attack signals and $c_{k-\tau}=y_{k-\tau-1}+s_{k-\tau-1}$.

Theorem 3. Considering the system (8) subjected to replay attack (33), using (12) and (13), the ADR of χ^2 detector tends to $\lim_{k\to\infty}\mathbb{P}_k^a = 1$.

Proof. Under the replay attack (33), the decryption process is given as follows:

$$\begin{cases} y_k^d = y_k^{rc} + d_k, \\ d_{k+1} = y_k^d + s_k. \end{cases} \tag{34}$$

Combining with (33) and (34), it can be obtained by iterating from time 0 to k that

$$\begin{cases} y_k^d = y_k^{rc} + d_k = y_{k-\tau} - c_{k-\tau} + d_k \\ \quad\quad = y_{k-\tau} - y_{k-\tau-1} - s_{k-\tau-1} + y_{k-1}^d + s_{k-1}, = y_{k-\tau} + y_{-1} - y_{-\tau-1} + \sigma_k, \\ d_{k+1} = y_k^d + s_k, \end{cases} \tag{35}$$

where $\sigma_k \sim \mathcal{N}(0, 2(k+1)S)$ and its variance can be easily obtained according to the derivation process of (35) under replay attack (33). Then, it follows from (35) that the decrypted measurement output y_k^d is divergent when $k \to \infty$ under replay attack (33), which implies that $z_k^a = y_k^d - C\hat{x}_{k|k-1}^a$ is also divergent and $\lim_{k\to\infty}\|\Delta z_k^a\| = \infty$. Therefore, it can be obtained from Lemma 2 that $\lim_{k\to\infty}\mathbb{P}_k^a = 1$.

Remark 5. In [30], a residual-based SUM detector is designed as follows:

$$g(z_k^s) = \tfrac{1}{k}(z_k^s)^{\mathrm{T}} P_z^{-1} z_k^s, \tag{36}$$

which is proven to have superior detection capabilities to the χ^2 detector, where $z_k^s = \sum_{i=0}^k z_i$, and z_i is the estimation residual. According to the conclusion in [30], the SUM detector is capable of detecting the FDI attack in Definition 1 due to $\lim_{k\to\infty}\|\Delta z_k\| \neq 0$, and it cannot detect the completely stealthy FDI attack in Definition 2 due to $\lim_{k\to\infty}\|\Delta z_k\| = 0$. Noticed that, using the DES (12) and (13), it can be easy obtained that $\lim_{k\to\infty}\|\Delta z_k^s\| = \lim_{k\to\infty}\|\sum_{i=0}^k \Delta z_i^a\| = \infty$. According to the conclusion in [30], if $\lim_{k\to\infty}\|\Delta z_k^s\| = \infty$, one has $\lim_{k\to\infty}\mathbb{P}_k^a = 1$. Therefore, using the DES (12) and (13), it can be easily proven that the completely stealthy FDI attack can be detected by the SUM detector. The stealthiness analysis of FDI attacks and replay attacks under both χ^2 detector and SUM detector will be given in Section 5.

So far, we have demonstrated that the proposed DES (12) and (13) is effective in detecting FDI attacks (including completely stealthy FDI attacks) and replay attacks, and this approach does not damage the system performance without attacks. Particularly, our proposed approach is developed independently of the anomaly detector, so it can be directly applied based on the anomaly detector.

5 Simulations

In this section, an IEEE 6 bus power system [25], illustrated in Figure 4, is adopted to illustrate the effectiveness of the DES (12) and (13) in detecting FDI and replay attacks. According to [37], the system model is represented in the form of model (1) as follows:

$$A = \begin{bmatrix} I_3 & T\cdot I_3 \\ -M_g^{-1}(L_{gl}L_{ll}^{-1}L_{\mathrm{lg}} - L_{gg})T & I_3 - M_g^{-1}D_gT \end{bmatrix}, B = \begin{bmatrix} 0 \\ M_g^{-1}T \end{bmatrix}, \tag{37}$$

where $T = 0.1$ s is sampling period, $x_k = [\delta_k^{\mathrm{T}}, \omega_k^{\mathrm{T}}]^{\mathrm{T}}$. $\delta_k = [\delta_1(k), \delta_2(k), \delta_3(k)]^{\mathrm{T}}$ and $\omega_k = [\omega_1(k), \omega_2(k), \omega_3(k)]^{\mathrm{T}}$ represent the generator rotor angles and frequencies. $u_k = P_\omega(k) - L_{gl}L_{ll}^{-1}P_\theta(k) - \omega_d(k)D_g$ is equivalent control inputs, in which $\omega_d(k)$, $P_\omega(k)$ and $P_\theta(k)$ are the expected generator frequency, mechanical power input and real power, respectively. Let $M_g = diag\{0.125, 0.034, 0.016\}$, $D_g = diag\{0.125, 0.068, 0.48\}$, $L_{gg} = diag\{0.058, 0.063, 0.059\}$ and

$$L_{gl} = \begin{bmatrix} -0.058 & 0 & 0 & 0\ 0\ 0 \\ 0 & -0.063 & 0 & 0\ 0\ 0 \\ 0 & 0 & -0.059 & 0\ 0\ 0 \end{bmatrix}, L_{lg} = L_{gl}^{\mathrm{T}},$$

$$L_{ll} = \begin{bmatrix} 0.235 & 0 & 0 & -0.085 & -0.092 & 0 \\ 0 & 0.296 & 0 & -0.161 & 0 & -0.072 \\ 0 & 0 & 0.329 & 0 & -0.17 & -0.101 \\ -0.085 & -0.161 & 0 & 0.246 & 0 & 0 \\ -0.092 & 0 & -0.17 & 0 & 0.262 & 0 \\ 0 & -0.072 & -0.101 & 0 & 0 & 0.173 \end{bmatrix}.$$

