IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 17, 2021, accepted February 8, 2022, date of publication February 15, 2022, date of current version March 21, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3151907

Cyber-Physical Attack Conduction and Detection in Decentralized Power Systems

MOSTAFA MOHAMMADPOURFARD[1,2], YANG WENG[3], (Senior Member, IEEE), ABDULLAH KHALILI[4], ISTEMIHAN GENC[1], (Member, IEEE), ALIREZA SHEFAEI[5,6], AND BEHNAM MOHAMMADI-IVATLOO[5,7], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Istanbul Technical University, 34469 Istanbul, Turkey
[2]Department of Electrical and Computer Engineering, Sahand University of Technology, Tabriz 5147896, Iran
[3]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281, USA
[4]Department of Electrical and Computer Engineering, University of Hormozgan, Bandar Abbas, Hormozgan 3995, Iran
[5]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666, Iran
[6]Technical University of Delft, 2628 CD Delft, Netherlands
[7]Department of Electrical and Electronics Engineering, Muğla Sıtkı Koçman University, 48000 Muğla, Turkey

Corresponding authors: Mostafa Mohammadpourfard (mohammadpourfard@itu.edu.tr) and Yang Weng (yang.weng@asu.edu)

This work was supported by the TÜBİTAK and European Commission Horizon 2020 Marie Skłodowska-Curie Actions Co-Fund Program under Project 120C080.

ABSTRACT The expansion of power systems over large geographical areas renders centralized processing inefficient. Therefore, the distributed operation is increasingly adopted. This work introduces a new type of attack against distributed state estimation of power systems, which operates on inter-area boundary buses. We show that the developed attack can circumvent existing robust state estimators and the convergence-based detection approaches. Afterward, we carefully design a deep learning-based cyber-anomaly detection mechanism to detect such attacks. Simulations conducted on the IEEE 14-bus system reveal that the developed framework can obtain a very high detection accuracy. Moreover, experimental results indicate that the proposed detector surpasses current machine learning-based detection methods.

INDEX TERMS Deep learning, cyber-attacks, distributed state estimation, smart grids.

I. INTRODUCTION

Expansion of power systems over large geographical areas has made it challenging to implement centralized processing methods [1]. Extending the power system to a wide area requires a complex and extensive network for centralized operation and near real-time processing of the collected measurements [2], [3]. This has accelerated the move towards distributed operation [4]. We specifically focus on distributed state estimation (DSE) in this work. In DSE, the grid is divided into many smaller areas, and each area independently collects measurements from its nodes and estimates the per-area system state. The power grid's overall state is computed by exchanging the per-area system state through an iterative process [5]. Given the growing interest in DSE, understanding its potential vulnerabilities is essential. Cyber-attacks against DSE can cause serious consequences, e.g., cascading failures [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Jahangir Hossain.

A. RELATED WORK

A data integrity attack against the DSE was proposed in [6] wherein the communication infrastructure used to convey information between different areas is compromised. Authors of [7] proposed a data integrity attack on the generation units and the controllable loads. The attacker requires access to sensors, actuators, and unsecured loads in only one area of a distributed network. Common cyber vulnerabilities in distributed power systems were discussed in [8] and the impact of cyber attacks on microgrids was presented. Reference [9] focused on time delay attacks against DSE to destabilize the system.

Several works have been done to detect the cyber-attacks on DSE to protect them against such attacks. [6] proposes a denial of service (DoS) attack detection technique based on the development of mean squared disagreement across areas and a mitigation mechanism based on individual areas' opinions about the attack point. The paper [10] focuses on the distributed state estimation security against DoS attacks. In [11], the distributed resilient filtering challenge for a distributed power system subject to DoS attacks is addressed.

False data injection attack (FDIA) [28] is the most commonly used attack in the cybersecurity of the power systems. In this type of attack, some values are added to the measurement vector so that the bad data detection (BDD) algorithm does not differentiate between normal and manipulated values. A novel DSE approach in networked DC microgrids is presented in [12] to detect the FDIA in the microgrid control network. In the context of DC microgrids, a Kullback-Liebler divergence-based criterion is proposed in [13] to detect the attacks to control the unit of distributed energy resource. A series of papers in this area carried out state estimation and detection together [14]–[18]. In [14], a fully distributed dynamic state estimation algorithm using phasor measurements unit (PMU) data, which is jointly designed with a novel attack detection scheme to limit communication overhead, is presented. The authors of [15] addressed the challenge of joint attack detection and state estimation by using hybrid Bernoulli random set densities to aggregate prior information about signal attacks and system status. The subject of [16] is distributed anomaly identification and reliable estimation of networked cyber-physical systems against FDIAs and jamming attacks. A secure DSE algorithm via consensus-based distributed non-convex optimization protocols is developed in [17]. In [18], an optimal filter and graph theory have been used to develop a distributed state estimation algorithm for a power distribution system integrating several synchronous generators, which have been modeled as a state-space framework. In [19], authors proposed a penalty-based adaptive estimating approach for distributed power systems under FDIAs that can dynamically modify the penalty parameter depending on the area and boundary bus errors.The authors in [20] have used a nonlinear input observer to propose a distributed detection method against FDIAs. Reference [21] compares centralized and decentralized detection methods.

In recent years machine learning has become widely used in all areas, including attack detection algorithms. As a well-known work, the authors in [22] recontextualized the intrusion identification issue as a machine learning issue and analyzed the performance of various learning methods for various attack scenarios. In [23], a detector based on the generalized likelihood ratio is established for cyber-anomalies. [24] proposes an unsupervised method for the detection of distributed false data injection attacks. However, it could be used offline and could not be utilized in real-time mode for power systems.

B. PROBLEM MOTIVATION & CONTRIBUTIONS

Existing work neglect that in an interconnected power system composed of several control areas, boundary buses lose some of the redundant measurements since each area works more independently than in the centralized mode. This can negatively affect the security of boundary buses leading to cyber-attacks. In this work, we take advantage of the reduced redundancy to propose a coordinated FDIA to disrupt the DSE. We consider an attacker who can manipulate measurements of different areas simultaneously. Specifically, by focusing on the boundary buses of neighboring areas, we introduce a distributed FDIA on the DSE. In the proposed attack, the attacker targets one of the boundary buses and injects false data into all measurements of that bus, causing an error in its estimated state. It is shown that the attack must be conducted separately in each area but simultaneously to bypass BDD methods implemented in a distributed manner such as the robust DSE [5]. Moreover, it can be proved that if the injected measurements of each area are equal, the attack can also circumvent the convergence-based detection methods (e.g. [6]) (see the illustrative example in Section III-B). Moreover, to detect the attack, we utilize deep learning techniques to develop a real-time intelligent attack detection that captures the temporal correlation in power systems between consecutive time slots to differentiate malicious measurements from the normal ones. The main contributions of this paper are listed as follows.

1) We design a distributed FDIA that can bypass the current robust distributed estimator as well as the convergence-based detection method.
2) We then formulate a deep learning-based algorithm to capture the temporal correlation in power system measurements and detect the introduced attack.
3) We provide a comparison between the conventional classification algorithms and the carefully designed framework.

The rest of the paper is organized as follows: The model of multi-area power system and the manner of state estimation carried out on it are explained in Section II. Section III presents the conducted distributed FDI attack and how it circumvents existing detectors. A deep learning-based framework is developed in Section IV to detect the proposed attack. Section V discusses the numerical results of the attack and the detection method, and Section VI concludes the paper.

