Extreme Learning Machine-Based State Reconstruction for Automatic Attack Filtering in Cyber Physical Power System

Ting Wu, Wenli Xue, Huaizhi Wang, *Member, IEEE*, C. Y. Chung, *Fellow, IEEE*, Guibin Wang, *Member, IEEE*, Jianchun Peng, *Senior Member, IEEE*, and Qiang Yang, *Senior Member, IEEE*

***Abstract*—Successful detection of false data injection attacks (FDIAs) and removal of state bias due to FDIAs are essential for ensuring secure power grids operation and control. This paper first extends the approximate dc model of FDIA to a more general ac model that can handle both traditional and synchronized measurements. To automatically filter out the established FDIAs, we propose a state reconstruction scheme consisting of a contaminated state separation method, an enhanced bad data identification approach and a state recovery algorithm. In this scheme, a classifier is developed by aggregating a series of extreme learning machines (ELMs) to detect anomaly states caused by FDIAs. Gaussian random distribution and Latin hypercube sampling are adopted to initialize the input weights of base ELMs, which can provide more diversities to enhance the ensemble performance. Then, to identify the exact locations of the compromised measurements, a state forecasting-based bad data identification approach is proposed by exploiting the consistency between the forecasted and the received measurements. Finally, an effective state recovery algorithm applies quasi-Newton method and Armijo line search to address the possible system unobservable problem due to the removal of attacked measurements. Numerical tests on serval IEEE standard test systems verify the efficiency of the proposed FDIA model and state reconstruction scheme.**

***Index Terms*—Cyber physical power system, false data injection attack, state estimation, extreme learning machine, quasi-Newton method.**

I. Introduction

With the incorporation of the remarkable advancements in sensing, monitoring, control technologies, and also the tight integration with cyber infrastructure and advance computing and communication technologies, the traditional electric grid is gradually evolving towards a deeply intertwined cyber physical power system (CPPS) which tends to be much more reliable, efficient, and intelligent [1, 2]. The CPPS relies greatly on the associated cyber network that has indeed revolutionized power grid efficiency and operational performance due to the two-way communication between utilities and consumers [3]. However, the integration of physical and cyber components also gives rise to cyber-attack threats in a power system, which can result into huge economical loss, power outage and even system blackouts [4] .

As an essential tool for providing accurate system snapshots to several crucial applications in energy management systems (EMS) [5], state estimation (SE) is also susceptible to cyber-attacks. False data injection attack (FDIA) [6] is an important type of typical malicious cyber-attacks, which can cause the state estimator to output erroneous state values to the system operator, and thus make either physical or economic impacts on the power system. The existing studies with regard to FDIAs on CPPS can be categorized into two groups. The first group concentrates on how to optimally construct a valid FDIA. Network topology attacks [7], load redistribution attacks [8], denial of service attacks [9], and state attacks [10] have been constructed to fulfill various malicious objectives.

However, most existing FDIA strategies have been formulated on a dc model, which has a greater chance of introducing errors in the measurements and thus trigger bad data detection (BDD), e.g., the largest normalized residual test and the hypothesis testing identification (HTI) method [11], in a real ac power grid. A graph theory-based algorithm proposed in [12] determined how many and which measurements need to be modified in order to minimize the efforts in keeping the attack hidden from BDD in ac SE. Operation scenario-based two-stage sparse cyber-attack models of ac smart grid with complete and incomplete network information were proposed in [13]. A dynamic cyber-attack model based on ac SE was proposed in [3] to account for dynamic characteristics of attack behaviors. These presented FDIA strategies are designed for supervisory control and data acquisition (SCADA) system.

With increasing use of synchronized phasor measurement units (PMUs) for wide area situational awareness in recent years, the measurement redundancy and SE accuracy have been significantly improved because PMUs can provide synchronized voltage and current phasors. Due to financial constraints, the number of PMUs developed in real systems worldwide is still insufficient to make the system fully observable [14]. Thus, the PMU measurements should be used along with SCADA measurements and provide a unified view of system operation

This work was supported in part by the National Natural Science Foundation of China under Grant 51907126 and Grant 51707123, and in part by the Foundations of Shenzhen Science and Technology Committee under Grant JCYJ20170817100412438 and Grant JCYJ20190808141019317.