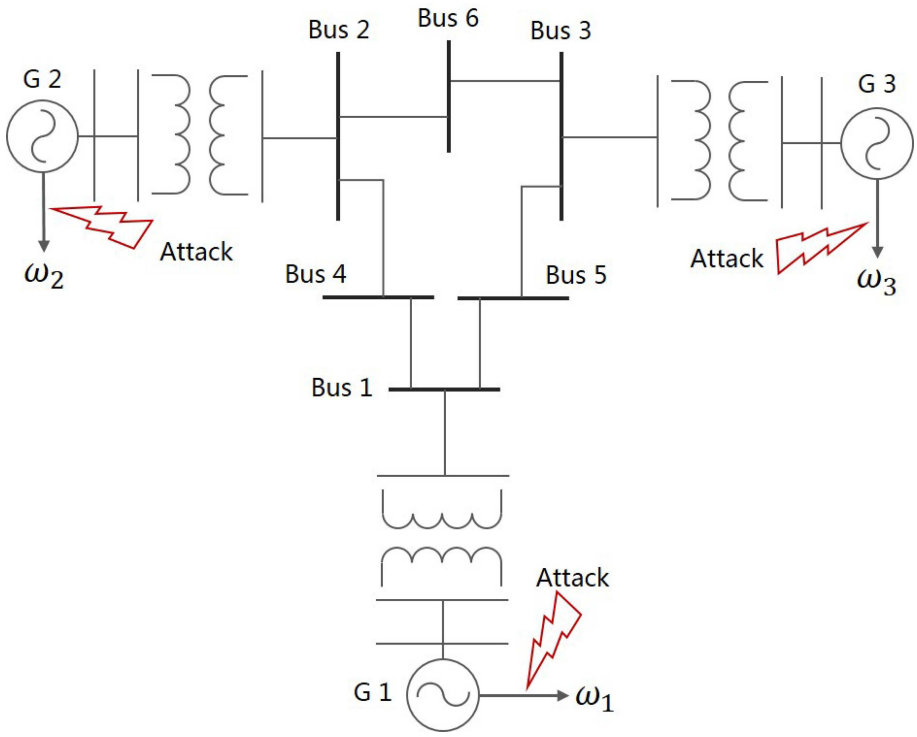

Figure 4 (Color online) The IEEE 6 bus power systems under attacks.

The measurement output is $y_k = [I_3, I_3]x_k + v_k$. Consider $w_k \sim \mathcal{N}(0, I_6)$ and $v_k \sim \mathcal{N}(0, I_3)$. The controller gain K and the filter gain L are designed as follows:

$$K = \begin{bmatrix} -0.058 & 0 & 0 & 0 & 0 & 0 \\ -0.695 & 0.015 & 0.016 & -0.696 & 0 & 0 \\ 0.015 & -0.327 & 0.015 & 0 & -0.278 & 0 \\ 0.014 & 0.014 & -0.173 & 0 & 0 & 0.299 \end{bmatrix}, L^{\mathrm{T}} = \begin{bmatrix} 0.394 & 0 & 0.002 & 0.339 & -0.002 & -0.003 \\ -0.02 & 0.449 & 0 & 0.018 & 0.288 & -0.003 \\ -0.003 & 0.002 & 0.253 & 0.001 & -0.005 & 0.672 \end{bmatrix}.$$

Consider a scenario where the measurement outputs y_k are compromised by an FDI attack. Meanwhile, the χ^2 detector (6) is employed to identify system anomalies. The threshold η for both the χ^2 detector (6) and SUM detector (36) is set to $\eta = 12.84$ corresponding to the FAR $\theta = 0.005$.

5.1 Detection results of FDI attacks

In this subsection, we present the detection results of χ^2 detector and the SUM detector under FDI attacks using the DES (12) and (13).

5.1.1 *FDI attack I with complete stealthiness satisfying Definition 2*

It is obvious that A contains unstable eigenvalues: -2.0059, 1.0417 and 1.0178. The eigenvector of unstable eigenvalue $\lambda = 1.0417$ is $v = [0.2275, -0.879, 0.0474, 0.1157, -0.3364, 0.0198]^{\mathrm{T}}$. Let the attack selection matrix be $\Gamma = I_3$. Choosing $\rho_1 = -0.9811$ and $\varepsilon = 4$, it can be obtained from Lemma 3 that $\rho_2 = -0.1822$. Then, according to (30), the completely stealthy FDI attack sequence a_k in attack model (7) is given as follows [25]:

$$\begin{cases} a_k - a_{k-1} = 0.9811 \times 1.0417^{k+1} a^*, 1 \leqslant k \leqslant 4, \\ a_k - a_{k-1} = 0.1822 \times 1.0417^{k+1} a^*, k > 4, \end{cases} \tag{38}$$

where a^* satisfying $Cv = \Gamma a^*$, and $a_0 = -\rho_1 \lambda a^*$.

First, we give the results of Δe_k, Δz_k and $\sum_{l=0}^{k} \Delta z_l$ shown in Figure 5 to illustrate the complete stealthiness of FDI attack sequence (38). It can be observed from Figure 5 that $\lim_{k\to\infty} \|\Delta e_k\| = \infty$, $\|\Delta z_k\| \leqslant \alpha$ and $\lim_{k\to\infty} \|\Delta z_k\| = 0$ (i.e., the complete stealthiness condition in Definition 2), which means that FDI attack sequence (38) is completely stealthy under χ^2 detector (6) and SUM detector (36), as shown in Figure 6.