II. SYSTEM MODEL

The power system model used in this paper, along with the distributed state estimation (DSE), is described in this section. Assume an interconnected power grid with K area where in each area, a nonlinear connection between M_k samples denoted by $\mathbf{z}_k$ and N_k states denoted by $\mathbf{x}_k$ exists as shown in (1). In this equation, $\mathbf{h}_k(.)$ is the function vector that defines interconnections between measurements and system states, and $\mathbf{w}_k$ represents the vector of measurement error of area k.

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{w}_k \tag{1}$$

If the estimation method vector of area k is represented by $\mathbf{f}_k(\mathbf{x}_k; \mathbf{z}_k, \mathbf{h}_k(\mathbf{x}_k))$, the following optimization problem can be solved to estimate the states of each area $|\mathrm{s}|\mathbf{x}_k\mathbf{f}_k(\mathbf{x}_k; \mathbf{z}_k, \mathbf{h}_k(\mathbf{x}_k))$, that could be expanded to estimate all network state variables in a centralized mode as follows. $|\mathrm{s}|\mathbf{x}\sum_{k=1}^{K}\mathbf{f}_k(\mathbf{x})$ The boundary buses are buses that belong to more than one area. In other words, their state variables fall into the set of state variables of at least two areas, which we call neighboring areas. These areas are

included in $\mathbf{x}$ of (II). It is essential to gain a dependency on boundary busses in order to achieve a distributed state estimation. Neighboring areas must thus share the state variable of their boundary buses. A constraint for boundary buses of each of the two adjacent areas is established, and system states of both adjacent areas are made equivalent to reaching the following optimization problem. $|s|\mathbf{x}_k \sum_{k=1}^{K} \mathbf{f}_k(\mathbf{x}_k)\mathbf{x}_{k,k'} = \mathbf{x}_{k',k}, \quad \forall k' \in \mathcal{N}_k, \forall k$, where $\mathbf{x}_{k,k'}$ (or $\mathbf{x}_{k',k}$) denotes the system state-variable vector shared by area k and k' and $\mathcal{N}_k$ represents the adjacent areas of region k. $\mathbf{x}_{k,b}$ denotes the state-variable vector shared between area k and its neighbouring areas. In order to have a fully DSE, the optimization problem (II) should be arranged so that techniques for distributed optimization problems like the alternating direction method of multipliers (ADMM) [26] can be used. For doing this, Lagrange multipliers $\mathbf{v}_{k,k'}$ are proposed for the constraints of (II) [5]. To solve this, an iterative strategy is introduced as follows.

$$\mathbf{x}_k^{(t+1)} = (\mathbf{H}_k^{(t)T}\mathbf{H}_k^{(t)} + c\mathbf{D}_k)^{-1}(\mathbf{H}_k^{(t)T}\mathbf{z}_k + c\mathbf{D}_k\mathbf{p}_k^{(t)}) \quad (2a)$$

$$\mathbf{s}_k^{(t+1)} = \mathbf{U}_{\mathbf{x}_k}.\sum_{\forall k' \in \mathcal{N}_k} \mathbf{Y}_{k,k'}.\mathbf{x}_{k',k}^{(t+1)} \quad (2b)$$

$$\mathbf{p}_k^{(t+1)} = \mathbf{p}_k^{(t)} + \mathbf{s}_k^{(t+1)} - \frac{1}{2}(\mathbf{Y}_{k,b}.\mathbf{Y}_{k,b}^T.\mathbf{x}_k^{(t)} - \mathbf{s}_k^{(t)}), \quad (2c)$$

where $c > 0$ is a predetermined value, $\mathbf{H}_k^{(t)}$ is the Jacobian of function vector $\mathbf{f}_k$, $\mathbf{D}_k$ represents a diagonal matrix in which element $d_{i,i}$ counts the regions sharing state i of $\mathbf{x}_k$, $\mathbf{U}_{\mathbf{x}_k}$ is a diagonal matrix in which element $u_{i,i}$ is inverse of $d_{i,i}$, and non-diagonal elements are zero. In addition, $\mathbf{Y}_{k,k'}$ defines the dependencies between $\mathbf{x}_k$ and $\mathbf{x}_{k,k'}$ and $y_{i,j}$ element of $\mathbf{Y}_{k,k'}$ equals to one if state variable i of $\mathbf{x}_k$ conforms to the element j of $\mathbf{x}_{k,k'}$; otherwise, $y_{i,j}$ will be zero. Similarly, $\mathbf{Y}_{k,b}$ (or $\mathbf{Y}_{b,k}$) is a matrix that depicts the relationships between $\mathbf{x}_k$ and $\mathbf{x}_{k,b}$ ($\mathbf{x}_{b,k}$) vectors. Elements of matrix $\mathbf{Y}_{k,b}$ are similar to those of $\mathbf{Y}_{k,k'}$.

The DSE converges when the difference between the predicted values of two consecutive iterations of the ADMM algorithm is less than a predetermined threshold. Then we have $\forall k \in \mathcal{K}, \quad \|\mathbf{x}_k^{(t^*+1)} - \mathbf{x}_k^{(t^*)}\|_\infty \leq \epsilon$ when $t^* + 1$ is the stopping iteration and ϵ is known as the *convergence threshold*. For more information about the application of the ADMM method on nonlinear models of power systems, interested readers can refer to [27].

III. ATTACK DESCRIPTION

This section begins with a quick overview of the detection methods for attacks on the distributed estimators is presented. After that, the proposed distributed false data injection (FDI) attack is explained.

A. CYBERSECURITY OF DSE

Amongst attacks on DSE, the technique presented in [6] is chosen for description and analysis. The intruder in [6], by hampering the DSE from converging, reaches his or her goal for disabling the DSE. For this, the intruder compromises the telecommunications systems of zone $k_a \in \mathcal{K}$ so that he/she can manipulate some of the inputs to the DSE (e.g. system states $\mathbf{x}_{k,k_a}^{(t)}$ and $\mathbf{x}_{k_a,k}^{(t)}$ exchanged among k_a and its neighboring areas $k \in \mathcal{N}_{k_a}$). Vector $a_{k,k_a}^{(t)}$ (equals to $a_{k_a,k}^{(t)}$) represents the attack on the state variables exchanged between areas in $k \in \mathcal{N}_{k_a}$ and area k_a (specifically, from k_a to set k) at iteration t. Then, the consequent debased state variables can be denoted as a vector $\tilde{\mathbf{x}}_{k,k_a}^{(t)}$

$$\tilde{\mathbf{x}}_{k,k_a}^{(t)} = \mathbf{x}_{k,k_a}^{(t)} + \mathbf{a}_{k,k_a}^{(t)}. \quad (3)$$