T. Wu is with the College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: wuting@szu.edu.cn).

W. Xue, H. Wang (*Corresponding Author*), G. Wang and J. Peng are with the College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China (email: wanghz@szu.edu.cn).

C. Y. Chung is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: c.y.chung@usask.ca).

Q. Yang is with the College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: qyang@zju.edu.cn).

state through a hybrid state estimator. Some existing works on detecting cyber-attacks [15] assumed that PMU measurements are secure and focused on obtaining the minimum number of PMUs and their placement that are required to detect an attack. However, this assumption is not realistic, because an attacker can access and modify the PMU data in three ways: attacking the PMUs, tampering with the communication network, and breaking into the synchrophasor system through the control center office [16]. Consequently, the design of FDIAs for the hybrid SE would help us to understand the in-depth of cyber-security and implement valid prevention measurements, which deserves further investigations and endeavors.

Since disturbing the normal operation of CPPS by launching FDIAs has tremendous financial and security effects, the second group of the researches related to FDIAs focuses on dissecting countermeasures against cyber-attack by applying various techniques, such as statistical method [17], physical method [15], sparse optimization [4], interval SE [3, 13], time-series simulation [18], and machine learning method [19]. These strategies all demonstrate satisfactory detection performance and false-alarm rates against FDIAs. However, these methods can only discriminate whether there is an attack in the CPPS or not, and it is impossible for them to identify which buses or states in CPPS are contaminated and further to recovery these states. As one of the important purposes of detecting the FDIA, state reconstruction (SR) that aims at identifying and eliminating the compromised measurements as well as providing the most likely system operation states for operators to make decisions has not been involved yet. Developing a SR scheme can broaden and perfect the capabilities of SE, because it ensures the security assessment functions and corresponding corrective actions reliably implemented even the power system suffers the biased state values caused by FDIAs. A sparse state recovery method based on the mixed convex programming was developed in [20] to filter out additive measurement noise. A state forecasting-based method was proposed in [21] to detect FDIAs by considering nodal state temporal correlations. These methods address only the detection of anomaly states, without eliminating the adverse impacts of false positive and false negative. Moreover, in some rare cases, the removal of attacked measurements may cause the system to be unobservable, resulting in the failure of the traditional state estimator [21]. This problem has not been completely considered and addressed in the existing literature.

To this end, this paper proposed a novel SR scheme based on the extreme learning machine (ELM) that is a promising learning algorithm developed for training single-hidden layer feedforward neural networks (SLFNs). Unlike the conventional neural networks that are based on iterative learning algorithm, ELM randomly initializes the input weights and bias of hidden layer neurons and analytically determines the output weights via direct matrix computations [22]. Due to its extremely fast learning speed and excellent generalization capability, ELM has been successfully applied in many different application domains [22-24]. However, the randomness of input weights and bias may result in unstable and diverse results [25]. Ensemble learning has been used in many studies to address this issue through aggregating multiple base learners to boost the performance [25-27]. In this paper, the performance of the ensemble ELM is further promoted, and a novel approach is applied in SR scheme to filter out the impacts of FDIAs on hybrid SE, and hence enforce the cyber security of CPPS.

The main contributions of this paper are threefold: 1) Most of the existing FDIAs assume an approximate dc model associated to the SCADA measurements, which is not comprehensive and accurate when the PMU measurements are incorporated in the ac SE. This paper extends this model to a more general FDIA model which can effectively handle both the SCADA and PMU measurements in a hybrid ac state estimator. 2) An enhanced ensemble ELM (E^3LM) approach is developed for contaminated state separation that is the first step of the proposed SR scheme. In this approach, Gaussian random distribution (GRD) and Latin hypercube sampling (LHS) techniques are used for the initialization of input weights of base ELMs, which can increase the solution diversity of the base learners, and hence improve the generalization performance of ensemble learning. 3) To identify the exact locations of the compromised measurements and remove them all at once, an enhanced bad data identification method is proposed by exploiting the consistency between the forecasted and the received measurements. Then, a state recovery approach, combining quasi-Newton (QN) method and Armijo line search (ALS), reconstructs the system states and addresses the possible system unobservable problem due to the removal of attacked critical measurements.