Furthermore, using the DES (12) and (13), Δe_k^a, Δz_k^a and $\sum_{l=0}^{k} \Delta z_l^a$ are shown in Figure 7. The detection results under χ^2 detector (6) and the SUM detector (36) are exhibited in Figure 8. It can be seen from Figure 7 that

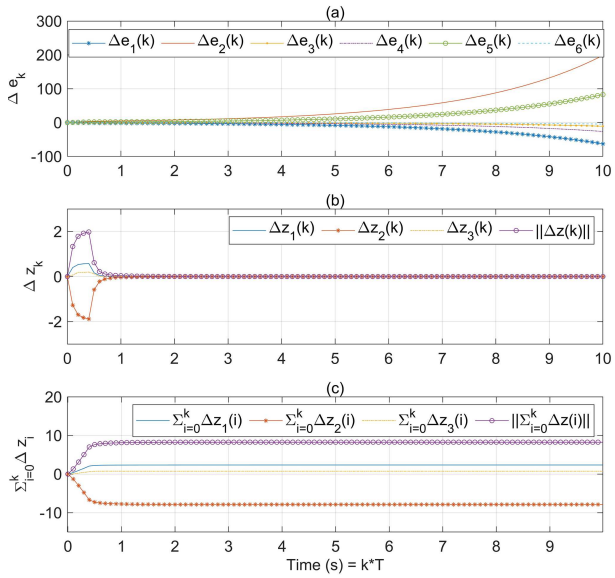

Figure 5 (Color online) The curves of (a) Δe_k, (b) Δz_k, and (c) $\sum_{l=0}^{k} \Delta z_l$ under FDI attack I.

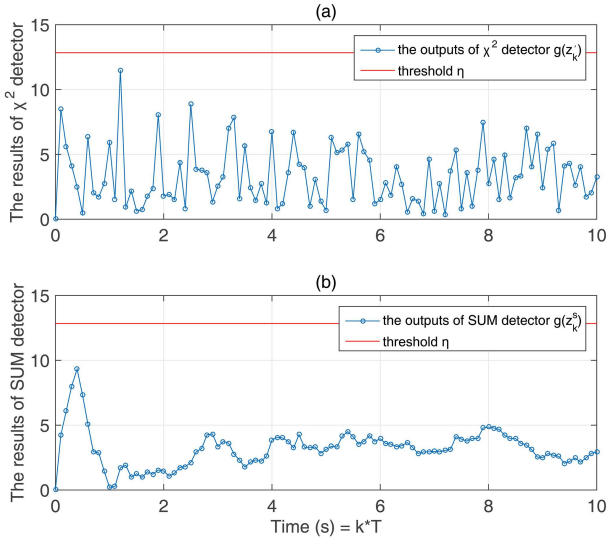

Figure 6 (Color online) The detection results of (a) χ^2 detector and (b) SUM detector under FDI attack I.

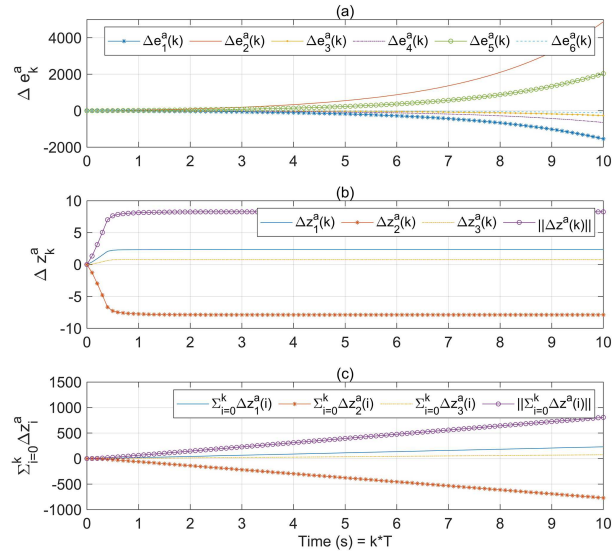

Figure 7 (Color online) The curves of (a) Δe_k^a, (b) Δz_k^a, and (c) $\sum_{l=0}^{k} \Delta z_l^a$ under FDI attack I using the DES.

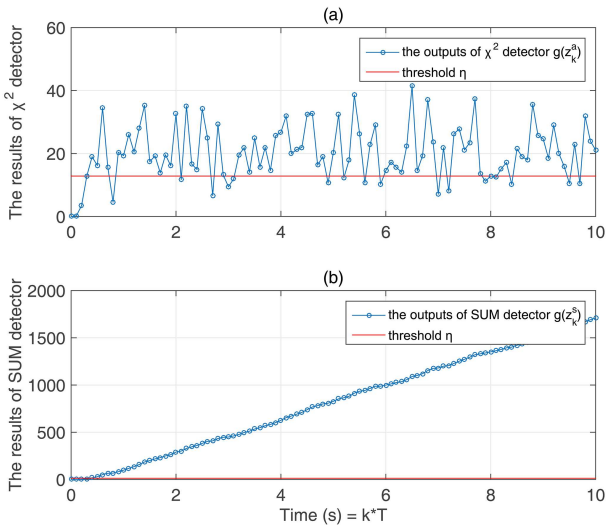

Figure 8 (Color online) The detection results of (a) χ^2 detector and (b) SUM detector under FDI attack I using the DES.

$\lim_{k\to\infty} \|\Delta e_k^a\| = \infty$, $\lim_{k\to\infty} \|\Delta z_k^a\| \neq 0$ and $\|\sum_{l=0}^{k} \Delta z_l^a\|$ are divergent. Figure 8 shows that the ADR of the χ^2 detector increases significantly, and the ADR of the SUM detector tends to 1. Thus, it can be concluded from Figures 5–8 that the FDI attack sequence (38) is no longer completely stealthy using the DES (12) and (13). This aligns with the results in Theorem 2, Corollary 1, and Remark 5.