As stated in [26, Appendix A, p. 106-110], ADMM converges when the following conditions are met: $\forall k \in \mathcal{K}$, $\mathbf{f}_k(\mathbf{x}_k; \mathbf{z}_k, \mathbf{h}_k(\mathbf{x}_k))$ function be acceptable, convex, closed and also the augmented Lagrangian

$$\mathcal{L} = \sum_{\forall k \in \mathcal{K}} \Bigg[\mathbf{f}_k(\mathbf{x}_k) + \sum_{k' \in \mathcal{N}_k} \left(\mathbf{v}_{k,k'}^T(\mathbf{x}_{k,k'}^{(t)} - \mathbf{x}_{k',k}^{(t)}) + c\| \frac{\mathbf{x}_{k,k'}^{(t)} - \mathbf{x}_{k',k}^{(t)}}{2} \|_2^2 \right) \Bigg] \quad (4)$$

is possessed a saddle point, where $\| \frac{\mathbf{x}_{k,k'}^{(t)} - \mathbf{x}_{k',k}^{(t)}}{2} \|_2^2 \to 0$ as $t \to \infty$. By considering the conditions, when the iteration counter t is increased and the DSE gets closer to the saddle point (i.e. solution), it is expected that the difference in the systems states of different regions k and k' (and all the other neighbors) is reduced. Regarding this, a convergence-based algorithm to catch cyber-anomalies on the DSE can be extended by calculating *mean squared disagreement* (MSD) for regions k and k' at the tth step as follows

$$d_{k,k'}^{(t)} = \frac{\|(\mathbf{x}_{k,k'}^{(t)} - \mathbf{x}_{k',k}^{(t)})/2\|_2^2}{\left|\mathbf{x}_{k,k'}^{(t)}\right|}, \quad (5)$$

in which, number of component in $\mathbf{x}_{k,k'}^{(t)}$ vector is represented by $\left|\mathbf{x}_{k,k'}^{(t)}\right|$. When the ADMM convergence requirements are met, a convergence issue (an attack) occurs, but for large values of t, there exists some k and $k' \in \mathcal{N}_k$ where $\sup\{d_{k,k'}^{(t')} : t' > t\} > 0, \quad \|(\mathbf{x}_k^{(t+1)} - \mathbf{x}_k^{(t)})\|_\infty > \epsilon$, and $\nexists n \in \mathbb{N}$ so that

$$\sup\{d_{k,k'}^{(t')} : t' > t\} > \sup\{d_{k,k'}^{(t')} : t' > t + n\}, \quad (6)$$

then there is a convergence issue (an attack).

B. THE PROPOSED DISTRIBUTED FDIA

This subsection presents the theoretical explanation of the developed distributed FDIA. The attack vector is designed to prevent the convergence of DSE to the correct estimation. The goal is accomplished utilizing neighboring areas' boundary buses. The FDIA proposed in this section is implemented such that, in addition to the centralized ones, it passes decentralized bad data detectors. Meanwhile, it can be unidentifiable after applying the convergence-based attack detection

methods. The attacker is expected to have full information on the power system which means that $\forall k \in \mathcal{K}$ areas, measurements vector $\mathbf{z}_k$ and the measurements to states mapping vector function $\mathbf{h}_k(\mathbf{x}_k)$ are known to the attacker. To implement a nonlinear attack, the attacker must know the estimated value of each state variable. Let us assume that the set of boundary buses targeted by the attacker is represented by $\mathcal{B}_a$. This set is the joint of all neighbors shown as $\mathcal{K}_a$. If the attacker aims to orchestrate an attack by injecting a value into the target buses' state, she/he must simultaneously launch $K_a = |\mathcal{K}_a|$ separate FDI attacks (one attack in each area). More specifically, if the boundary buses targeted by the attacker via area $i \in \mathcal{K}_a$ are represented as $\mathbf{x}_a$ and the vector added to the corresponding system state vectors is denoted as $\mathbf{c}_i$, $\mathbf{a}_i$ represents the attack vector injected into the measurements of area i. Using this definition, $\hat{\mathbf{x}}_i$ denotes the estimated values of state variables of area i under attack. The distributed FDIA can, therefore, be carried out as the following, which is consistent with the [29]

$$\mathbf{a}_i = \mathbf{h}_i(\hat{\mathbf{x}}_i + \mathbf{c}_i) - \mathbf{h}_i(\hat{\mathbf{x}}_i), \quad \forall i \in \mathcal{K}_a \tag{7}$$

$$\mathbf{z}_{a_i} = \mathbf{z}_i + \mathbf{a}_i, \quad \forall i \in \mathcal{K}_a. \tag{8}$$

It must be emphasized that while in each area i, the vectors $\mathbf{a}_i$ are injected separately to the vector of measurements $\mathbf{z}_i$, this happens simultaneously and on a unified observation of samples. If procedures of [29] are followed, it can be proved that the proposed FDIA can pass the bad data detectors. For this purpose, first and foremost, a look at the robust DSE [5] is taken as a distinguished instance of detectors of these types. For each area $k \in \mathcal{K}$, unknown vectors $\mathbf{o}_k$ are required to denote possible bad data in $\mathbf{z}_k$. In each area k, $\mathbf{o}_k$ variable vectors alter the estimation function $\mathbf{f}_k(\mathbf{x}_k, \mathbf{o}_k; \mathbf{z}_k, \mathbf{h}_k(\mathbf{x}_k))$; since these vectors only belong to region k (not shared with the neighbors) and conform to the measurements, $\mathbf{z}_k - \mathbf{o}_k$ notation is used for them in DSE equations. Two other equations are required to update $\mathbf{o}_k$ in the robust DSE as follows [5].

$$\mathbf{o}_k^{(t+1)} = \left[\mathbf{z}_k - \mathbf{H}_k^{(t)}\mathbf{x}_k^{(t+1)}\right]_{\lambda}^{+} \tag{9}$$

$$\begin{cases} x + \lambda & x < -\lambda \\ 0 & |\, x \,| \leq \lambda \\ x - \lambda & x > \lambda \end{cases} \tag{10}$$

Variable x in (10) also equals to $\mathbf{z}_k - \mathbf{H}_k^{(t)}\mathbf{x}_k^{(t+1)}$. After reformulating (1) as explained before, we have

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{o}_k + \mathbf{w}_k, \tag{11}$$

which is solved by using the iterative normal equation method with the first order optimality condition

$$-2\mathbf{H}_k^{(T)}(\hat{\mathbf{x}}_k)(\mathbf{z}_k - \mathbf{o}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k)), \tag{12}$$

in which $\mathbf{H}_k$ is the Jacobian matrix derived from the function vector $\mathbf{h}_k$ that its $m \times n$th entry is equal to $\frac{\partial h_m}{\partial x_n}$.