The remainder of this paper is organized as follows. In Section II, the generic FDIA model of ac grids is presented. Section III elaborates the proposed SR scheme with detailed explanation on the architecture and implementation issues. Section IV demonstrates the numerical results on the tested power systems. Finally, we conclude this work in Section V.

II. Generic False Data Injection Attack Model

In this section, we review the existing models of FDIA and propose a generic FDIA model against the hybrid ac SE from the adversary's perspective.

A. Review of the FDIA Models

Liu *et al.* [28] demonstrated for the first time that in the case of a full power grid topology and parameter information, the adversaries can inject pre-designed false data into the SE without being detected by the traditional BDD procedure. Inspired by Liu's work, plenty of researches have been done to reveal the mechanisms of undetectable FDIAs and design effective attack strategies. Particularly, Liu *et al.* in [29] proved that the attacker does not have to get access to the information of the entire power network, but only needs to have the network information of the local area to launch such an undetectable FDIA in [28]. A heuristic algorithm was proposed in [30] to find an optimal attacking region which requires the reduced network information. A practical attack strategy against ac SE was developed in [31] based on a few measurements in the attacking region associated with boundary buses. Deng *et al.* in [32] proved that adversaries can launch FDIAs to modify the state variable on a bus only if they know the susceptance of

every transmission line that is incident to the target bus. The above FDIA strategies with incomplete network information further highlight the vulnerability of the CPPS.

FDIAs may have a catastrophic effect on the normal operation of power systems. For instance, Yuan *et al.* [33] considered a special FDIA, called load redistribution attack, which caused the power system to enter an uneconomic operating state with a wide load shedding. A cyber-attack strategy was proposed in [34] to overload transmission lines by considering the multiple solutions to security constrained economic dispatch (SCED). In order to carry out financial misconduct, an FDIA strategy was constructed in [35] to manipulate power flow estimates and to shift real-time locational marginal prices in a desired direction. An attack strategy was proposed in [36] to withhold generation capacity for profit by manipulating the ramp constraints of the generators during look-ahead dispatch. Mengis *et al.* in [37] proved that even if the network dynamics have limited uncertainty, the attack can still manipulate the nodal prices of the real-time markets without being detected. For multiphase and unbalanced smart distribution systems, the constructions of three-phase coupled, perfect three-phase decoupled, and imperfect three-phase decoupled FDIAs were proposed and discussed in [38]. Tan *et al.* in [39] proposed an attack strategy for automatic generation control (AGC) to cause frequency excursion that can trigger remedial actions, such as disconnection customer loads or generators, leading to blackouts, and potentially costly equipment damage.

B. Proposed Generic FDIA Model

To the best of our knowledge, none of the existing work can simultaneously overcome the following drawbacks of existing FDIA models: 1) To evade the BDD in control center, attack strategy should be designed to completely satisfy the underlying system model. However, in the literature, the approximately simplified and linearized dc model derived from the complex nonlinear power flow equations is widely used for FDIA construction, which is neither accurate nor general, thus making it likely for this kind of FDIAs to be detected by control center. 2) Existing related studies commonly assume that the attacker is kind of omnipotent and has all-encompassing knowledge of the network information, such as grid topology and system parameters. However, in practice, getting access to the complete network information for FDIA construction is expensive and unrealistic, because this information is generally kept confidential and highly protected in a control center. It is more realistic to consider attacks with incomplete network information. 3) Due to the increasing installation of PMU devices, the state estimator based on the pure SCADA measurements is gradually evolving towards a hybrid state estimator. From the adversaries' point of view, the FDIA model should be modified accordantly, otherwise control center will take advantage of PMUs to detect cyber-attacks. However, most existing FDIA models only associate with the SCADA measurements, which are not accurate and comprehensive when PMU measurements are included.