5.1.2 *FDI attack II satisfying Definition 1*

Let $\rho_1 = \rho_2 = -0.9811$, the completely stealthy FDI attack sequence (38) degenerates into the FDI attack sequence as follows [8]:

$$a_k - a_{k-1} = 0.9811 \times 1.0417^{k+1} a^*, k \geqslant 1, \tag{39}$$

which only satisfies the condition (10) in Definition 1.

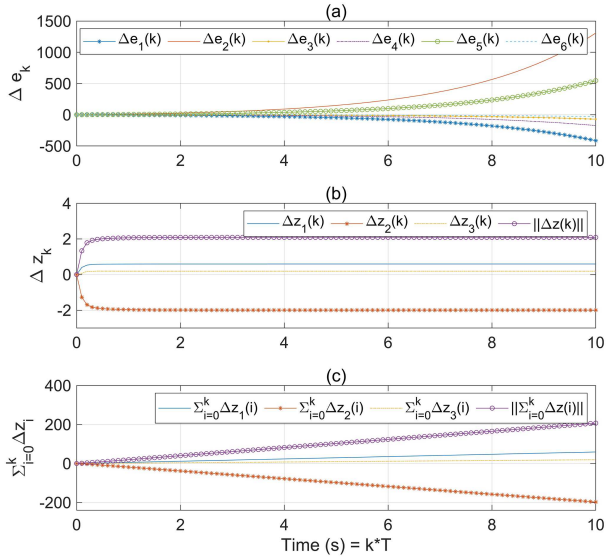

Figure 9 (Color online) The curves of (a) Δe_k, (b) Δz_k, and (c) $\sum_{l=0}^{k}\Delta z_l$ under FDI attack II.

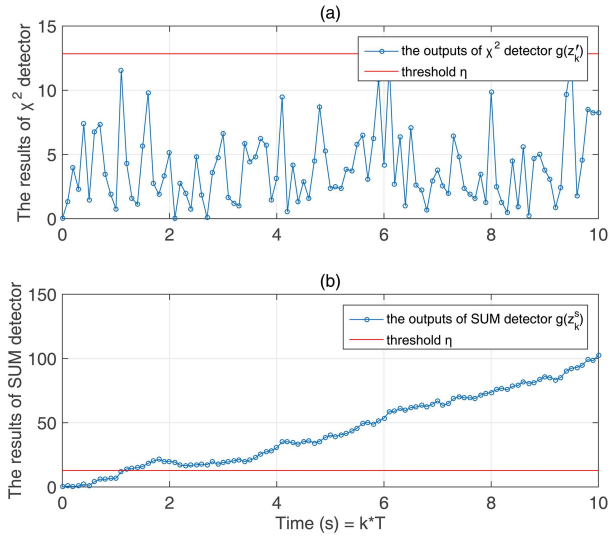

Figure 10 (Color online) The detection results of (a) χ^2 detector and (b) SUM detector under FDI attack II.

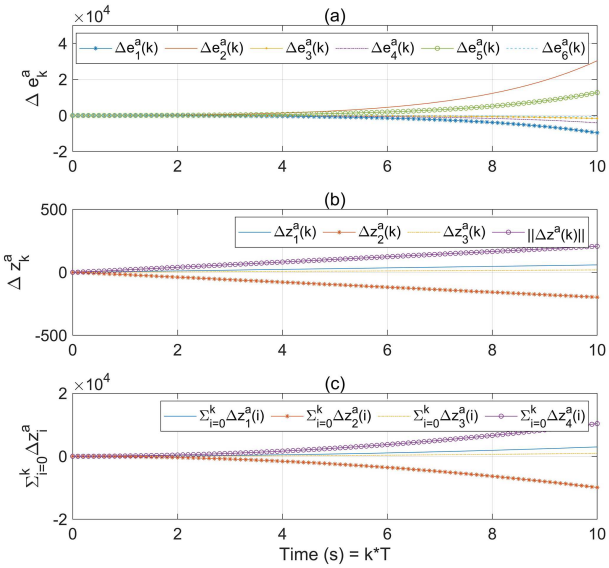

Figure 11 (Color online) The curves of (a) Δe_k^a, (b) Δz_k^a, and (c) $\sum\limits_{l=0}^{k}\Delta z_l^a$ under FDI attack II using the DES.

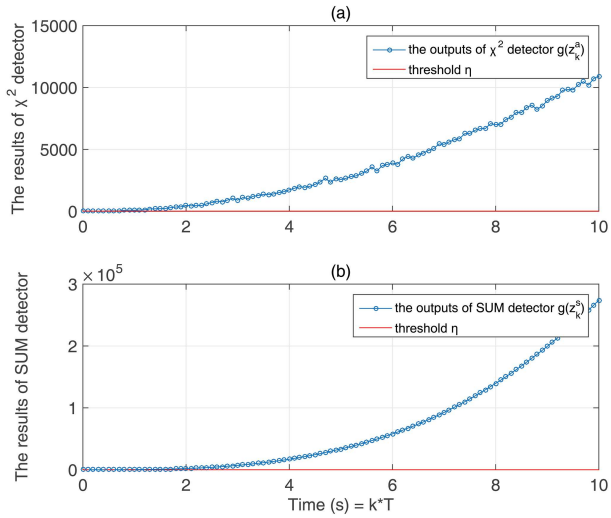

Figure 12 (Color online) The detection results of (a) χ^2 detector and (b) SUM detector under FDI attack II using the DES.

Figure 9 shows the results of Δe_k, Δz_k and $\sum_{l=0}^{k}\Delta z_l$ under the FDI attack (39). It can be seen from Figure 9 that $\lim_{k\to\infty}\|\Delta e_k\|=\infty$, $\lim_{k\to\infty}\|\Delta z_k\|\neq 0$ and $\|\sum_{l=0}^{k}\Delta z_l\|$ are divergent, which means that the condition (10) in Definition 1 is satisfied. Meanwhile, according to Remark 5, the ADR of the SUM detector tends to 1 due to $\lim_{k\to\infty}\|\Delta z_k\|\neq 0$, which can be verified by the results in Figure 10.