After the attack was carried out, the unchanged value of $\mathbf{o}_i$ vector holds if the following is met:

$$\left|\mathbf{z}_i - \mathbf{h}_i(\hat{\mathbf{x}}_i)\right| = \left|\mathbf{z}_{a_i} - \mathbf{h}_{a_i}(\hat{\mathbf{x}}_{a_i})\right|. \tag{13}$$

So, with replacing (7) and (8) in the equation (13) and some simplifications, we have:

$$\begin{aligned} &\left|\mathbf{z}_i - \mathbf{h}_i(\hat{\mathbf{x}}_i)\right| \\ &\quad = \left|\mathbf{z}_{a_i} - \mathbf{h}_{a_i}(\hat{\mathbf{x}}_{a_i})\right| \Rightarrow \\ &\left|\mathbf{z}_i - \mathbf{o}_i - \mathbf{h}_i(\hat{\mathbf{x}}_i)\right| \\ &\quad = \left|\mathbf{z}_{a_i} - \mathbf{o}_{a_i} - \mathbf{h}_{a_i}(\hat{\mathbf{x}}_{a_i})\right| \\ &\quad = \left|\mathbf{z}_i + \mathbf{a}_i - \mathbf{o}_{a_i} - \mathbf{h}_i(\hat{\mathbf{x}}_i + \mathbf{c}_i)\right| \\ &\quad = \left|\mathbf{z}_i + (\mathbf{h}_i(\hat{\mathbf{x}}_i + \mathbf{c}_i) - \mathbf{h}_i(\hat{\mathbf{x}}_i)) - \mathbf{o}_{a_i} - \mathbf{h}_i(\hat{\mathbf{x}}_i + \mathbf{c}_i)\right| \Rightarrow \\ &|\mathbf{o}_i| = \left|\mathbf{o}_{a_i}\right| \end{aligned} \tag{14}$$

where attacked $\mathbf{o}_i$ is represented by $\mathbf{o}_{a_i}$. Above equations demonstrate that $\mathbf{o}_k$ will not alter after the attack, so the FDIA proposed here can circumvent the robust DSE. However, additional conditions explained in [6] are required for this. These conditions are as follows.

Proposition: Assume that the distributed FDI attack described above is carried out successfully. When the distance between the items with non-zero values of $\mathbf{c}_i$ vectors, corresponding to the system states of buses in $\mathcal{B}_a$, is lower than the predetermined value ϵ, the FDIA can circumvent the convergence-based detection technique proposed in [6].

Proof: Let us show the MSD between every two regions i and j of $\mathcal{K}_a$ in iteration t prior to the conduction of attack as:

$$d_{i,j}^{(t)} = \frac{\|(\mathbf{x}_{i,j}^{(t)} - \mathbf{x}_{j,i}^{(t)})/2\|_2^2}{|\mathcal{B}_a|}. \tag{15}$$

Now suppose an attack is injected in such a way that the distance between the non-zero elements of $\mathbf{c}_i$ and $\mathbf{c}_j$ for every distinct $\{i, j\} \in \mathcal{K}_a$ is lower than ϵ (i.e. the predetermined threshold of convergence). Thus, after the attack, the MSD equation could be shown as

$$\begin{aligned} d_{i,j}^{(t)} &= \frac{\left\|\left((\mathbf{x}_{i,j}^{(t)} + \mathbf{c}_i) - (\mathbf{x}_{j,i}^{(t)} + \mathbf{c}_j)\right)/2\right\|_2^2}{|\mathcal{B}_a|} \\ &= \frac{\left\|\left((\mathbf{x}_{i,j}^{(t)} - \mathbf{x}_{j,i}^{(t)}) + (\mathbf{c}_i - \mathbf{c}_j)\right)/2\right\|_2^2}{|\mathcal{B}_a|}, \end{aligned} \tag{16}$$

accordingly, if for each pair $\{i, j\} \in \mathcal{K}_a$, $\mathbf{x}_{i,j}^{(t)}$ and $\mathbf{x}_{j,i}^{(t)}$ representing $\mathbf{x}_a$ satisfy the MSD conditions before the FDIA, then considering $(\mathbf{c}_i - \mathbf{c}_j) < \epsilon$, $d_{i,j}^{(t)}$ value remains unchanged after this attack and consequently, FDIA can get through the method proposed in [6]. A point needs to be made clear here: $\forall i \in \mathcal{K}_a$, it is not required that the non-zero elements of $\mathbf{c}_i$ vectors have the same value, but only, the distance between the elements of $\mathbf{c}_i$ vectors, $\forall i \in \mathcal{K}_a$ which correspond to a determined bus in $\mathcal{B}_a$ have to be lower than ϵ. As one will see in Section V, 10% incremental and decremental attacks are simulated.

Illustrative Example: Assume that $\mathbf{z}_1$ and $\mathbf{z}_2$, the normal measurements for areas 1 and 2, respectively, bypass the distributed BDD (robust DSE) in [5]. The manipulated measurements $\mathbf{z}_{a_1} = \mathbf{z}_1 + \mathbf{a}_1$ and $\mathbf{z}_{a_2} = \mathbf{z}_2 + \mathbf{a}_2$ for the areas 1 and 2, respectively, can bypass the distributed BDD if

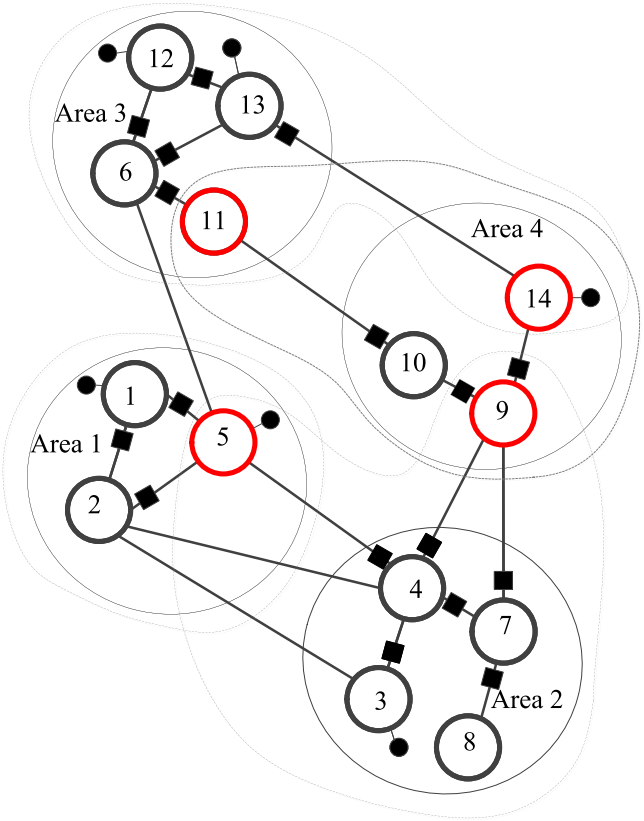

FIGURE 1. The IEEE 14-bus test system divided into 4 zones [5].

$\mathbf{a}_1 = \mathbf{h}_1(\hat{\mathbf{x}}_1 + \mathbf{c}_1) - \mathbf{h}_1(\hat{\mathbf{x}}_1)$ and $\mathbf{a}_2 = \mathbf{h}_2(\hat{\mathbf{x}}_2 + \mathbf{c}_2) - \mathbf{h}_2(\hat{\mathbf{x}}_2)$, in which $\hat{x}_1$ and $\hat{x}_2$ are the estimated values of state vectors of area 1 and 2, respectively. Now, if the distance between the components of vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ be lower than the threshold value introduced in [6], the FDIA will get through the detection method proposed in [6]. This description can be extended to boundary buses with over two adjacent areas.

For the bus number 5 in the system presented in Fig. 1 and for attack size α, $\mathbf{c}_1$ and $\mathbf{c}_2$ are vectors of size 6×1 and 12×1 where the non-zero vectors relevant to bus number 5 equal to α. Vector $\mathbf{a}_1$ has size 10×1 where its components representing the voltage value of bus number 5 and line currents (1,5) and (2,5) are not zero. In addition, $\mathbf{a}_2$ is a 14×1 vector which its elements corresponding to the line current (4,5) is non-zero. The size of other vectors of test system are brought in Table 1. If we obtain the average of estimates, the distance between the estimated values for the voltage of bus 5 calculated based on area 1 is 1.062018326 and through area 2 is 1.062018324 which is much lower than the ϵ in [5] (10^{-3}). This case is also observed for the voltage angle of $5th$ bus over the zones 1 (0.109613) and 2 (0.109614).

