International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

False Data Injection Attack Detection and Estimation in Smart Grid: An Observer-based Approach

Zahra Molavi-Nafchi[1] | AbdorezaRabiee[1] | Maryam Shahriari-kahkeshi[1*]

Electrical Engineering Department, Faculty of Engineering, Shahrekord University, Shahrekord, Iran.[1]
*Corresponding author's email: m.shahriyarikahkeshi@alumni.iut.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: **********
Received in revised form: **********
Accepted: **********
Published online: **********

Keywords: Adaptive observer, Attack detection and estimation, Attack estimator, Cyber-attacks in smart grids.

ABSTRACT

In this paper, an adaptive observer-based cyber-attack detection and estimation problem is studied for smart grid under false data injection attack (FDIA). By invoking mathematical model of the smart grid, sliding mode observer (SMO) is proposed for estimating grid states and generating residual signal. By monitoring grid status, evaluating residual signal and comparing it with the appropriate threshold level, attack detection is done and alarm signal is generated. Upon alarm generation, attack estimation algorithm is activated to estimate the FDIA occurred at the vulnerable buses. In the proposed detection scheme, only the frequency deviations of the generator buses are required; also, no off-line learning phase and no prior knowledge about attack signal are required. Moreover, the proposed scheme is able to detect FDIA and exactly estimates the severity of the detected attack. The Lyapunov stability theorem is used to guarantee stability of the proposed state observer and the proposed attack estimation algorithm, separately. Simulation results for the IEEE six-bus network under single-point and multi-point FDIAs show that upon attack occurrence, the defined evaluation function exceeds from a defined threshold level, which makes attack detection after some milliseconds. Also, the reported mean square error shows that the proposed attack estimation strategy can precisely estimate the attack shape and severity and identify the under-attack bus.

I. Introduction

Smart grid is a typical cyber-physical system characterized by cooperation between cyber space and physical infrastructure. It integrates the power generation and distribution devices on the utility side into the smart meters on the consumer side by using the communication networks and information technology [1, 2]. Dispatching and production management in the smart grids highly depend on the communication system. Therefore, any anomaly in information exchange as well as any cyber-attack influence the safe operation of the smart grid, disrupt the grid operation and impose risks such as unpredictable outage and financial losses [3].

One of the important cyber-attacks targeting the power system is the false data injection attack (FDIA). In the FDIA, attacker injects manipulated data to change the load as the main target in the transmission channels among the physical system and cyber space and disturb normal operation [4-6]. As an example of FDIA at the power system, in 2015, the hackers penetrate the supervisory control and data acquisition (SCADA) system and caused a wide blackout in Kiev and west Ukraine, another cyber-attack in 2016 caused the shut-down of 200 MW power generation in the part of Kiev [7]. The 2019 attack on Venezuela's national grid caused outages of all 11 states except the capital, the 2020 attack in India caused a 12-hour blackout of the power grid and the 2022 attack in Germany disrupted the communication system used to monitor and control 2000 wind turbines are some examples of the latest successful cyber-attacks on the power systems [8, 9].

Studying security problems in smart grids, and especially detection of possible malicious attacks due to the complexity, and vulnerability of the smart grids is an important problem [10-12]. Therefore, at present, FDIA detection problem as a mature research field has received great attention [13-21]. In general, FDIA detection approaches can be categorized into data-based and model-based schemes. One prevalent scheme among the model-

based approaches is to design an observer (also known as estimator) to estimate the systems states or any modeling uncertainty [13]. In [13], q-analogue of the Bernstein-Schurer-Stancu operator was proposed to approximate model uncertainty and then model-free observer was proposed to estimate velocity, after that an adaptive control scheme was proposed for cooperative robots. In [14], an unknown input interval observer-based FDIA detection and isolation algorithm was proposed for smart grid. In [15], an optimized sliding mode observer (SMO)-based attack detection scheme was proposed for power systems under FDIA. It uses evolutionary algorithm to tune observer parameters offline. In [16], an adaptive sliding mode observer was proposed for detection and reconstruction of cyber-attacks in power systems. In [17], an observer-based cyber-attack detection and isolation algorithm was suggested for large scale smart grid systems. In [19], the power system security problem under the malicious process attack was studied and a resilient control approach based on the integral sliding mode surface and intermediate variable observer was proposed. In [21], the maximum a posteriori estimation method was used for target node localization in wireless sensor networks.

Load change attack in smart grid can be categorized into load redistribution attack (LRA) [22] and load altering attack (LAAs) [23, 24]. In the LRA, attacker disrupts optimal power flow decisions after the system reaches a steady state condition. It affects the demand response and demand side management. While LAA denotes the direct manipulation of the load at end-user devices installed at customer side, it directly controls and changes the unsafe controllable loads to overload the circuit and cause deviation of power system frequency from its nominal value. Load altering attack in power system occurs in static or dynamic forms. The static LAA denotes the cyber-attack that changes the load amount suddenly. In contrast, the dynamic LAA changes the damaged load and runs the trajectory under the damaged load.

Recently some schemes for dynamic LAA detection have been proposed in [25-32]. In [25], dynamic LAA detection and localization strategy was proposed for power system described as a linear discrete-time system and then a bank of unknown input observers was designed for attack localization. In [26], a bank of functional observers was proposed for LAA detection in smart grids. The attack localization scheme based on the bank of observers idea designs one observer, one residual generator and one residual evaluation unit for each bus possible to attack for isolating the under-attack bus from the others. In [27, 28], SMO-based dynamic LAA detection strategy was presented for power systems. Compared with [28], in [27] attack reconstruction is done by invoking the SMO. In [29], data driven and model-based methods have been invoked to develop an attack detection and localization scheme for smart grid. In [30], an adaptive sliding mode controller-based consensus algorithm was proposed for fractional-order linear time invariant cyber physical power systems under the cyber-attacks, external disturbance and modeling uncertainties. It mitigates attack effects in a passive manner and does not focus on attack detection, isolation and estimation purpose. In [31, 32], discrete model of the smart grid in the presence

TABALE I Comparison review between the proposed scheme and existing works

Ref. No.	Power system model	Proposed scheme	Attack estimation	Attack localization	Limitations
[15]	Continuous-time model	Sliding mode observer	No	No	• Requires off-line learning phase to tune observer parameters • Requires training data.
[19]	Continuous-time model	No	Yes	No	• Attack occurrence time and attack location are not determined.
[25]	Discrete-time model	Unknown input observer	No	Yes	• Unable to estimate attack. • Complex attack localization and facing scalability problem due to designing a bank of observers. • Only checking vulnerable buses.
[26]	Continuous-time model	Functional Observer	No	Yes	• Unable to estimate detected attack. • Facing scalability issue for attack localization in large scale systems.
[27]	Continuous-time model	Sliding mode observer	Yes	No	• Unable to identify attack location. • Unable to reconstruct complicated attacks.
[28]	Continuous-time model	Sliding mode observer	No	No	• Unable to determine attack location. • Unable to reconstruct attack signal. • All state variables are required to be measurable.
[29]	Discrete-time model	Unknown input observer	No	Yes	• Requires off-line learning to train neural networks. • Complexity due to the combination of model-based and data-based approaches.
[31, 32]	Discrete-time model	Adaptive fading Kalman filter	No	No	• Unable to identify attack location. • Unable to reconstruct attack signal.
Proposed scheme	Continuous-time model	Adaptive observer	Yes	Yes	• Dealing with dynamic LAA

of dynamic LAA was used to design an attack detection scheme based on the adaptive fading Kalman filter.