Then, using the DES (12) and (13), Δe_k^a, Δz_k^a and $\sum_{l=0}^{k}\Delta z_l^a$ are shown in Figure 11, and the detection results of χ^2 detector (6) and SUM detector (36) are depicted in Figure 12. Figure 11 exhibits $\lim_{k\to\infty}\|\Delta e_k\|=\infty$, $\lim_{k\to\infty}\|\Delta z_k\|=\infty$ and $\lim_{k\to\infty}\|\sum_{l=0}^{k}\Delta z_l\|=\infty$, which implies that the ADR of both χ^2 detector (6) and SUM detector (36) obeys $\lim_{k\to\infty}\mathbb{P}_k^a=1$, as shown in Figure 12. Thus, it can be concluded from Figures 9–12 that the FDI attack sequence (39) is no longer stealthy using the DES (12) and (13), which is consistent with the results in Theorem 1 and Remark 5. In addition, it should be pointed out that with the increase of attack intensity, the attack detection probability increases. When the attack intensity is constant, if the noise intensity is large enough,

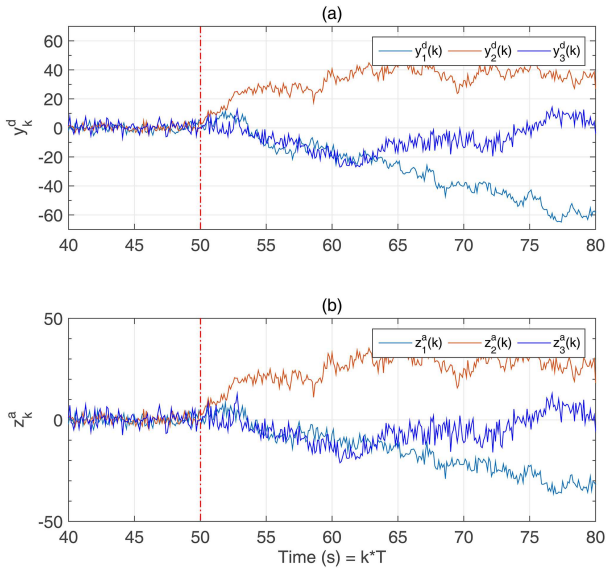

Figure 13 (Color online) The curves of (a) y_k^d and (b) z_k^a under replay attack using the DES.

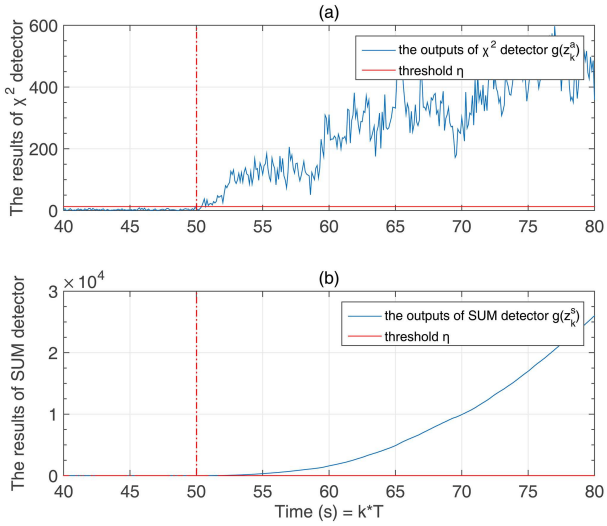

Figure 14 (Color online) The detection results of (a) χ^2 detector and (b) SUM detector under replay attacks using the DES.

it will make it difficult for the detector to distinguish between attack and noise, thus reducing the attack detection probability. In practical applications, the noise intensity is usually required to be at a low level to ensure the estimation accuracy.

5.2 Detection results of replay attacks

In this subsection, we would like to show the detection results with the χ^2 detector and SUM detector under replay attacks using the DES (12) and (13). The attacker records measurement outputs from $k = 0$ to $k = 499$ as attack signals. Then, attack signals replayed at $k = 500$ (shown by the red dotted line in Figure 13). The threshold η of χ^2 detector (6) and SUM detector (36) is also set to $\eta = 12.84$ corresponding to the FAR $\theta = 0.005$.

Figure 13 shows the decrypted measurement outputs y_k^d and estimation residuals z_k^a. The detection results of the χ^2 detector (6) and SUM detector (36) are displayed in Figure 14. It observes from Figure 13 that the distributions of decrypted measurement outputs y_k^d and estimation residuals z_k^a are normal without replay attack (from time $k = 0$ to $k = 499$), which are consistent with the results in Proposition 1. When replay attacks occur at $k = 500$, the distributions of y_k^d and z_k^a are obviously divergent, which will increase the ADR of the χ^2 detector (6) and SUM detector (36). Actually, it is obtained from Figure 14 that the outputs of χ^2 detector and the SUM detector are completely larger than the threshold η, which implies that the ADR of both the χ^2 detector and the SUM detector tends to 1. Accordingly, the results presented in Figures 13 and 14 align with the results drawn in Theorem 3 and Remark 5.