IV. DEEP LEARNING-BASE DETECTION METHOD

Decentralized bad data detection methods are shown vulnerable to the proposed distributed FDIA based on section II. The issue with the state estimation and convergence-based detection methods is that a power system's temporal pattern is not observed and depends only on the specified topology and measurements. Therefore, if a detection method can be designed, which can predict based on the measurement history, it can not be easily circumvented. Recurrent neural networks (RNNs) offer such potentiality. RNN is an ideal candidate because of its established ability to learn contextual connections in terms of its success in machine translation and speech recognition [30]. This section summarizes several key ideas for RNNs and long short-term memory (LSTM) block structure.

A. THE LSTM MODEL

Compared to conventional feedforward neural networks assuming inputs/outputs are mutually independent, RNNs are a particular type of neural networks that use sequential information to forecast the output [30]. The temporal correlations between previous and current knowledge can be defined in RNN models. This implies that, with the issue of time series, the decision taken by the RNN at stage $t-1$ could have an impact on the decision taken at a later time step t. This feature of RNNs is ideal since there is a temporal correlation in power systems between consecutive time slots. Utilizing this characteristic, in this section, a real-time FDIA identification mechanism is proposed using recent advances in RNNs.

RNNs are trained by Backpropagation Through Time (BPTT) Algorithm by adjusting the weights of the network [31]. BPTT expands the Backpropagation Learning (BPL) Algorithm over a time series in which the gradient of each output is calculated by both current and prior steps. However, RNN is limited in learning long-term dependency due to gradient loss or explosion problems, as the gradients back propagated over certain steps [32], [33]. In order to avoid the vanishing gradient and allow RNN to learn the long-term dependency, a variant of RNN known as long short-term memory (LSTM) architecture has been introduced by [34]. It has been the most effective implementation of the RNN and has achieved tremendous prominence in many subsequent applications and usually outperforms the conventional RNNs.

Reference [35] has carried out an in-depth review of LSTM's overall structure and recent development. We follow a naming convention identical to [35] to present the idea of the LSTM briefly. To remember long-time temporal dependencies, LSTM describes and retains an internal memory cell state C during the whole life cycle as the most important LSTM system component. Besides the memory cell state, the LSTM architecture also defines input node g_c, input gate i_g, output gate o_g, and forget gate f_g. The following equations give the formulations of an LSTM block:

$$g_c(t) = \phi(W_{gq}TS(t) + W_{gh}h(t-1) + b_g) \quad (17)$$
$$i_g(t) = \sigma(W_{iq}TS(t) + W_{ih}h(t-1) + b_i) \quad (18)$$
$$f_g(t) = \sigma(W_{fq}TS(t) + W_{fh}h(t-1) + b_f) \quad (19)$$
$$o_g(t) = \sigma(W_{oq}TS(t) + W_{oh}h(t-1) + b_o) \quad (20)$$
$$C(t) = g_c(t) \odot i_g(t) + C(t-1) \odot f_g(t) \quad (21)$$
$$h(t) = \phi(C(t)) \odot o_g(t) \quad (22)$$

TABLE 1. The used vectors in the illustrative example.

Area	Size of Vectors		Attacked Measurements of Buses			
	$\mathbf{a}_i$	$\mathbf{c}_i$	5	9	11	14
1	6×1	10×1	bus voltage 5, line current (1,5), and (2,5)			
2	12×1	14×1	line current (4,5)	line current (4,9), and (7,9)		
3	10×1	14×1			line current (6,11)	line current (13,14)
4	8×1	8×1		line current (9,10), and (9,14)	line current (10,11)	bus voltage 14

where $\mathbf{TS}(t)$ is input sequence for an LSTM at time point t; $\mathbf{W}_{gq}$, $\mathbf{W}_{iq}$, $\mathbf{W}_{fq}$, $\mathbf{W}_{oq}$, $\mathbf{W}_{gh}$, $\mathbf{W}_{ih}$, $\mathbf{W}_{fh}$, $\mathbf{W}_{oh}$ indicate the correlation weight matrices between the corresponding inputs of the network activation functions; $h(t)$ displays the LSTM module outputs at step t; $\odot$ represents element-wise multiplication; b_g, b_i, b_f, and b_o are biases of corresponding nodes and gates; $\phi(x)$ indicates the *tanh* function, and $\sigma(x)$ is the *sigmoid* activation function.

B. THE LSTM BASED ATTACK DETECTOR

In a power system, measurements and state vectors could be classified as long temporal sequences. Therefore, in the proposed FDIA detector, LSTM architecture is adopted to carefully design a real-time deep learning-based attack detection framework. Specifically, the proposed detection mechanism works as a sequence classifier to discriminate attacks from normal measurements by learning the temporal data correlation of a sequence of input data. In this paper, system states are selected as the inputs to our attack detector. This means that we used a minimal set of features to find attacks. Given a set of samples S and labels Y (normal versus attack), the proposed method's objective is developing a learning function $S \rightarrow Y$ that maps the feature to its corresponding label and classifying the measurements into two classes, normal and tampered. Therefore, the output of the proposed FDIA detector is as follows:

$$y_i = \begin{cases} 1, & \text{FDIA in the input data} \\ 0, & \text{otherwise} \end{cases}. \tag{23}$$

To carefully develop an LSTM based FDIA detector, we need to meticulously tune its hyper-parameters, namely, the input layer's number of neurons, number of LSTM hidden layer(s), number of fully-connected (dense) layers, and the number of neurons in the output layer, to obtain an acceptable attack detection accuracy. The structure of the employed model for constructing attack detection is shown in Fig. 2. The input layer is in charge of obtaining observations from the dataset and delivering them to the first hidden layer for analysis. Non-linear relationships between data are represented in the LSTM hidden layers. To properly classify a system with increasingly intricate relationships, it is necessary to add more hidden layers. The dense layers take data from the previous hidden layer and calculate the likelihood of samples belonging to each class. Lastly, the output layer gets the probabilities of the final fully-connected (FC) layer and assigns samples labels. Based on observations from many experiments, the constructed model comprises one Dense layer with 11 neurons, three LSTM hidden layers, and 28 neurons for the input layer.

Other parameters are required to be tuned in this LSTM model to ensure that FDIA detector achieves a promising detection accuracy. These parameters are dropout and sequence length. Dropout is a regularization method for neural networks to minimize the risk of overfitting [36]. The sequence length specifies the number of samples that should precede the existing observation when it is entered into the LSTM. In the constructed LSTM, dropout is applied at a 40% ratio and the sequence length is set to 3 since it provides the best results based on many conducted experiments. It is also worth noting that Adam optimizer [37] is utilized to define optimum values of the learning parameters of the constructed LSTM model, i.e., the values for the correlation weight matrices in (20)-(25). Finally, since the LSTM is sensitive to data scale, all inputs should be scaled to the range of (0, 1) before the system states enter the proposed model's input layer.