In Table 1 comprehensive review of the related existing works has been presented. From Table 1, (1) The existing schemes investigate attack detection and localization problem based on the discrete-time model of the power systems which has larger dimension than the continuous one; (2) most of the existing works that consider the continuous-time model of the power systems, invoke SMO for attack detection and none of them can simultaneously detect attack occurrence time, estimate attack severity and identify attack location; (3) Majority of the existing schemes design a bank of observers for diagnosing attack location, which might face scalability issues as the system size increases; (4) Almost all of the existing attack localization schemes only check a set of vulnerable bus status, so they fail to identify attack location when the attacked bus falls outside the assumed vulnerable set. (5) None of the existing works estimate attack severity and magnitude during real time operation of the system which could be used to design defense strategies for attack mitigation.

Motivated by the above discussion, this paper proposes an observer-based FDIA detection and estimation scheme for smart grid. The proposed scheme uses the continuous-time mathematical model of the smart grid derived by invoking DC power flow equations, generator swing equation and governor controller and load-frequency controller. To estimate the gird states, SMO is designed and output estimation error is considered as the residual signal and used for monitoring the grid status. By evaluating the residual signal and comparing it with appropriate threshold, attack detection is done and alarm signal is generated. Upon generating detection alarm, the proposed adaptive attack estimation algorithm is activated to estimate the FDIA affected the victim loads at the vulnerable buses. Stability analysis of the proposed state observer and the proposed attack estimation algorithm guarantees that all signals of the closed-loop system are uniformly ultimately bounded.

Main contributions of this work are: (1) Compared with the discrete-time model based schemes, the proposed strategy is based on the continuous-time model of the power systems, which has lower dimension than the discrete one and therefore it has simpler structure with less computational complexity. (2) The proposed attack estimation algorithm does not require any prior information about attack signal such as occurrence time, attack severity and attack location. (3) The presented approach only requires the frequency deviations of the generator buses, and therefore all grid states are not required to be available. (4) The proposed attack estimation algorithm estimates the severity and time characteristic of the occurred attack without requiring any off-line learning phase and any training data set.

II. Problem Statement

A. System Model

Consider a smart grid with $N = G \cup L$ buses where G and L represent the sets of generator and load buses, respectively. Considering the DC power flow equations of the generator buses, the model of the turbine governor and the load frequency controller, the following descriptor model of the power system is obtained [24]:

$$\begin{bmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\omega}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ K^i - L_{gg} & K^p - D & -L_{gl} \\ -L_{lg} & 0 & -L_{ll} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\omega} \\ \boldsymbol{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{P_L} \end{bmatrix}. \quad (1)$$

Table 2 describes all variables presented in (1) where $L = \begin{bmatrix} L_{gg} & L_{gl} \\ L_{lg} & L_{ll} \end{bmatrix}$, and M, D, K^i and K^p are positive diagonal matrices.

TABLE II System parameters

Physical Parameters	Definition
$\boldsymbol{\delta}$(rad)	Voltage phase angle at the generator buses
$\boldsymbol{\omega}$(p.u.)	Rotor angular frequency deviations at the generator buses
$\boldsymbol{\theta}$(rad)	Voltage angles at the load buses
M(p.u.)	Inertia coefficient of the generators
D(p.u.)	Damping coefficient of the generators
L(p.u.)	Imaginary part of the transmission line admittance matrix
K^i	Integral controller coefficient
K^p	Proportional controller coefficient
$\boldsymbol{P_L}$	Load power demand

Form (1), vector of voltage phase angle is obtained as

$$\boldsymbol{\theta} = -L_{ll}^{-1}L_{lg}\boldsymbol{\delta} + L_{ll}^{-1}\boldsymbol{P_L}. \quad (2)$$

By substituting (2) in (1), and doing some manipulation on the result, the following mathematical model is achieved

$$\begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} 0 \\ -M^{-1}L_{gl}L_{ll}^{-1} \end{bmatrix}\boldsymbol{P_L} + \begin{bmatrix} 0 & I \\ M^{-1}\left(K^i - L_{gg} + L_{gl}L_{ll}^{-1}L_{lg}\right) & M^{-1}(K^p - D) \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\omega} \end{bmatrix}. \quad (3$$

Now, the state vector is defined as $\boldsymbol{x} \triangleq [\boldsymbol{\delta} \quad \boldsymbol{\omega}]^T$ and the state space model of the power system is derived as

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{P_L}, \\ \boldsymbol{y} &= C\boldsymbol{x}, \end{aligned} \quad (4)$$

where $\boldsymbol{y}$ is the measured output vector, $A = \begin{bmatrix} 0 & I \\ M^{-1}\left(K^i - L_{gg} + L_{gl}L_{ll}^{-1}L_{lg}\right) & M^{-1}(K^p - D) \end{bmatrix}$ is the system matrix, $B = \begin{bmatrix} 0 \\ M^{-1}L_{gl}L_{ll}^{-1} \end{bmatrix}$ is the input matrix and C is the output matrix.

B. Attack Model

As explained in [23], dynamic LAA is a kind of FDIA that targets load changes at vulnerable buses and it may occur in the open loop and closed-loop forms. In the open loop form of the dynamic LAA, attacker does not access to the tools to monitor the grid condition and only uses historical data to apply the attack trajectory to the victim load. While in the closed-loop form, the attacker locates frequency sensor at some particular buses to continuously monitor the grid condition for taking feedback from the current status of the power grid and designing the attack trajectory by using some simple feedback controllers like proportional one. Now, assume that the load at bus is the victim load. The closed-loop dynamic LAA designed by a proportional controller is modeled as:

$$f_v = -K_{v,s}\omega_s + \varepsilon_v, \tag{5}$$

where $K_{v,s} > 0$ is the proportional controller gain, ω_s measured by the attacker's sensor at bus s, denotes the frequency deviation from its nominal value, and ε_v represents the small load fluctuations in power consumption on the demand side. So, the load power at bus v can be expressed as

$$P_{L,v} = P_{L,s}^{v} - K_{v,s}\omega_s + \varepsilon_v, \tag{6}$$

where $P_{L,s}^{v}$ is the secure part of the load power at bus v. Thereby, the power system model in (4) under dynamic LAA can be written as

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{P}_{L,s} + E\boldsymbol{f}, \\ \boldsymbol{y} &= C\boldsymbol{x}, \end{aligned} \tag{7}$$

where $\boldsymbol{x} \in R^n$ is the state vector, $\boldsymbol{P}_{L,s} \in R^{m\times 1}$ is the secure load power, $E \in R^{n\times p}$ is an input matrix corresponding to the grid buses possible to attack, and $E \subset B$, $\boldsymbol{f} \in R^{p\times 1}$ represents the dynamic LAA which is unknown.