6 Conclusion

This paper examines attack detection in CPSs under FDI attacks using data encryption. A DES has been proposed to assist anomaly detectors in detecting FDI attacks. The effectiveness of the DES in detecting FDI attacks has been proven through theoretical analysis. It observes that the DES can significantly improve the ADR of the anomaly detector and that such a scheme does not degrade system performance without attacks. In particular, our proposed approach is proven effective at detecting replay attacks. Finally, we conducted simulations on an IEEE 6-bus power system to verify the effectiveness of our proposed approach in detecting both FDI attacks and replay attacks. Our future work will consider the joint problem of attack detection and distributed security estimation under communication bandwidth constraints to weaken the impact of attacks on the estimation accuracy.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62503145, 92367205), Joint Funds of the National Natural Science Foundation of China (Grant No. U24A20258), Zhejiang Provincial Natural Science Foundation of China (Grant No. LRG25F030001), Key Research and Development Projects of Zhejiang Province (Grant Nos. 2025C2007, 2023C01022), and Funding of Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2023R01006).

References

1 Duo W, Zhou M C, Abusorrah A. A survey of cyber attacks on cyber physical systems: recent advances and challenges. IEEE CAA J Autom Sin, 2022, 9: 784–800
2 Ding D R, Han Q-L, Ge X H, et al. Privacy-preserving filtering, control and optimization for industrial cyber-physical systems. Sci China Inf Sci, 2025, 68: 141201
3 Chen B, Ho D W C, Zhang W A, et al. Distributed dimensionality reduction fusion estimation for cyber-physical systems under DoS attacks. IEEE Trans Syst Man Cybern Syst, 2017, 49: 455–468
4 Li T, Chen B, Yu L, et al. Active security control approach against DoS attacks in cyber-physical systems. IEEE Trans Automat Contr, 2020, 66: 4303–4310
5 Xu W, Hu G, Ho D W C, et al. Distributed secure cooperative control under denial-of-service attacks from multiple adversaries. IEEE Trans Cybern, 2019, 50: 3458–3467
6 Wen H, Li Y M, Tong S C. Distributed adaptive resilient formation control for nonlinear multi-agent systems under DoS attacks. Sci China Inf Sci, 2024, 67: 209201
7 Su W, Mu C X, Zhu S, et al. Event-triggered leader-follower bipartite consensus control for nonlinear multi-agent systems under DoS attacks. Sci China Inf Sci, 2025, 68: 132206
8 Mo Y, Sinopoli B. False data injection attacks in control systems. In: Proceedings of the 1st Workshop on Secure Control Systems, 2010. 1: 1–6
9 Mo Y, Sinopoli B. On the performance degradation of cyber-physical systems under stealthy integrity attacks. IEEE Trans Automat Contr, 2015, 61: 2618–2624
10 Hu L, Wang Z, Han Q L, et al. State estimation under false data injection attacks: security analysis and system protection. Automatica, 2018, 87: 176–183
11 Sui T, Mo Y, Marelli D, et al. The vulnerability of cyber-physical system under stealthy attacks. IEEE Trans Automat Contr, 2020, 66: 637–650
12 Li T, Wang Z, Zou L, et al. A dynamic encryption-decryption scheme for replay attack detection in cyber-physical systems. Automatica, 2023, 151: 110926
13 Mo Y, Sinopoli B. Secure control against replay attacks. In: Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing, 2009. 911–918
14 Ye D, Zhang T Y, Guo G. Stochastic coding detection scheme in cyber-physical systems against replay attack. Inf Sci, 2019, 481: 432–444
15 Guo H, Pang Z-H, Sun J, et al. An output-coding-based detection scheme against replay attacks in cyber-physical systems. IEEE Trans Circuits Syst II Express Briefs, 2021, 68: 3306–3310
16 Fidler D P. Was Stuxnet an act of war? Decoding a cyberattack. IEEE Secur. Privacy, 2011, 9: 56–59
17 Zhang D, Wang Q-G, Feng G, et al. A survey on attack detection, estimation and control of industrial cyber-physical systems. ISA Trans, 2021, 116: 1–16
18 Li X-M, Zou T, Lu R Q, et al. Event-based non-fragile state estimation for time-varying systems under deception attacks. Sci China Inf Sci, 2025, 68: 142204
19 Mehra R K, Peschon J. An innovations approach to fault detection and diagnosis in dynamic systems. Automatica, 1971, 7: 637–640
20 Guo Z, Shi D, Johansson K H, et al. Worst-case stealthy innovation-based linear attack on remote state estimation. Automatica, 2018, 89: 117–124
21 Shang J, Yu H, Chen T. Worst-case stealthy innovation-based linear attacks on remote state estimation under Kullback-Leibler divergence. IEEE Trans Automat Contr, 2021, 67: 6082–6089
22 Ye D, Yang B, Zhang T Y. Optimal stealthy linear attack on remote state estimation with side information. IEEE Syst J, 2021, 16: 1499–1507
23 Li Y G, Yang G H. Optimal stealthy innovation-based attacks with historical data in cyber-physical systems. IEEE Trans Syst Man Cybern Syst, 2019, 51: 3401–3411
24 Cheng D, Shang J, Chen T. Finite-horizon strictly stealthy deterministic attacks on cyber-physical systems. IEEE Control Syst Lett, 2021, 6: 1640–1645
25 Zhang T Y, Ye D. False data injection attacks with complete stealthiness in cyber-physical systems: a self-generated approach. Automatica, 2020, 120: 109117
26 Miao F, Zhu Q, Pajic M, et al. Coding schemes for securing cyber-physical systems against stealthy data injection attacks. IEEE Trans Control Netw Syst, 2016, 4: 106–117
27 Ferrari R M G, Teixeira A M H. A switching multiplicative watermarking scheme for detection of stealthy cyber-attacks. IEEE Trans Automat Contr, 2020, 66: 2558–2573
28 Shang J, Chen M, Chen T. Optimal linear encryption against stealthy attacks on remote state estimation. IEEE Trans Automat Contr, 2020, 66: 3592–3607
29 Huang J, Ho D W C, Li F, et al. Secure remote state estimation against linear man-in-the-middle attacks using watermarking. Automatica, 2020, 121: 109182
30 Ye D, Zhang T Y. Summation detector for false data-injection attack in cyber-physical systems. IEEE Trans Cybern, 2019, 50: 2338–2345
31 Mo Y, Chabukswar R, Sinopoli B. Detecting integrity attacks on SCADA systems. IEEE Trans Contr Syst Technol, 2013, 22: 1396–1407
32 Miao F, Pajic M, Pappas G J. Stochastic game approach for replay attack detection. In: Proceedings of the 52nd IEEE Conference on Decision and Control, 2013. 1854–1859
33 Fang C, Qi Y, Cheng P, et al. Optimal periodic watermarking schedule for replay attack detection in cyber-physical systems. Automatica, 2020, 112: 108698
34 Menezes A J, van Oorschot P C, Vanstone S A. Handbook of Applied Cryptography. Boca Raton: CRC Press, 2018
35 Blum M, Micali S. How to generate cryptographically strong sequences of pseudo random bits. In: Providing Sound Foundations for Cryptography: on the Work of Shafi Goldwasser and Silvio Micali. New York: Association for Computing Machinery, 2019. 227–240
36 Katz J, Lindell Y. Introduction to Modern Cryptography: Principles and Protocols. New York: Chapman & Hall/CRC, 2007
37 Wei Q, Liu D, Lin Q, et al. Discrete-time optimal control via local policy iteration adaptive dynamic programming. IEEE Trans Cybern, 2016, 47: 3367–3379