C. MEASURES

F-Measure is used to validate the developed LSTM RNN-based FDI detector. The F-Measure is calculated in the following manner [38]:

$$F_1 = \left(\frac{2 \times P_r \times R_e}{P_r + R_e}\right), \tag{24}$$

where R_e is the recall and P_r is the precision, which are determined in the following manner:

$$P_r = \left(\frac{TP}{TP + FP}\right) \quad R_e = \left(\frac{TP}{TP + FN}\right), \tag{25}$$

where true positive (TP) represents the number number of correctly recognized attack measurements, false positives (FP) represents the number of incorrectly observed attacks, true negative (TN) represents the proportion of successfully identified normal findings, and false negative (FN) represents the number of overlooked attack vectors. The F-Measure value 1 shows that every sample labeled as normal is truly normal, and each observation designated as an attack is genuinely a manipulated one.

V. NUMERICAL RESULTS

The developed distributed state estimator, attack and detection method are numerically tested using MATLAB and PyTorch in this section. IEEE 14-bus system is used as a benchmark for completing the simulations. The test system's

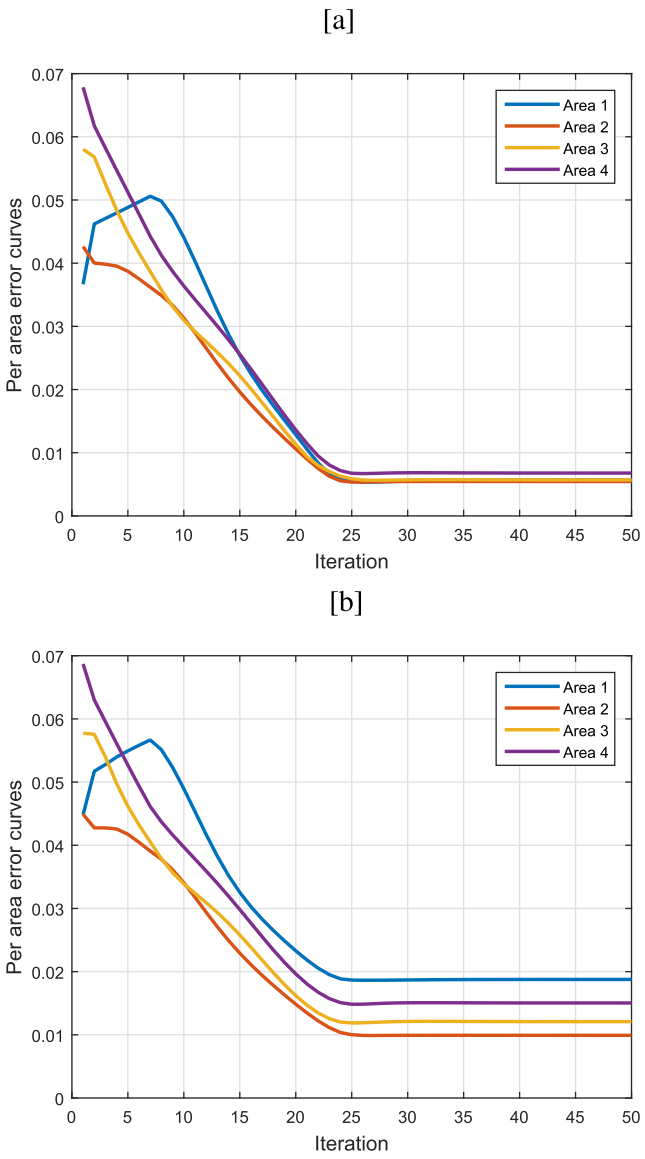

FIGURE 2. Estimation error per area for four regions: a) the normal state, and b) attack to buses 5 and 14.

load data is derived from the New York independent system operator (NYISO) over a period of two days. Parameters' values and the system's states are obtained using MATPOWER [39]. As depicted in Fig. 1, the IEEE 14-bus grid is partitioned into 4 areas; buses 1, 2, and 5 are in area 1, buses 3, 4, 7, and 8 are in area 2, buses 6, 11, 12, and 13 are in area 3, and buses 9, 10, and 14 are in area 4. Boundary buses are marked in red. The state vector contains the magnitude and phase angles of all bus voltages. Measurements consist of PMU recordings on 6 bus voltages (circles in Fig. 1) and 17 line currents (squares in Fig. 1), expressed in rectangular coordinates too. The measurement noise is modeled as a zero-mean Gaussian with a standard deviation of 0.02 and 0.05. It is noteworthy that the current measurements on tie-lines connected to the boundary buses specify each area's actual range. In this regard, according to Fig. 1, it is specified using the gray dashed lines that the actual

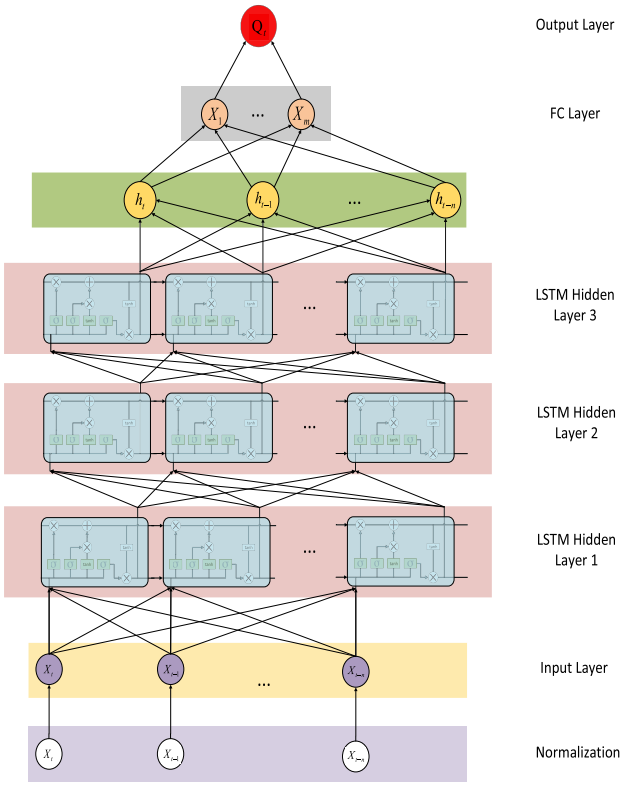

FIGURE 3. Proposed LSTM-based attack detection model.

range of area 2 encircles buses 5 and 9, the actual range of area 3 encircles bus 14, and the actual range of area 4 encircles bus 11.

A. DSE AND DETECTION METHODS

The developed distributed FDIA is employed on each of the test system's boundary buses and multiple buses 5 and 11 in order to assess its effect. Furthermore, a coordinated attack on two border buses is also modeled. The magnitude of attack is always 0.1 of the real measurement of the manipulated variable. In particular, two injection quantities of 90% and 110% of the true value are simulated for each attacked state variable. 90% signifies that the modified state variable has been set to a value that is 10% less than the true value. 100 random DSE and attack scenarios are conducted, and the average standard deviation of state variables are calculated throughout these runs. As described in the III-B, the suggested attack is capable of evading the bad data detectors of [5]. Figure 3(a)-(b) shows the per-area estimate error relative to the genuine data in the normal condition and attack to two buses 5 and 14, respectively. It should be mentioned that where areas are operated independently, none of their operators is capable of obtaining neighborhood-specific area errors and consequently, the graphics shown here can not be used to detect an attack.