Before presenting the proposed approach, the following assumptions are considered.

Assumption 1. It is assumed that: (i) The pain (A, C) is observable, (ii) $rank(CE) = p$, and (iii) Invariant zeros of (A, E, C) lie in open left half plane.

Assumption 2. The attack function $\boldsymbol{f}$ is assumed to be bounded, i.e., $\|\boldsymbol{f}\| \le \bar{f}$.

Assumption 3. The parameters of all vulnerable buses E in (7) are assumed to be known.

Remark 1- The norm of the attack vector in assumption 2 is an unknown constant. It is only used for analytical purposes and implementing the proposed algorithm does not require its numerical value.

III. Proposed Attack Detection and Estimation Scheme

In this section, at first attack detection scheme is presented and then the proposed attack estimation algorithm is explained.

A. Proposed Attack Detection Scheme

In this work, the following SMO is proposed to estimate state variables:

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}} &= A\hat{\boldsymbol{x}} + B\boldsymbol{P}_{L,s} + \boldsymbol{\nu} + \chi E\hat{\boldsymbol{f}} - L(\hat{\boldsymbol{y}} - \boldsymbol{y}), \\ \hat{\boldsymbol{y}} &= C\hat{\boldsymbol{x}}, \end{aligned} \tag{8}$$

where $\hat{\boldsymbol{x}} \in R^n$ denote the estimated state, $\hat{\boldsymbol{y}} \in R^q$ is estimated output, $L \in R^{n\times q}$ is the observer gain matrix, $\hat{\boldsymbol{f}} \in R^p$ denotes the attack estimation which is zero before any attack detection and $\boldsymbol{\nu} \in R^p$ is defined as

$$\boldsymbol{\nu} = -\rho\frac{F_1\tilde{\boldsymbol{y}}}{\|F_1\tilde{\boldsymbol{y}}\| + \delta}, \tag{9}$$

where positive constants δ and ρ are design parameters, $\tilde{\boldsymbol{y}}$ is the output estimation error defined as $\tilde{\boldsymbol{y}} = \hat{\boldsymbol{y}} - \boldsymbol{y}$, and $F_1 = C^T P$ where P will be defined in theorem 1. Let us defined the state estimation error as:

$$\tilde{\boldsymbol{x}} = \hat{\boldsymbol{x}} - \boldsymbol{x}. \tag{10}$$

For attack detection, output estimation error $\tilde{\boldsymbol{y}}$ is considered as a residual signal. Before attack occurrence, we have $\boldsymbol{f} = 0$ and $\hat{\boldsymbol{f}} = 0$. Considering (7) and (8), state estimation error dynamics before attack is obtained as:

$$\dot{\tilde{\boldsymbol{x}}} = (A - LC)\tilde{\boldsymbol{x}} + \boldsymbol{\nu}. \tag{11}$$

Substituting (9) in (11), gives:

$$\dot{\tilde{\boldsymbol{x}}} = (A - LC)\tilde{\boldsymbol{x}} - \rho\frac{F_1\tilde{\boldsymbol{y}}}{\|F_1\tilde{\boldsymbol{y}}\| + \delta}. \tag{12}$$

Theorem 1. Consider the power system mode in (7), and the proposed SMO in (8). Under assumption 1, if for a given positive definite matrix $Q \in R^{n\times n}$, there exists a symmetric positive definite matrix $P \in R^{q\times q}$ such that the following matrix inequality holds, then the proposed SMO guarantees that all signals of the closed-loop system are bounded and estimation error converges to zero asymptotically,

$$(A - LC)^T C^T PC + C^T PC(A - LC) \le -Q. \tag{13}$$

Proof. Consider the Lyapunov function as follows

$$V_1 = \tilde{\boldsymbol{x}}^T C^T PC\tilde{\boldsymbol{x}}. \tag{14}$$

Differentiating (14) with respect to time and using (12), results in

$$\dot{V}_1 = \tilde{\boldsymbol{x}}^T\big((A - LC)^T C^T PC + C^T PC(A - LC)\big)\tilde{\boldsymbol{x}} \tag{15}$$

$$-2\rho\frac{\|F_1\tilde{\boldsymbol{y}}\|^2}{\|F_1\tilde{\boldsymbol{y}}\|+\delta} \le -\tilde{\boldsymbol{x}}^T Q\tilde{\boldsymbol{x}}.$$

From (15), it is obtained that $\tilde{\boldsymbol{x}}$ is bounded and therefore $\tilde{\boldsymbol{y}}$ is bounded. This together with (9) implies that $\dot{\tilde{\boldsymbol{x}}}$ is bounded. Applying the Barbalat's Lemma to these results leads to the conclusion that $\tilde{\boldsymbol{x}} \to 0$ as $t \to \infty$ and subsequently $\tilde{\boldsymbol{y}} \to 0$ as $t \to \infty$.

In the following, for attack detection, output estimation error $\tilde{\boldsymbol{y}}$ is considered as the residual signal and the following residual evaluation function is proposed

$$J(\tilde{\boldsymbol{y}}) = \sqrt{\frac{1}{T}\int_0^T \tilde{\boldsymbol{y}}^T(\tau)\,\tilde{\boldsymbol{y}}(\tau)d\tau}. \tag{16}$$

By integrating both side of (15), we will have

$$\begin{aligned}&\underline{\lambda}_Q \int_0^t \tilde{\boldsymbol{x}}^T(\tau)\,\tilde{\boldsymbol{x}}(\tau)d\tau \le \int_0^t \tilde{\boldsymbol{x}}^T(\tau)\,Q\tilde{\boldsymbol{x}}(\tau)d\tau \\ &\le V_1\big(\tilde{\boldsymbol{y}}(0)\big) - V_1\big(\tilde{\boldsymbol{y}}(t)\big).\end{aligned} \tag{17}$$

Inequality (17) proves that $\int_0^t \tilde{\boldsymbol{x}}^T(\tau)\,\tilde{\boldsymbol{x}}(\tau)d\tau$ is bounded and therefore there exists positive constant U such that $C^TC \le UI$, where I is the identity matrix with appropriate dimension, and therefore one can easily obtain:

$$\begin{aligned}&\int_0^t \tilde{\boldsymbol{y}}^T(\tau)\,\tilde{\boldsymbol{y}}(\tau)d\tau = \int_0^t \tilde{\boldsymbol{x}}^T(\tau)\,C^TC\tilde{\boldsymbol{x}}(\tau)d\tau \\ &\le U\int_0^t \tilde{\boldsymbol{x}}^T(\tau)\,\tilde{\boldsymbol{x}}(\tau)d\tau.\end{aligned} \tag{18}$$