Appendix A Proof of Corollary 1

To present the proof of Corollary 1, we first give the following lemma and proposition.

Lemma A1 ([25]). If $A-LCA$ is strictly stable, i.e., $\rho(A-LCA)<1$, then $y-(A-LCA)y=c$ has a unique solution for a constant vector c.

Proposition A1. Considering the FDI attack (30), the dynamics of Δe_k^a and Δz_k^a is derived as follows:

$$\Delta e_{k+1}^a-\Delta e_k^a=\sum_{i=1}^{k-\varepsilon+1}\Phi^i(\rho_1-\rho_2)\lambda^{\varepsilon+1}v-\sum_{i=1}^{k+2}\Phi^i\rho_1 v+\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^{i+1}v+\sum_{i=0}^{\varepsilon}\rho_1\lambda^{i+1}v, \tag{A1}$$

$$\Delta z_{k+1}^a-\Delta z_k^a=(\rho_1-\rho_2)\lambda^{\varepsilon+1}CA\sum_{i=0}^{k-\varepsilon}\Phi^i v-\rho_1 CA\sum_{i=0}^{k+1}\Phi^i v, \tag{A2}$$

where $\Phi=(A-LCA)$.

Proof. For $0\leqslant k<\varepsilon$, according to (21), it can be obtained that

$$\begin{aligned}\Delta e_{k+1}^a-\Delta e_k^a&=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a\right)-L\Gamma a_{k+1}=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a\right)+L\sum_{i=0}^{k+1}\rho_1\lambda^i CAv\\&=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a-\sum_{i=0}^{k+1}\rho_1\lambda^i v\right)+\sum_{i=0}^{k+1}\rho_1\lambda^{i+1}v,\end{aligned} \tag{A3}$$

which can be further expressed by iteration as follows:

$$\Delta e_{k+1}^a-\Delta e_k^a=\Phi^{k+1}\left(\Delta e_0^a-\Delta e_{-1}^a-\rho_1\lambda v\right)-\sum_{i=1}^{k+1}\Phi^i\rho_1 v+\sum_{i=0}^{k+1}\rho_1\lambda^{i+1}v=-\sum_{i=0}^{k+1}\Phi^{i+1}\rho_1 v+\sum_{i=0}^{k+1}\rho_1\lambda^{i+1}v, \tag{A4}$$

where $\Delta e_{-1}^a=0$ and $\Delta e_0^a=-L\Gamma a_0=\rho_1 LCAv$. For $k\geqslant\varepsilon$,

$$\begin{aligned}\Delta e_{k+1}^a-\Delta e_k^a&=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a\right)-L\Gamma a_{k+1}=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a\right)+L\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^i CAv+L\sum_{i=0}^{\varepsilon}\rho_1\lambda^i CAv\\&=\Phi\left(\Delta e_k^a-\Delta e_{k-1}^a-\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^i v-\sum_{i=0}^{\varepsilon}\rho_1\lambda^i v\right)+\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^{i+1}v+\sum_{i=0}^{\varepsilon}\rho_1\lambda^{i+1}v.\end{aligned} \tag{A5}$$

Because the attack parameters are converted from ρ_1 to ρ_2 at $\varepsilon+1$, one has

$$\begin{aligned}\Delta e_{k+1}^a-\Delta e_k^a=\Phi^{k-\varepsilon+1}&\left(\Delta e_\varepsilon^a-\Delta e_{\varepsilon-1}^a-\rho_2\lambda v-\sum_{i=0}^{\varepsilon}\rho_1\lambda^i v\right)-\sum_{i=1}^{k-\varepsilon}\Phi^i(\rho_2-\rho_1)\lambda^{\varepsilon+1}v-\sum_{i=1}^{k-\varepsilon}\Phi^i\rho_1 v\\&+\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^{i+1}v+\sum_{i=0}^{\varepsilon}\rho_1\lambda^{i+1}v.\end{aligned} \tag{A6}$$

Substituting (A4) into (A6), Eq. (A1) can be obtained. For $0\leqslant k<\varepsilon$, combining with (A4), it can be obtained that

$$\begin{aligned}\Delta z_{k+1}^a-\Delta z_k^a&=CA\left(\Delta e_k^a-\Delta e_{k-1}^a\right)+\Gamma a_{k+1}=CA\left(\Delta e_k^a-\Delta e_{k-1}^a\right)-\Gamma\sum_{i=0}^{k+1}\rho_1\lambda^{i+1}a^*\\&=CA\left(-\sum_{i=0}^{k}\Phi^{i+1}\rho_1 v+\sum_{i=0}^{k}\rho_1\lambda^{i+1}v\right)-\sum_{i=0}^{k+1}\rho_1\lambda^{i+1}Cv=-\sum_{i=0}^{k+1}\rho_1 CA\Phi^i v.\end{aligned} \tag{A7}$$