Consequently, in this subsection, we take all above practical issues into consideration and originally propose a more general nonlinear cyber-attack model with incomplete network information that can handle both SCADA and PMU measurements, according to the following equations:

$$\textit{Objective}: \qquad \min \left\| \boldsymbol{z}^{\mathfrak{V}_{\mathrm{A}}} - \boldsymbol{h}^{\mathfrak{V}_{\mathrm{A}}}\left(\boldsymbol{x}^{\mathrm{a}}\right) \right\|_0 \tag{1}$$

$$\textit{s.t.}: \qquad P_i + \Delta P_i^{\mathrm{a}} = V_i^{\mathrm{a}} \sum_{j \in \aleph_i} V_j^{\mathrm{a}} \left(G_{ij} \cos\theta_{ij}^{\mathrm{a}} + B_{ij} \sin\theta_{ij}^{\mathrm{a}} \right) \tag{2}$$

$$Q_i + \Delta Q_i^{\mathrm{a}} = V_i^{\mathrm{a}} \sum_{j \in \aleph_i} V_j^{\mathrm{a}} \left(G_{ij} \sin\theta_{ij}^{\mathrm{a}} - B_{ij} \cos\theta_{ij}^{\mathrm{a}} \right) \tag{3}$$

$$P_{ij} + \Delta P_{ij}^{\mathrm{a}} = \left(g_{\mathrm{sh},i} + g_{ij} \right) \left(V_i^{\mathrm{a}} \right)^2 - V_i^{\mathrm{a}} V_j^{\mathrm{a}} \left(g_{ij} \cos\theta_{ij}^{\mathrm{a}} + b_{ij} \sin\theta_{ij}^{\mathrm{a}} \right) \tag{4}$$

$$Q_{ij} + \Delta Q_{ij}^{\mathrm{a}} = -\left(b_{\mathrm{sh},i} + b_{ij} \right) \left(V_i^{\mathrm{a}} \right)^2 - V_i^{\mathrm{a}} V_j^{\mathrm{a}} \left(g_{ij} \sin\theta_{ij}^{\mathrm{a}} - b_{ij} \cos\theta_{ij}^{\mathrm{a}} \right) \tag{5}$$

$$V_i + \Delta V_i^{\mathrm{a}} = V_i^{\mathrm{a}}, \ \theta_i + \Delta\theta_i^{\mathrm{a}} = \theta_i^{\mathrm{a}} \tag{6}$$

$$\begin{aligned} I_{ij}^{\mathrm{re}} + \Delta I_{ij}^{\mathrm{re,a}} = {} & V_i^{\mathrm{a}} \left[\left(g_{\mathrm{sh},i} + g_{ij} \right) \cos\theta_i^{\mathrm{a}} - \left(b_{\mathrm{sh},i} + b_{ij} \right) \sin\theta_i^{\mathrm{a}} \right] \\ & - V_j^{a} \left[g_{ij} \cos\theta_j^{\mathrm{a}} - b_{ij} \sin\theta_j^{\mathrm{a}} \right] \end{aligned} \tag{7}$$

$$\begin{aligned} I_{ij}^{\mathrm{im}} + \Delta I_{ij}^{\mathrm{im,a}} = {} & V_i^{\mathrm{a}} \left[\left(g_{\mathrm{sh},i} + g_{ij} \right) \sin\theta_i^{\mathrm{a}} + \left(b_{\mathrm{sh},i} + b_{ij} \right) \cos\theta_i^{\mathrm{a}} \right] \\ & - V_j^{a} \left[g_{ij} \sin\theta_j^{\mathrm{a}} + b_{ij} \cos\theta_j^{\mathrm{a}} \right] \end{aligned} \tag{8}$$