Inequality (18) implies that $\int_0^t \tilde{\boldsymbol{y}}^T(\tau)\,\tilde{\boldsymbol{y}}(\tau)d\tau$ is bounded. Therefore evaluation function $J(\tilde{\boldsymbol{y}})$ in (16) is bounded before attack occurrence. If the maximum bound of the evaluation function under secure operation of power system is defined as J_{th}, then decision making for attack detection can be done based on the following logic:

$$\chi = \begin{cases}0 : J \le J_{th}\\ 1 : J > J_{th}\end{cases} \tag{19}$$

where χ is called detection indicator. In (19), evaluation function is compared with the threshold level to decide whether or not cyber-attack has occurred. This step determines the time at which the vulnerable buses are subjected to some attacks. As obtained from (19), once the evaluation function exceeds from the threshold level, χ is triggered from zero to one and announces the attack occurrence and simultaneously activates attack estimator to estimate the occurred attack and to identify the under attack bus among the vulnerable buses.

B. Proposed Attack Estimator

Once the attack is detected and attack indicator triggered from zero to one, attack estimator is activated and therefore, dynamics of the estimator error is obtained as

$$\dot{\tilde{\boldsymbol{x}}} = (A-LC)\tilde{\boldsymbol{x}} + \boldsymbol{\nu} + E\tilde{\boldsymbol{f}}. \tag{20}$$

where $\tilde{\boldsymbol{f}} = \hat{\boldsymbol{f}} - \boldsymbol{f}$ is the attack estimation error. Substituting (9) in (20) results in

$$\dot{\tilde{\boldsymbol{x}}} = (A-LC)\tilde{\boldsymbol{x}} - \rho\frac{F_1\tilde{\boldsymbol{y}}}{\|F_1\tilde{\boldsymbol{y}}\|+\delta} + E\tilde{\boldsymbol{f}}. \tag{21}$$

Theorem 2. Consider the power system model in (7) and adaptive SMO in (8). After attack detection, under assumptions 1 and 2, if there exists symmetric positive definite matrix $P \in R^{q\times q}$ such that the following matrix inequality holds

$$\Xi = \begin{bmatrix}-\big((A-LC)^TC^TPC + C^TPC(A-LC)\big) & 0\\ 0 & -\boldsymbol{\sigma}\end{bmatrix} < 0. \tag{22}$$

then the proposed attack estimation algorithm

$$\dot{\hat{\boldsymbol{f}}} = \boldsymbol{\Gamma}\big(-E^TC^TP\tilde{\boldsymbol{y}} - \sigma\hat{\boldsymbol{f}}\big). \tag{23}$$

guarantees that $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{f}}$ are uniformly ultimately bounded where symmetric positive definite matrix $\boldsymbol{\Gamma} \in R^{p\times p}$ is the learning rate, and σ is a design parameter chosen by the designer.

Proof. Consider the following Lyapunov function:

$$V_2 = \tilde{\boldsymbol{x}}^TC^TPC\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{f}}^T\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{f}}. \tag{24}$$

Differentiating (24) with respect to time and substituting (21) in the result, gives

$$\begin{aligned}\dot{V}_2 &= \tilde{\boldsymbol{x}}^T\big((A-LC)^TC^TPC + C^TPC(A-LC)\big)\tilde{\boldsymbol{x}} \\ &+2\tilde{\boldsymbol{x}}^TC^TPC\boldsymbol{\nu} + 2\tilde{\boldsymbol{f}}^TE^TC^TPC\tilde{\boldsymbol{x}} + 2\tilde{\boldsymbol{f}}^T\boldsymbol{\Gamma}^{-1}\dot{\hat{\boldsymbol{f}}}.\end{aligned} \tag{25}$$

Substituting (9) into (25) gives:

$$\begin{aligned}\dot{V}_2 &= \tilde{\boldsymbol{x}}^T\big((A-LC)^TC^TPC + C^TPC(A-LC)\big)\tilde{\boldsymbol{x}} \\ &-2\rho\tilde{\boldsymbol{y}}^TF^T\frac{F\tilde{\boldsymbol{y}}}{\|F\tilde{\boldsymbol{y}}\|+\delta} + 2\tilde{\boldsymbol{f}}^T\Big(E^TC^TP\tilde{\boldsymbol{y}} + \boldsymbol{\Gamma}^{-1}\dot{\hat{\boldsymbol{f}}}\Big).\end{aligned} \tag{26}$$

Substituting the attack estimation algorithm in (23) into (26), gives

$$\dot{V}_2 \le -\tilde{\boldsymbol{x}}^TQ\tilde{\boldsymbol{x}} - 2\rho\frac{\|F\tilde{\boldsymbol{y}}\|^2}{\|F\tilde{\boldsymbol{y}}\|+\delta} - 2\boldsymbol{\sigma}\tilde{\boldsymbol{f}}^T\hat{\boldsymbol{f}}. \tag{27}$$

Substituting $\tilde{\boldsymbol{f}} = \hat{\boldsymbol{f}} - \boldsymbol{f}$ in (26) and applying inequality $-2\tilde{\boldsymbol{f}}^T\boldsymbol{f} \le \tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{f}} + \boldsymbol{f}^T\boldsymbol{f}$ to the result, gives

$$\begin{aligned}\dot{V}_2 &\le -\tilde{\boldsymbol{x}}^TQ\tilde{\boldsymbol{x}} - 2\rho\frac{\|F\tilde{\boldsymbol{y}}\|^2}{\|F\tilde{\boldsymbol{y}}\|+\delta} - \sigma\tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{f}} + \sigma\boldsymbol{f}^T\boldsymbol{f} \\ &\le \boldsymbol{\psi}^T\Xi\boldsymbol{\psi} + \sigma\boldsymbol{f}^T.\end{aligned} \tag{28}$$

where $\boldsymbol{\psi} = [\tilde{\boldsymbol{x}} \quad \tilde{\boldsymbol{f}}]^T$. From (28), it is obtained that for $\Xi < 0$ we have $\dot{V}_2(t) < -\underline{\lambda}(-\Xi)\|\boldsymbol{\psi}(t)\| + \sigma\boldsymbol{f}^T\boldsymbol{f}$, where $\underline{\lambda}$ denotes

the minimum eigenvalue of matrix $-\Xi$. It follows that $\dot{V}_2(t) < 0$ for $\underline{\lambda}(-\Xi)\|\boldsymbol{\psi}(t)\| > \sigma \boldsymbol{f}^T \boldsymbol{f}$, which shows that state estimation error and attack estimation errors, i.e., $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{f}}$, are uniformly bounded and they converge to a small set near the origin.