In addition to decentralized BDD methods, as shown in Sec. III-B and numerically made plain in that section as an illustrative example, the developed attack could circumvent

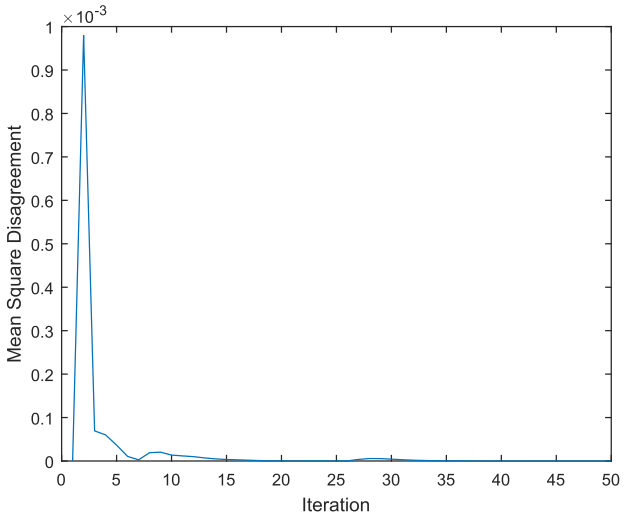

FIGURE 4. The MSD of coordinated attack to buses 14 and 5.

TABLE 2. Training and validation accuracy of the methods.

	Proposed Method		C4.5		AdaBoost (C5.0)		MLP		Naïve Bayes		Random Forest	
Inject	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec
θ_5	1	1	1	1	1	1	1	1	0.99	0.99	1	1
θ_9	1	1	1	0.99		1	1	1	0.99	0.99	1	1
θ_{11}	1	1	1	1	1	1	1	1	1	1	1	1
θ_{14}	1	1	0.99	0.99	1	1	1	1	0.99	0.98	1	1
$\theta_{5,\ 11}$	1	1	0.99	0.99	1	0.99	1	1	0.98	0.98	1	1
V_5	1	1	1	1	1	1	1	1	1	1	1	1
V_9	1	1	1	1	1	1	1	1	1	1	1	1
V_{11}	1	1	1	1	1	1	1	1	1	1	1	1
V_{14}	1	1	1	1	1	1	1	1	1	1	1	1
$V_{5,\ 11}$	1	1	1	1	1	1	1	1	1	1	1	1

the convergence-based detection techniques. Convergence-based detection approaches, as detailed in [6], identify an assault by monitoring oscillations in the development of MSD. In each of the attack cases, this index is examined, and it is found that it is zero for targeting buses 14 and 11. The MSD diagrams for the coordinated attack to buses 5 and 14 over the iterations are shown in Fig. 4. As it is evident from this diagram, the MSD has no oscillation and by increasing the iterations, It eventually reaches zero or a particular value.

B. PROPOSED DETECTION METHODS

Tables 1 and 2 present the detailed results of the proposed method over different attack scenarios. As one can see, the proposed method is compared with the conventional machine learning-based FDIA detectors including C4.5, Adaboost with C4.5-based classifier, multilayer perceptron, Naive Bayes, and Random Forest [38]. For the proposed LSTM-based FDIA detector, the dataset was divided for the training and test sets. The ratio for the training set is 70%, the ratio for the testing set is 30%. For the training set, we used a 80% – 20% train-validation pattern. For other classification methods, we used tenfold cross-validation. Table 1 presents

TABLE 3. Test accuracy of the methods.

	Proposed Method		C4.5		AdaBoost (C4.5)		MLP		Naïve Bayes		Random Forest	
Inject	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec	Inc	Dec
θ_5	0.99	1	0.34	0.47	0.39	0.47	1	0	0.87	0.88	0.58	0.94
θ_9	1	1	0.38	0.72	0.38	0.58	0.55	0.99	0.94	0.72	0.87	0.62
θ_{11}	1	1	0.40	0.45	0.98	0.45	0	0.57	0.99	0.94	0.88	0.44
θ_{14}	1	1	0.73	0.40	0.69	0.40	0.99	1	0.82	0.94	0.72	0.97
$\theta_{5,\ 11}$	1	1	0.61	0.37	0.61	0.37	0.94	0.94	0.89	0.89	0.73	0.5
V_5	1	1	1	1	1	1	1	1	1	1	1	1
V_9	1	1	1	1	1	1	1	1	1	1	1	1
V_{11}	1	1	1	1	1	1	1	1	1	1	1	1
V_{14}	1	1	1	1	1	1	1	1	1	1	1	1
$V_{5,\ 11}$	1	1	1	1	1	1	1	1	1	1	1	1

the results for the validation accuracy of the proposed method and the training accuracy of other algorithms. Table 2 presents the results of applying the developed models on the test dataset. As can be seen from the results, the proposed method is superior to other methods in detecting attacks on phase angles while for voltage magnitudes, all methods have been able to detect all attacks. This is because the variance of changes in the phase angles is much more than voltage magnitudes. In other words, the data distribution change in voltage magnitudes is not too much to make any manipulation easier to detect.

VI. CONCLUSION

While emphasizing the importance of boundary buses in multi-area power systems, a distributed FDIA has been developed in this paper. It has been shown that the developed attack strategy can bypass distributed BDD methods and convergence-based detection ones. The first is provided by attacking the state of boundary buses through all corresponding measurements of those buses in neighboring areas, and the latter is achieved by equalizing the injected values by each area. Afterward, an LSTM-based framework is carefully designed to detect FDIA, which considers the temporal dependency of consecutive system states. Numerical results showed that the proposed method could gain excellent detection accuracy and deliver remarkable accuracy improvements to existing FDIA detectors.

REFERENCES

[1] Y. Wang, S. Wang, and L. Wu, "Distributed optimization approaches for emerging power systems operation: A review," *Electr. Power Syst. Res.*, vol. 144, pp. 127–135, Mar. 2017.
[2] T. Odun-Ayo and M. L. Crow, "Structure-preserved power system transient stability using stochastic energy functions," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1450–1458, Aug. 2012.
[3] S. Savulescu, *Real-Time Stability in Power Systems*. Berlin, Germany: Springer, 2014.
[4] D. K. Molzahn, F. Dárfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.
[5] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.

[6] O. Vuković and G. Dün, "Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1500–1508, Jul. 2014.

[7] F. Ahmadloo and F. R. Salmasi, "A stealth integrity targeted cyber-attack in distributed electric power networks with local model information," *Asian J. Control*, vol. 21, no. 1, pp. 545–558, Jan. 2019.

[8] Z. Li, M. Shahidehpour, and F. Aminifar, "Cybersecurity in distributed power systems," *Proc. IEEE*, vol. 105, no. 7, pp. 1367–1388, Jul. 2017.

[9] A. Sargolzaei, K. K. Yen, and M. N. Abdelghani, "Preventing time-delay switch attack on load frequency control in distributed power systems," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1176–1185, Mar. 2016.

[10] J. Liu, W. Suo, L. Zha, E. Tian, and X. Xie, "Security distributed state estimation for nonlinear networked systems against DoS attacks," *Int. J. Robust Nonlinear Control*, vol. 30, no. 3, pp. 1156–1180, Feb. 2020.

[11] W. Chen, D. Ding, H. Dong, and G. Wei, "Distributed resilient filtering for power systems subject to denial-of-service attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 8, pp. 1688–1697, Aug. 2019.