$$V_i^{\min} \le V_i^{\mathrm{a}} \le V_i^{\max} \tag{9}$$

$$P_{Gk}^{\min} \le P_k + \Delta P_k^{\mathrm{a}} \le P_{Gk}^{\max}, \ Q_{Gk}^{\min} \le Q_k + \Delta Q_k^{\mathrm{a}} \le Q_{Gk}^{\max} \tag{10}$$

$$\sqrt{\left(P_l + \Delta P_l^{\mathrm{a}} \right)^2 + \left(Q_l + \Delta Q_l^{\mathrm{a}} \right)^2} \ge S_l^{\max} \tag{11}$$

$$x_e = x_e^0, \quad e \in \mathfrak{V}_{\mathrm{B}} \tag{12}$$

where $\boldsymbol{z}^{\mathfrak{V}_{\mathrm{A}}}$ is a measurement vector in the attacking region $\mathfrak{V}_{\mathrm{A}}$ that includes SCADA measurements (active/reactive power injection P_i and Q_i, active/reactive power flow P_{ij} and Q_{ij}) and PMU measurements (voltage phase angle θ_i, voltage magnitude V_i, real and imaginary parts of current phasor I_{ij}^{re} and I_{ij}^{im}). The superscripts "re" and "im" represent real and imaginary parts respectively; the subscripts "i" and "j" are the bus indexes in the attacking region. $\boldsymbol{h}^{\mathfrak{V}_{\mathrm{A}}}(\cdot)$ denotes the vector functions (2)-(8) that specify the relationships between the compromised measurements and state variables after attack $\boldsymbol{x}^{\mathrm{a}} = [\boldsymbol{\theta}^{\mathrm{a}}; \boldsymbol{V}^{\mathrm{a}}]$. θ_i^{a} and V_i^{a} are the ith element of $\boldsymbol{\theta}^{\mathrm{a}}$ and $\boldsymbol{V}^{\mathrm{a}}$ respectively. ΔP_i^{a}, ΔQ_i^{a}, $\Delta P_{ij}^{\mathrm{a}}$, $\Delta Q_{ij}^{\mathrm{a}}$, ΔV_i^{a}, $\Delta\theta_i^{\mathrm{a}}$, $\Delta I_{ij}^{\mathrm{re,a}}$ and $\Delta I_{ij}^{\mathrm{im,a}}$ are the incremental changes in corresponding measurements for FDIA construction. Moreover, G_{ij}+jB_{ij} is the ijth element of the complex bus admittance matrix; g_{ij}+jb_{ij} is the admittance of the series branch connecting buses i and j; $g_{\mathrm{sh},i}$+$jb_{\mathrm{sh},i}$ is the admittance of the shunt branch connected at bus i; $\aleph_i$ is the set of bus numbers that are directly connected to bus i. $V_i^{\min}$ and $V_i^{\max}$ are the lower and upper limits for V_i. $P_{Gk}^{\min}$, $P_{Gk}^{\max}$, $Q_{Gk}^{\min}$ and $Q_{Gk}^{\max}$ are the kth generator's active and reactive power capacity limits, $k \in \mathfrak{V}_{\mathrm{G}}$, $\mathfrak{V}_{\mathrm{G}}$ is the set of generators. $S_l^{\max}$ is the apparent power limits for the lth branch, $l \in \mathfrak{V}_{\mathrm{L}}$, $\mathfrak{V}_{\mathrm{L}}$ is the set of transmission lines. $\mathfrak{V}_{\mathrm{B}}$ is the set of boundary buses on the edge of the attack region. x_e is the state variable of the eth boundary bus, and x_e^0 denotes its initialized state value.

Objective function (1) aims at minimizing the total number of nonzero elements in the attack vector and so the attack vector exhibits high sparsity. SCADA measurement equations (2)-(5) and PMU measurement equations (6)-(8) describe the ac grid model subject to bus voltage operational limits (9) and generator capacity limits (10). To make the FDIA more realistic, we

assume that the attacker alters the incremental changes in SCADA and PMU measurements to launch an FDIA by overloading transmission lines as shown in (11) [40]. Moreover, (12) is used to guarantee that the values of boundary bus states consisting of voltage magnitudes and phase angles remain unchanged before and after the attack implementation [30].