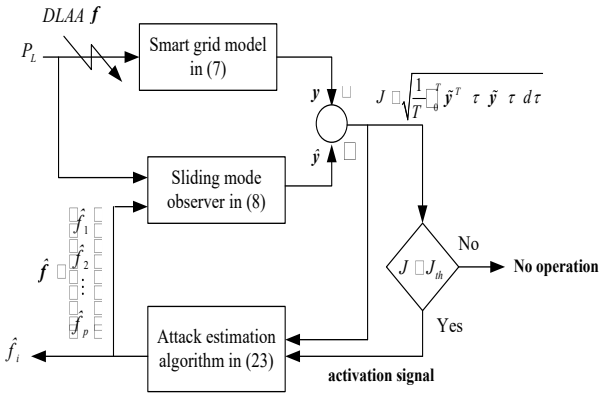

Fig. 1. Schematic of the proposed attack detection and estimation scheme.

Remark 2. The proposed attack estimation algorithm in theorem 2 identifies the attack and estimates its magnitude online. It estimates severity of load changes in all vulnerable buses possible to attack. The non-zero element of attack estimator output corresponds to the load changes at under-attack bus. This step isolates the under-attack bus from the secure ones.

Figure 1 shows the schematic of the proposed attack detection and estimation scheme.

IV. Simulation Results

In this section, performance of the proposed scheme is verified on the IEEE-6 bus grid network under cyber-attack. For simulation, system parameters set to $M = diag\{0.125, 0.034, 0.016\}$, $D = diag\{0.125, 0.068, 0.032\}$, and the PI controller coefficients are set to $K_i = diag\{-35, -40, -35\}$ and $K_p = diag\{-2, -9, -3\}$. The transmission line admittance is chooses as the same as that in [25], initial conditions of the system is set to $\boldsymbol{x}(0) = [-0.05 \quad -0.045 \quad -0.07 \quad 0 \quad 0.1 \quad 0.01]^T$ and initial conditions of the observer is set to zero.

For evaluating effectiveness of the proposed scheme, three different attacks at three different instants have been considered. In case 1, a single-point closed-loop dynamic LAA is considered and in case 2-1 and case 2-2, two different multi-point closed-loop dynamic LAAs are applied to the system. In the single-point dynamic LAA, attacker manipulates the load consumption only at one victim load bus while in the multi-point dynamic LAAs, attacker manipulates a group of vulnerable loads at several load buses.

Figure 2 shows the considered attack cases applied to the IEEE 6-bus power system.

Case 1: Single-point closed-loop dynamic LAA- In this case, a single-point closed-loop attack at victim bus $v = 4$ with sensor bus $s = 1$ is applied at $t = 4S$. The proportional control gain and load fluctuations for attack signal design set to $K_{4,1} = 7$ and $\varepsilon_4 = 0.4$, respectively. Figures 3-5 show the simulation results. Figure 3-(a) shows

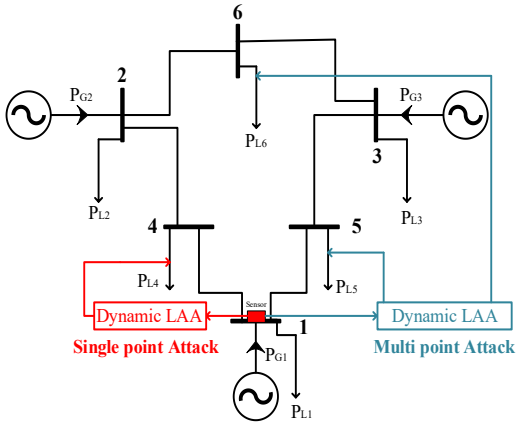

Fig. 2. IEEE 6-bus power system under single-point and multi-point dynamic LAAs.

the victim load variation at the vulnerable bus 4 and Fig. 3-(b) illustrates the system frequency deviation from its nominal value.

As obtained from the results, while the attack is occurred, grid frequency deviates from its nominal value. In Fig. 4-(a), evaluation function and threshold level are given. As obtained from the result, once the residual signal exceeds the threshold level, detection indicator is activated to announce the attack occurrence (Fig. 4-(b)). As obtained from Fig. 4-(b), attack is detected after 2mS. Upon attack detection, attack estimation algorithm is activated to estimate the load changes at the vulnerable buses 4, 5, and 6, simultaneously. Figure 5 shows the output of the attack estimation algorithm and actual attack occurred at the vulnerable buses. Results verify that the proposed algorithm estimates the occurred attack accurately. Output of the attack estimation algorithm shows that cyber-attack has occurred at the vulnerable bus 4 and other buses are free from attack.

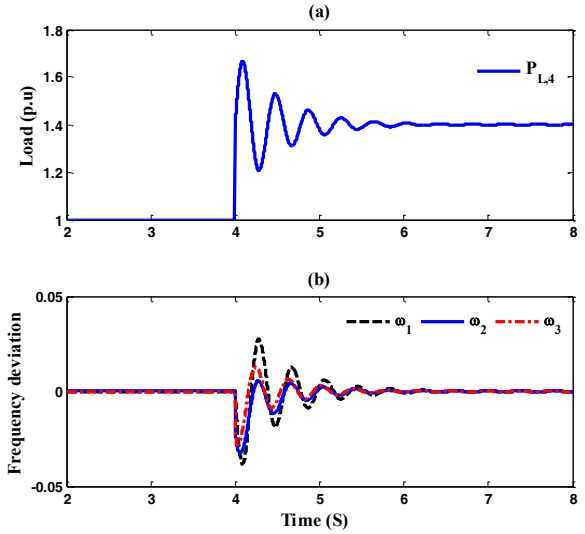

Fig. 3. (a) The victim load change, and (b) Frequency deviation (case 1).

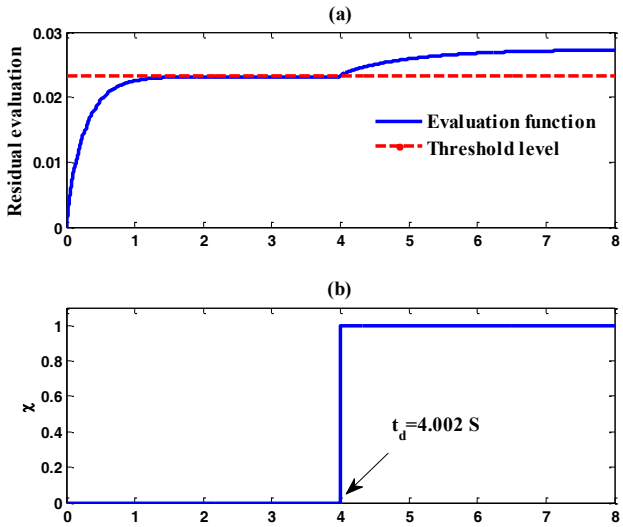

Fig. 4. (a) Evaluation function and threshold level, and (b) Detection indicator (case 1).