Next, for $k\geqslant\varepsilon$, it is derived from (A6) that

$$\begin{aligned}\Delta z_{k+1}^a-\Delta z_k^a&=CA\left(\Delta e_k^a-\Delta e_{k-1}^a\right)+\Gamma a_{k+1}=CA\left(\Delta e_k^a-\Delta e_{k-1}^a\right)-\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^{i+1}\Gamma a^*-\sum_{i=0}^{\varepsilon}\rho_1\lambda^{i+1}\Gamma a^*\\&=(\rho_1-\rho_2)\lambda^{\varepsilon+1}CA\sum_{i=0}^{k-\varepsilon}\Phi^i v-\rho_1 CA\sum_{i=0}^{k+1}\Phi^i v,\end{aligned} \tag{A8}$$

which implies that the proof is complete.

Proof of Corollary 1 follows.

Proof. According to (A1) in Proposition 4, it is derived that

$$\lim_{k\to\infty}\left(\Delta e^a_{k+1}-\Delta e^a_k\right)=\sum_{i=\varepsilon+1}^{k+1}\rho_2\lambda^{i+1}v+\sum_{i=0}^{\varepsilon}\rho_1\lambda^{i+1}v. \tag{A9}$$

Due to $\lambda>1$, it can be easy obtained from (A9) that

$$\lim_{k\to\infty}\|\Delta e^a_k\|=\infty. \tag{A10}$$

For $0\leqslant k<\varepsilon$, it follows from (A7) that

$$\Delta z^a_k=-\sum_{l=0}^{k}\sum_{i=0}^{l}\rho_1 CA\Phi^i v. \tag{A11}$$

Then, according to Lemma 6, we have

$$\|\Delta z^a_k\|\leqslant|\rho_1|\|CA\|\left\|\sum_{l=0}^{k}\sum_{i=0}^{l}\Phi^i v\right\|\leqslant\sqrt{n}|\rho_1|\|(D_\epsilon U)^{-1}\|\|CA\|\kappa_1(k), \tag{A12}$$

where $\kappa_1(k)\triangleq\sum_{l=0}^{k}\sum_{i=0}^{l}\|\Phi\|_w^i\|v\|_w$ is an increasing function, i.e., $\kappa_1(k)\leqslant\kappa_1(\varepsilon)$. Accordingly,

$$\|\Delta z^a_k\|\leqslant\sqrt{n}|\rho_1|\|(D_\epsilon U)^{-1}\|\|CA\|\kappa_1(\varepsilon). \tag{A13}$$

For $0\leqslant k<\varepsilon$, it can be derived from (A8) that

$$\Delta z^a_k=-\sum_{l=\varepsilon+1}^{k}\sum_{i=0}^{l-\varepsilon-1}(\rho_1-\rho_2)\lambda^{\varepsilon+1}CA\Phi^i v-\sum_{l=\varepsilon+1}^{k}\sum_{i=0}^{l}\rho_1 CA\Phi^i v-\sum_{l=0}^{\varepsilon}\sum_{i=0}^{l}\rho_1 CA\Phi^i v, \tag{A14}$$

which is known from condition (31) that

$$\|\Delta z^a_k\|\leqslant|\rho_1|\|CA\|\left(\left\|\sum_{l=0}^{\varepsilon}\sum_{i=0}^{l}\Phi^i v\right\|+\left\|\sum_{l=\varepsilon+1}^{k}\sum_{i=l-\varepsilon}^{l}\Phi^i v\right\|\right)\leqslant\sqrt{n}|\rho_1|\|(D_\epsilon U)^{-1}\|\|CA\|(\kappa_1(\varepsilon)+\bar\kappa_2), \tag{A15}$$

which $\kappa_2(k)\triangleq\sum_{l=\varepsilon+1}^{k}\sum_{i=l-\varepsilon}^{l}\|A-LCA\|_w^i\|v\|_w$. Obviously, $\kappa_2(k)$ is increasing with k, which means $\bar\kappa_2\triangleq\lim_{k\to\infty}\kappa_2(k)\geqslant\kappa_2(k)$. Due to $\|A-LCA\|_w<1$, it observes from Lemma A1 that

$$\kappa_2(k)-\|\Phi\|_w\kappa_2(k)=\sum_{i=0}^{\varepsilon+l}\|\Phi\|_w^i\|v\|_w \tag{A16}$$

has a unique constant vector solution, which indicates

$$\bar\kappa_2\triangleq\lim_{k\to\infty}\kappa_2(k)=\sum_{i=0}^{\varepsilon+1}\|\Phi\|_w^i\|v\|_w/(1-\|\Phi\|_w). \tag{A17}$$

Let $\bar\kappa=\kappa_1(\varepsilon)+\bar\kappa_2$. Then, we have

$$\lim_{k\to\infty}\|\Delta z^a_k\|\leqslant\sqrt{n}\,|\rho_1|\,\left\|(D_\epsilon U)^{-1}\right\|\|CA\|\,\bar\kappa. \tag{A18}$$

Therefore, it follows from Lemma 2 that $\lim_{k\to\infty}\mathbb{P}^a_k>\lim_{k\to\infty}\mathbb{P}_k$ due to $\lim_{k\to\infty}\|\Delta z^a_k\|\neq0$, which means that the FDI attack (30) is no longer completely stealthy under the χ^2 detector. The proof is complete.