According to the generic FDIA model (1)-(12), the attacker first performs local OPF to estimate the state information of the attack region, and then constructs optimal attack vector by solving the L0-norm minimization problem. In this model, the information required for FDIA construction is within the attack region. Furthermore, there is no special requirement for the attack region, so the attacker is free to choose any subnetwork with accessible network topology and parameters as the attack region. In addition, this generic model effectively integrates the SCADA and PMU measurements, and thus the control center cannot use the PMU measurements to directly verify the SE results. It is clear that the proposed generic FDIA model with limited network information exhibits higher practicability than the existing attack models, and further reduces the chance of being detected. After the attack vector has been implemented, the system operators will find the target lines are overloaded, which may result in a cascading failure of the target CPPS if there are no countermeasures against this kind of cyber-attack.

III. Proposed State Reconstruction Scheme

Traditional weighted least square (TWLS) method-based hybrid SE provides updated snapshots of system operating states to several crucial applications in EMS. However, it fails to diagnose the well-coordinated bad measurements from stealthy cyber-attack, which causes biased states and thus severely threatens the security of the power system. Therefore, a SR scheme, consisting of three steps, i.e., anomaly states separation, enhanced BDD and state recovery, is for the first time proposed in this paper to overcome the deficiencies of TWLS.

A. E³LM-Based Contaminated State Separation

In this section, ELM is applied for classification task to identify the relationship between input and output variables. For a training dataset with N total distinct instances $\{(\boldsymbol{v}_i, \boldsymbol{o}_i)\}_{i=1}^{N}$, where $\boldsymbol{v}_i \in \mathbb{R}^d$ with $\boldsymbol{v}_i = [v_{i1}, v_{i2}, \cdots, v_{id}]^T$ and $\boldsymbol{o}_i \in \mathbb{R}^c$ with $\boldsymbol{o}_i = [o_{i1}, o_{i2}, \cdots, o_{ic}]^T$, ELM with K hidden layer nodes and an active function ϑ can be mathematically represented as

$$\sum_{i=1}^{K} \boldsymbol{\beta}_i \vartheta\left(\boldsymbol{w}_i \cdot \boldsymbol{v}_j + \eta_i\right) = \boldsymbol{o}_j, \qquad j = 1, \cdots, N \tag{13}$$

where $\boldsymbol{w}_i = [w_{i1}, w_{i2}, \cdots, w_{id}]^T$ is the weight vector connecting the ith hidden node and the input nodes, $\boldsymbol{\beta}_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{ic}]^T$ is the weight vector connecting the ith hidden node and the output nodes, and η_i is the bias of the ith hidden node. The above N equations can be rewritten as

$$\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{O} \tag{14}$$

$$\boldsymbol{A} = \begin{bmatrix} \vartheta\left(\boldsymbol{w}_1 \cdot \boldsymbol{v}_1 + \eta_1\right) & \cdots & \vartheta\left(\boldsymbol{w}_K \cdot \boldsymbol{v}_1 + \eta_K\right) \\ \vdots & \vdots & \vdots \\ \vartheta\left(\boldsymbol{w}_1 \cdot \boldsymbol{v}_N + \eta_1\right) & \cdots & \vartheta\left(\boldsymbol{w}_K \cdot \boldsymbol{v}_N + \eta_K\right) \end{bmatrix} \tag{15}$$

where $\boldsymbol{A}$ is the hidden-layer output matrix of the ELM, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_K]^T$ is the output weight matrix, and $\boldsymbol{O} = [\boldsymbol{o}_1, \boldsymbol{o}_2, \cdots, \boldsymbol{o}_N]^T$ is the matrix of targets. Moreover, the ELM learning process is summarized as follows: 1) the input weights $\boldsymbol{w}_i$ and the biases η_i are randomly generated and not necessarily tuned; 2) the hidden-layer output matrix $\boldsymbol{A}$ is calculated according (15); 3) the output weights matrix is derived as $\boldsymbol{\beta} = \boldsymbol{A}^{\dagger}\boldsymbol{O}$, where $\boldsymbol{A}^{\dagger}$ is the Moor-Penrose generalized inverse of $\boldsymbol{A}$ and can be obtained by using the singular value decomposition method [41].