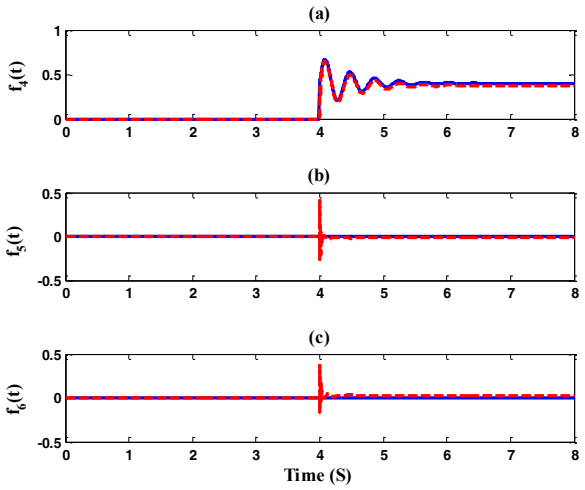

Fig. 5. Output of the attack estimation scheme and actual attack corresponding to (a) Bus 4, (b) Bus 5, and (c) Bus 6 (actual attack: solid line, and estimated attack: dashed line) (case 1).

Case 2-1: Multi-point closed-loop dynamic LAA at $t = 3S$- In this case, a multi-point closed-loop attack at vulnerable buses $v = 5,6$ with sensor bus $s = 1$ is applied. By using the proportional controller, the attack gain $K_{5,1} = K_{6,1} = 6$ and load fluctuations $\varepsilon_5 = 0.5$ and $\varepsilon_6 = 0.3$ are designed and applied to the victim load buses.

Figures 6-8 show the results. Figure 6-(a) shows the load changes at vulnerable buses 5 and 6 and Fig. 6-(b) illustrates the frequency deviation of the grid.

In Fig. 7-(a) evaluation function is compared with the threshold level. As obtained from Fig. 7-(b), at $t_d = 3.001S$ evaluation function exceeds from the threshold and alarm signal is activated and simultaneously attack estimation algorithm is activated. Figure 8 shows the output of the attack estimation algorithm and load variations at vulnerable buses. Output of the attack estimation algorithm

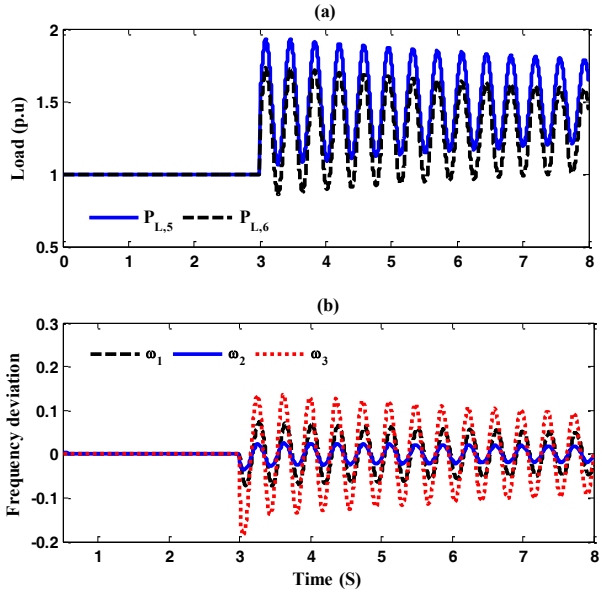

Fig. 6. (a) The victim load change, and (b) Frequency deviation (case 2-1).

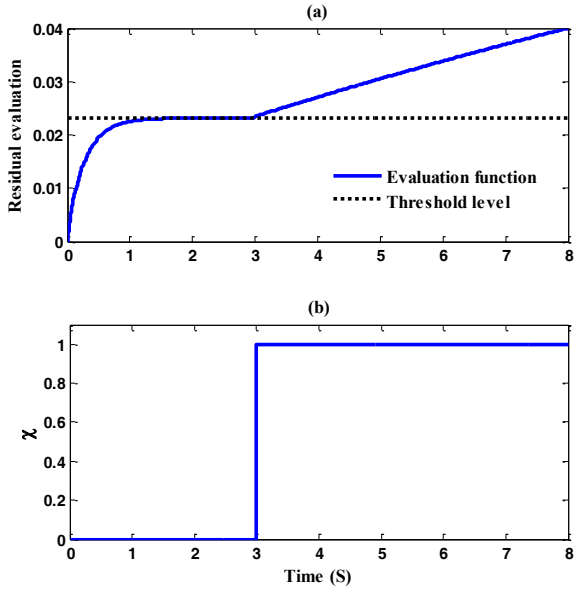

Fig. 7. (a) Evaluation function and threshold level, and (b) Detection indicator (case 2-1).

reveals that buses 5 and 6 are under load variations and bus 4 is free from cyber-attack.

Case 2-2: Multi-point closed-loop dynamic LAA at $t = 2S$ - In this case, performance of the proposed scheme against another multi-point closed-loop attack at vulnerable buses $v = 5, 6$ at $t = 2S$ is evaluated. The attack gains are $K_{5,1} = 5$ and $K_{6,1} = 7$, and load fluctuations are $\varepsilon_5 = 0.6$ and $\varepsilon_6 = 0.5$. Figure 9-(a) shows the load changes at vulnerable buses and Fig. 9- (b) shows the frequency deviations under the applied attack. Figure 10 illustrates the evaluation function and threshold level. It shows that upon attack occurrence at $t = 2S$, the evaluation function

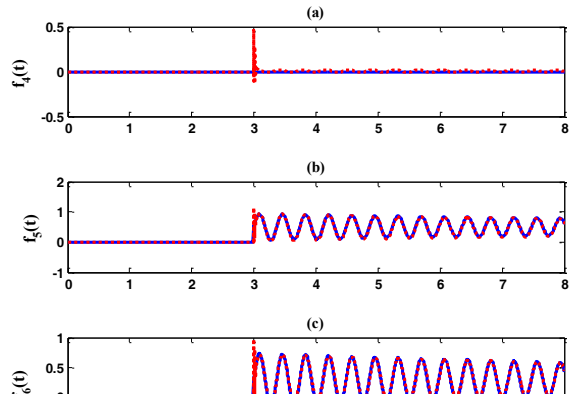

Fig. 8. Output of the attack estimation scheme and actual attack corresponding to (a) Bus 4, (b) Bus 5, and (c) Bus 6 (actual attack: solid line, and estimated attack: dotted line) (case 2-1)

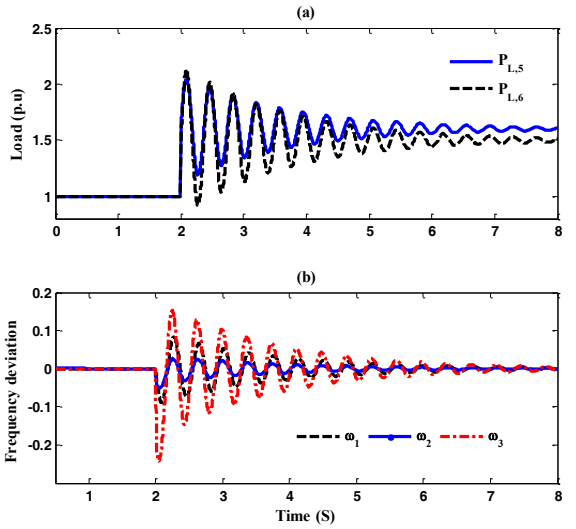

Fig. 9. (a) The victim load change, and (b) Frequency deviation (case 2-2).