[12] T. V. Vu, B. H. L. Nguyen, T. A. Ngo, M. Steurer, K. Schoder, and R. Hovsapian, "Distributed optimal dynamic state estimation for cyber intrusion detection in networked DC microgrids," in *Proc. 45th Annu. Conf. Ind. Electron. Soc.*, Lisbon, Portugal, Oct. 2019, pp. 4050–4055.

[13] B. P. Poudel, A. Mustafa, A. Bidram, and H. Modares, "Detection and mitigation of cyber-threats in the DC microgrid distributed control system," *Int. J. Electr. Power Energy Syst.*, vol. 120, Sep. 2020, Art. no. 105968.

[14] A. Minot, H. Sun, D. Nikovski, and J. Zhang, "Distributed estimation and detection of cyber-physical attacks in power systems," in *Proc. IEEE Int. Conf. Commun. Workshops*, Shanghai, China, Oct. 2019, pp. 1–6.

[15] N. Forti, G. Battistelli, L. Chisci, S. Li, B. Wang, and B. Sinopoli, "Distributed joint attack detection and secure state estimation," *IEEE Trans. Signal Inf. Process. over Netw.*, vol. 4, no. 1, pp. 96–110, Mar. 2018.

[16] Y. Guan and X. Ge, "Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 4, no. 1, pp. 48–59, Mar. 2018.

[17] L. An and G.-H. Yang, "Distributed secure state estimation for cyber–physical systems under sensor attacks," *Automatica*, vol. 107, pp. 526–538, Sep. 2019.

[18] M. M. Rana, R. Bo, and A. Abdelhadi, "Distributed grid state estimation under cyber attacks using optimal filter and Bayesian approach," *IEEE Syst. J.*, vol. 15, no. 2, pp. 1970–1978, Jun. 2021.

[19] M. Yang, H. Zhang, C. Peng, and Y. Wang, "A penalty-based adaptive secure estimation for power systems under false data injection attacks," *Inf. Sci.*, vol. 508, pp. 380–392, Jan. 2020.

[20] X. Wang, X. Luo, M. Zhang, and X. Guan, "Distributed detection and isolation of false data injection attacks in smart grids via nonlinear unknown input observers," *Int. J. Electr. Power Energy Syst.*, vol. 110, pp. 208–222, Sep. 2019.

[21] R. Anguluri, V. Katewa, and F. Pasqualetti, "Centralized versus decentralized detection of attacks in stochastic interconnected systems," *IEEE Trans. Autom. Control*, vol. 65, no. 9, pp. 3903–3910, Sep. 2020.

[22] M. Mohammadpourfard, A. Khalili, I. Genc, and C. Konstantinou, "Cyber-resilient smart cities: Detection of malicious attacks in smart grids," *Sustain. Cities Soc.*, vol. 75, Dec. 2021, Art. no. 103116.

[23] S. Li, Y. Yilmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.

[24] A. Shefaei, M. Mohammadpourfard, B. Mohammadi-ivatloo, and Y. Weng, "Revealing a new vulnerability of distributed state estimation: A data integrity attack and an unsupervised detection algorithm," *IEEE Trans. Control Netw. Syst.*, early access, Jun. 22, 2021, doi: 10.1109/TCNS.2021.3091631.

[25] Z. Li, M. Shahidehpour, and X. Liu, "Cyber-secure decentralized energy management for IoT-enabled active distribution networks," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 5, pp. 900–917, Sep. 2018.

[26] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Nov. 2010.

[27] R. S. Kar, Z. Miao, M. Zhang, and L. Fan, "ADMM for nonconvex AC optimal power flow," in *Proc. North Amer. Power Symp. (NAPS)*, Morgantown, WV, USA, Sep. 2017, pp. 1–6.

[28] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, p. 13, 2011.

[29] H. Moayyed, M. Mohammadpourfard, C. Konstantinou, A. Moradzadeh, B. Mohammadi-Ivatloo, and A. P. Aguiar, "Image processing based approach for false data injection attacks detection in power systems," *IEEE Access*, vol. 10, pp. 12412–12420, 2022.

[30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[31] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

[32] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[33] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, S. C. Kremer, and J. F. Kolen, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies," in *A Field Guide to Dynamical Recurrent Networks*. Piscataway, NJ, USA: IEEE Press, 2001.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[35] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," 2015, *arXiv:1506.00019*.

[36] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012, *arXiv:1207.0580*.

[37] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, San Diego, CA, USA, Jul. 2015, pp. 1–14.

[38] M. Mohammadpourfard, I. Genc, S. Lakshminarayana, and C. Konstantinou, "Attack detection and localization in smart grid with image-based deep learning," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2021, pp. 121–126.

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

MOSTAFA MOHAMMADPOURFARD received the Ph.D. degree from Shiraz University. He is currently a Marie Skłodowska-Curie Fellow at Istanbul Technical University and an Assistant Professor with the Sahand University of Technology. His research interests include power system cyber-security, machine learning, E-government, and computer security, with an emphasis on human factors and science for policy.

YANG WENG (Senior Member, IEEE) received the B.E. degree in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, the M.Sc. degree in statistics from the University of Illinois at Chicago, Chicago, IL, USA, the M.Sc. degree in machine learning of computer science, and the M.E. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University (CMU), Pittsburgh, PA, USA. He joined Stanford University, Stanford, CA, USA, as a TomKat Fellow of Sustainable Energy. He is currently an Assistant Professor in electrical, computer, and energy engineering with Arizona State University, Tempe, AZ, USA. His research interests include power systems, machine learning, and renewable integration. He was a recipient of the CMU Deans Graduate Fellowship, in 2010, the Best Paper Award at the International Conference on Smart Grid Communication (SGC), in 2012, the First Ranking Paper of SGC, in 2013, the Best Papers at the Power and Energy Society General Meeting, in 2014, the ABB Fellowship, in 2014, and the Golden Best Paper Award at the International Conference on Probabilistic Methods Applied to Power Systems, in 2016.

ABDULLAH KHALILI is currently an Assistant Professor with Hormozgan University. His current research interests include cyber-physical systems and machine learning.

ALIREZA SHEFAEI is currently pursuing the Ph.D. degree with TU Delft. His research interests include optimization and power systems.

ISTEMIHAN GENC (Member, IEEE) received the B.Sc. degree in electrical engineering from Istanbul Technical University, the M.Sc. degrees in electrical engineering, systems and control engineering, and systems science and mathematics from Istanbul Technical University, Bogazici University, and Washington University, respectively, and the Doctor of Science (D.Sc.) degree from Washington University in St. Louis, in 2001. He joined the Department of Electrical Engineering, Istanbul Technical University. He was a Visiting Scholar during his postdoctoral studies with the Department of Electrical Engineering, Arizona State University, and also a Visiting Professor with the Department of Electrical and Systems Engineering, Washington University.

BEHNAM MOHAMMADI-IVATLOO (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2006, and the M.Sc. and Ph.D. degrees (Hons.) from the Sharif University of Technology, Tehran, Iran, in 2008. He is currently a Professor with the Faculty of Electrical and Computer Engineering, University of Tabriz. His research interests include economics, operation, and planning of intelligent energy systems in a competitive market environment. He is an Associate Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE ACCESS, *IET Smart Grid*, and the *Sustainability*.