The randomness in ELM parameters make individual ELMs suffering from degradation of consistency and robustness, but on the other hand provides a good opportunity for designing ensemble models since the randomness can inherently increase the diversity of an ensemble, and hence significantly improve the classification accuracy [42]. Fig. 1 schematically illustrates the proposed framework of a tailored ELM ensemble classifier for contaminated state separation. This classifier consists of E individual ELMs that tend to compensate for each other and thus increase accuracy over the individual.

To improve the performance of ensemble learning, one of the most important techniques is to generate data diversity [43]. In the base ensemble ELM (BE²LM) approach, the input weights are initialized by assigning uniform random values from -1 to 1. However, such initialization method restricts the weight values into a certain narrow range, leading to a poor data diversity. To enhance the performance of contaminated state separation, an E³LM approach with the assist of GRD and LSH techniques for weight initialization is proposed in this work. Specifically, each weight value is sampled from the GRD with zero mean and unit variance, which can map the inputs to a random space with more diversity than the weight initialization adopted in BE²LM. Moreover, LHS is used to further increase the diversity of base learners. As a stratified-random procedure, LHS provides an efficient way of sampling variables from their distributions [44] and its implementation can be summarized as follows: 1) divide the cumulative distribution of each weight variable into E equiprobable intervals; 2) randomly select a value from each interval and represent the sampled cumulative probability for the ith interval as $\text{Prob}_i = (r_u + i - 1)/E$, where r_u is a random number ranging from 0 to 1; 3) transform the probability values into the weight value using the inverse of the distribution function F^{-1}, i.e., $w = F^{-1}(\text{Prob})$; 4) repeat the above procedures for each weight variables and randomly pair the obtained weight values to form E input weight matrices. LHS ensures a full coverage of the range of each weight variable by maximally stratifying each marginal distribution. Therefore, compared to BE²LM, E³LM can achieve a better generalization performance and provide more diversity for ensemble learning.

As shown in Fig. 1, an initial database with N total distinct instances $\{(\boldsymbol{v}_i, \boldsymbol{o}_i)\}_{i=1}^{N}$ is first derived by generating measurements with/without FDIA. Each ELM takes the measurement vector consisted of SCADA and PMU measurements as input $\boldsymbol{v}_i$. The corresponding desired target vector $\boldsymbol{o}_i$ is composed of c=2n-1 elements, where n denotes the bus number. And two decision boundaries for the "normal state" (represented by +1) and "contaminated state" (represented by -1) are defined. The normal or contaminated state is identified and flagged based on the result of comparing the estimated states obtained from SE

before and after FDIA. Specially, if the input $\boldsymbol{v}_i$ is not manipulated by an attacker, the corresponding desired target $\boldsymbol{o}_i$ is a zero vector with dimension of 2n-1. The whole dataset is randomly divided into a training set and a testing set. The training stage aims to identify the optimal learning parameter set that matches the input and target relationship presented in the training dataset. E^{3}LM initializes its weight matrices using GRD and LHS and each ELM randomly selects bias vector, a subset of training data, and a candidate activation function. The parameters of each ELM can be obtained after ELM learning process. For online separation of contaminated states, the previously trained parameters are utilized to establish the mathematical relationship between measurements and the flag vector indicating the normal and contaminated states. The output of each ELM is a vector with the same dimension as the target vector. And the classification rule of each single ELM is shown as follows:

$$\begin{cases} \boldsymbol{y}_i(j) \ge 0 \rightarrow \boldsymbol{y}_i(j) = 1 \\ \boldsymbol{y}_i(j) < 0 \rightarrow \boldsymbol{y}_i(j) = -1 \end{cases} \tag{16}$$