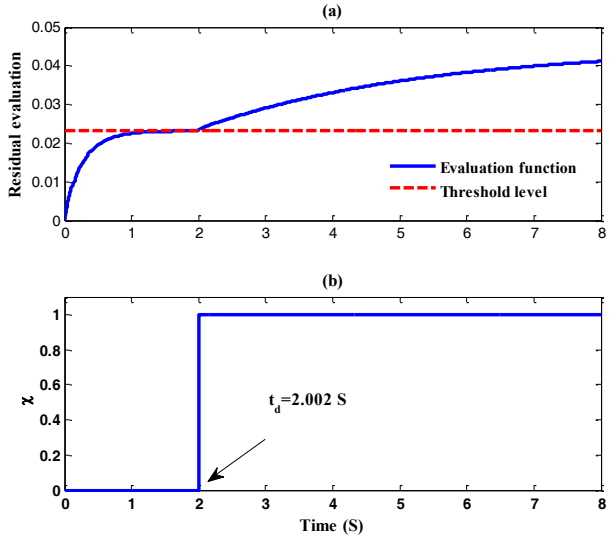

Fig. 10. Evaluation function and threshold level, and (b) Detection indicator (case 2-2).

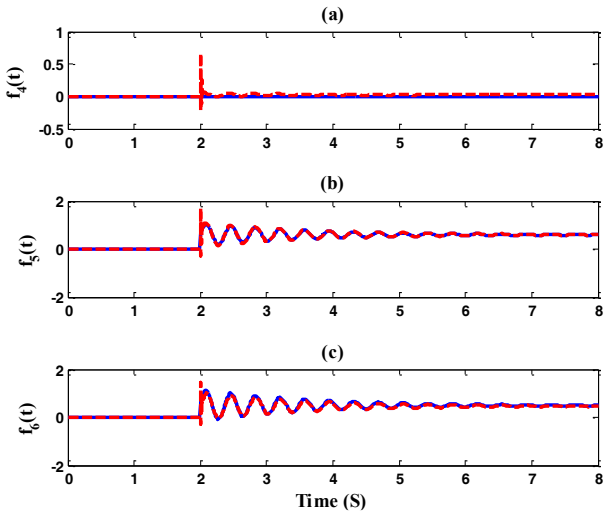

Fig. 11. Output of the attack estimation scheme and actual attack corresponding to (a) Bus 4, (b) Bus 5, and (c) Bus 6 (actual attack: solid line, and estimated attack: dotted line) (case 2-2)

exceeds from the threshold level and detects attack. As obtained from Fig. 10-(b) attack is detected at $t = 2.002S$.

Finally, Fig. 11 shows the actual attack and estimated attack at the vulnerable buses 4, 5, and 6, simultaneously.

As obtained from the result, upon the detection indicator announces the attack occurrence, the estimation unit is activated to estimate the detected attack online. Obtained results in Fig. 11 verify that the proposed attack estimator presents accurate estimation of the occurred attacks and simultaneously determines the under-attack buses.

In order to quantitatively evaluate performance of the proposed attack estimation algorithm, mean square error (MSE) of attack estimation for considered scenarios is calculated and reported in Table 3. Reported results confirm that the proposed estimation algorithm is able to estimate attack signals with acceptable precision.

Obtained simulation results confirm the ability of the proposed scheme to detect and estimate the dynamic load altering attack online at any arbitrary time instant.

V. Conclusions

In this paper, an adaptive observer-based attack detection and estimation algorithm was proposed for smart grid under cyber-attacks. The proposed scheme focused on detecting and estimating dynamic LAA as an important type of FDIA in smart grids. In the proposed scheme, output estimation error obtained from the SMO is defined as the residual signal and used for monitoring status of the smart grid for attack detection. By evaluating residual signal and comparing it with appropriate threshold level, attack detection is done. Once the evaluation function exceeds the threshold, detection indicator announces attack occurrence and simultaneously activates attack estimation algorithm. Attack estimation algorithm estimates the load variations in all vulnerable buses based on the adaptive law developed by using the Lyapunov direct method. Output of the attack estimation algorithm estimates the severity of the detected attack and identifies which vulnerable bus is under attack. Simulation results on the IEEE 6-bus system under three different DLAAs verify effectiveness of the proposed scheme.

TABLE III MSE of the estimated attacks

Physical Parameters	Case 1	Case 2-1		Case 2-2	
Attack Signal	$f_4(t)$	$f_5(t)$	$f_6(t)$	$f_5(t)$	$f_6(t)$
MSE of the estimated attack	4.63e-4	9.14e-4	6.5e-4	9.62e-4	2e-3

References

[1] X. Yu, and Y. Xue, "Smart grids: A cyber–physical systems perspective," *Proceedings of the IEEE*, vol. 104, no. 5, p. 1058-1070, May 2016.

[2] G. Dileep, "A survey on smart grid technologies and applications," *Renewable Energy*, vol. 146, p. 2589–2625, Feb. 2020.

[3] Y. Kim, S. Hakak, and A. Ghorbani, "Smart grid security: attacks and defence techniques," *IET Smart Grid*, vol. 6, no. 2, p. 103-123, 2023.

[4] S. Jin, "False data injection attack against smart power grid based on incomplete network information," *Electric Power Systems Research*, vol. 230, May 2024, 110294.

[5] Z. Qu, J. Yang, Y. Wang, and P. M. Georgievitch, "Detection of false data injection attack in power system based on hellinger distance," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 2, p. 2119-2128, Feb. 2024.

[6] M. Shahriari, S.H. Nourian, "Adaptive resilient control of uncertain nonlinear cyber-physical systems under deception attack," International Journal of Industrial Electronics, Control and Optimization, in press, Online Oct. 2025, doi. 10.22111/ieco.2025.52052.1691.

[7] W. Fu, Y. Yan, Y. Chen, Z. Wang, D. Zhu, and L. Jin, "Temporal false data injection attack and detection on cyber-physical power system based on deep reinforcement learning," *IET Smart Grid*, vol. 7, no. 1, p. 78-88, Feb. 2024.

[8] V.S. Rajkumar, A. Ştefanov, A. Presekal, P. Palensky, and J.L. Rueda Torres, "Cyber-attacks on power grids: Causes and propagation of cascading failures," *IEEE Access*, vol. 11, p. 103154–103176, 2023.

[9] V.S. Rajkumar, A. Ştefanov, J.L.R. Torres, and P. Palensky, "Dynamical analysis of power system cascading failures caused by cyber-attacks," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 6, p. 8807–8817, 2024.

[10] D. Du, M. Zhu, X. Li, M. Fei, S. Bu, L. Wu, and K. Li, "A review on cybersecurity analysis, attack detection, and attack defense methods in cyber-physical power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 3, p. 727-743, May 2023.

[11] S. Sahoo, F. Blaabjerg, and T. Dragicevic, Cyber security for microgrids, *Published by The Institution of Engineering and Technology (IET)*, London, United Kingdom, 2022.

[12] R.V. Yohanandhan, R.M. Elavarasan, P. Manoharan, and L. Mihet-Popa, "Cyber-physical power system (CPPS): A reviews on modeling, simulation, and analysis with cyber security applications," *IEEE Access*, vol. 8, 2020.

[13] A. Izadbakhsh, A. Deylami, and S. Khorashadizadeh, "Observer-based versus non-observer based adaptive control of electrically driven cooperative manipulators using q-analogue of the Bernstein-Schurer-Stancu operator as uncertainty approximator," *International Journal of Control, Automation, and Systems*, vol. 21, no. 8, pp. 2664-2673, 2023.

[14] X. Wang, X. Luo, M. Zhang, Z. Jiang, and X. Guan, "Detection and isolation of false data injection attacks in smart grids via unknown input interval observer," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3214-3229, 2020.

[15] M. Adeli, M. Hajatipour, M.J. Yazdanpanah, H. Hashemi-Dezaki, and M. Shafieirad, "Optimizaed cyber-attack detection method of power systems using sliding mode observer," *Electric Power Systems Research*, vol. 205, 107745, 2022.

[16] W. Ao, Y. Song, and C. Wen, "Adaptive cyber-physical system attack detection and reconstruction with application to power systems," *IET Control Theory & Applications*, vol. 10, no. 12, p. 1458-1468, 2016.

[17] X. Luo, Q. Yao, X. Wang, and X. Guan, "Observer-based cyber attack detection and isolation in smart grids," *Electrical Power and Energy Systems*, vol. 101, p. 127-138, 2018.
[18] G. Rinaldi, P.P. Menon, C. Edwards, A. Ferrara, and Y. Shtessel, "Adaptive dual-layer super-twisting sliding mode observers to reconstruct and mitigate disturbances and communication attacks in power networks," Automatica, vol. 129, 109656, 2021.
[19] J. Li, D. Yang, and Q. Su, "Sliding mode resilient control and application based on intermediate variable observer in smart grid," *International Journal of Control, Automation, and Systems*, vol. 21, no. 6, pp. 1803-1815, 2023.
[20] X. Yu, C. Gao, Y. Du, B. Gao, D. Tian, and T. Hou, "Adaptive residual observer-based detection and isolation framework against false data injection attack in large-scale power systems," *Scientific Reports*, vol. 15, 41070, 2025.
[21] M. Mohammadi, H. Zayyani, M. Bekrani, "RSS localization in the presence of Byzantine attacks using MAP estimation," International Journal of Industrial Electronics, Control and Optimization, in press, online from 03 July 2025, doi. 10.22111/ieco.2025.51214.1666.
[22] P. Verma, and C. Chakraborty, "Load redistribution attacks against smart grids- models, impacts, and defense: A review," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 8, p. 10192-10208, Aug. 2024.
[23] S. Amini, F. Pasqualetti, H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, p. 2862-2872, 2016.
[24] S. Amini, H. Mohseni-Rad, and F. Pasqualetti, Dynamic load altering attacks in smart grid, Proceeding of 2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 18-20 Feb. 2015, Washington, DC, USA.
[25] F. Najafi, S. Nobakht, M. Samimiat, A-K. Ahmadi, and A. Nateghi, "Detection and localization of dynamic load altering attacks in power systems," *Computers and Electrical Engineering*, vol. 123, 110207, 2025.
[26] A.e Sousa, N. Messai, N. Manamanni, "Load-altering attack detection on smart grid using functional observers," *International Journal of Critical Infrastructure Protection*, vol. 37, 100518, 2022.
[27] Q. Su, S. Li, Y. Gao, X. Huang, and J. Li, "Observer-based detection and reconstruction of dynamic load altering attack in smart grid," *Journal of the Franklin Institute*, vol. 358, p. 4013-4027, 2021.
[28] J. Li, H. Li, Q. Su, "Dynamic load altering attack detection for cyber physical power systems via sliding mode observer," *International Journal of Electrical Power and Energy Systems*, vol. 153, 109320, 2023.
[29] X. Wang, X. Wang, X. Luo, X. Guan, S. Wang, "Novel cyber-physical collaborative detection and localization method against dynamic load altering attacks in smart energy grids," *Global Energy Interconnection*, vol. 6, no. 5, p. 362-376, 2023.
[30] M. Nazifi, M. Pourgholi, "Adaptive fractional-order consnensus control of cyber-physical power systems in the presence of unbounded perturbations," *International Journal of Industrial Eectronics, Control and Optimization*, vol. 8, no. 2, 105-116, 2025.
[31] Q. Ma, Z. Xu, W. Wang, L. Lin, T. Ren, S. Yang, and J. Li, "Dynamic load-altering attack detection based on adaptive fading Kalman filter in power systems," *Global Energy Interconnection*, vol. 4, no. 2, p. 184-192, April 2021.
[32] J. Li, C. Sun, S. Yang, and Q. Su, "Dynamic load altering attack detection based on adaptive fading Kalman filter in smart grid," *IET Generation, Transmission & Distribution*, vol. 18, pp. 303-313, 2024.

Zahra Molavi received the B.Sc. degree in Electrical Engineering from Shahrekord University, Iran, in 2013, and the M.Sc. degree in Medical Radiation Engineering from Shahid beheshti university, Iran, in 2018. She is currently pursuing the Ph.D. degree in Electrical Engineering at Shahrekord University, Iran. Her research interests include cybersecurity in power systems and cyber-physical systems.

Abdorreza Rabiee received his B. Sc. degree in 2002 from Shahid Chamran University of Ahwaz and M. Sc. and PhD degrees in 2004 and 2009 from Iran University of Science and Technology (IUST), respectively, all in Electrical Engineering. Currently, he is a professor in the Electrical Engineering Department, Faculty of Engineering and Technology, Shahrekord University (SKU), Shahrekord, Iran. His research interests include power system operation and planning, electricity market, integration of renewable energy into the power system, power system dynamics, and protection.

Maryam Shahriari-kahkeshi received her M.Sc. and Ph.D degrees in control engineering from Isfahan University of Technology, Isfahan, Iran in 2010 and 2014, respectively. She is currently an Associate Professor with the Faculty of Engineering, Shahrekord University, Shahrekord, Iran. Her research interests include artificial intelligence, cyber-physical systems, fault tolerant control systems, and resilient control systems against cyber-attacks.