where $\boldsymbol{y}_i(j)$ is the jth element of the ith ELM unit sub-output, i=1,2,…,E, j=1,2,…,2n-1. Suppose there are ζ "+1" sub-outputs and μ "-1" sub-outputs for the jth state (ζ+μ=E), a decision-making mechanism is strategically designed to evaluate all the sub-outputs of the ELM ensemble classification and determine the final classification result, shown as follows:

$$\begin{cases} \zeta > \mu \rightarrow \boldsymbol{Y}(j) = 1 \\ \zeta \le \mu \rightarrow \boldsymbol{Y}(j) = -1 \end{cases} \tag{17}$$

where $\boldsymbol{Y}(j)$ is the final classification result for the jth state. With this learning rule, the accuracy of contaminated state identifier can be improved due to the generalized randomness of individual ELM and the extended diversity derived from ensemble learning as well as the proposed weight initialization method.

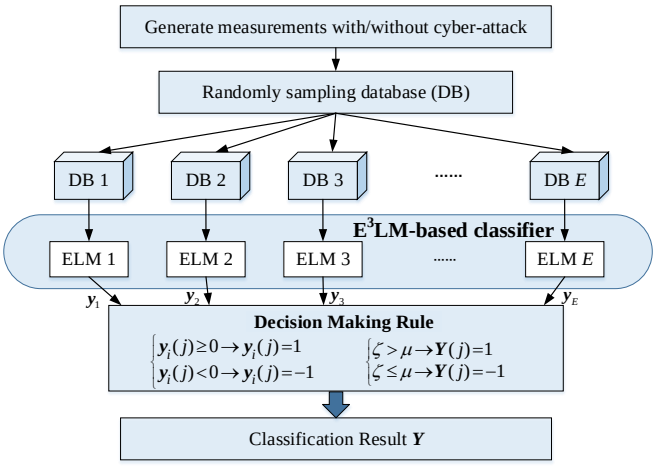


Fig. 1. Proposed E^{3}LM-based classifier for contaminated state separation.

B. Enhanced Bad Measurements Identification

To address the misdeclaration and missing alarm issues in previous procedure, we first replace these identified anomalies by their corresponding forecasted values. There are two approaches available in the literature to predict the system state for the next time step. One is to forecast system state directly by extracting the temporal and spatial relationships of system dynamics from the previous system states [45]. The other is to forecast the bus load and then obtain the system state via power flow analysis [46]. The simulation results confirmed the effectiveness and accuracy of both prediction solutions [45, 46].

In this paper, some sophisticated load forecasting method described in [46] and [47] can be directly applied to generate the active power increment $\Delta\hat{P}_{i.t+1}$ of bus i, between time t and t+1 in (18). Reactive power $\hat{Q}_{i.t+1}$ is changed accordingly to keep original power factor constant at the ith bus.

$$\hat{P}_{i,t+1} = \hat{P}_{i,t} + \Delta\hat{P}_{i.t+1} \tag{18}$$

where $\hat{P}_{i,t}$ is the forecasted active power at bus i and time t.

The forecasted loads, model parameters, network topologies, and generation schedules are collected and used to perform an economic dispatch by applying the interior point method. Then, the power-flow equations are iteratively solved using the Newton-Raphson method to convert forecasted load to forecasted state, i.e., voltage magnitudes and phase angles. Afterwards, the contaminated states identified in Section III-A are replaced by their corresponding forecasted values obtained from the above method to form a new operating state vector $\boldsymbol{x}_t^{\text{new}}$. To identify the compromised measurements caused by FDIAs all at once, $\boldsymbol{x}_t^{\text{new}}$ is taken into $\boldsymbol{h}(\cdot)$ to construct a new set of measurements $\boldsymbol{z}_t^{\text{new}}$, and BDD is implemented again, shown as

$$\gamma_t(i) = \left|\boldsymbol{z}_t(i) - \boldsymbol{z}_t^{\text{new}}(i)\right| \Big/ \sqrt{\boldsymbol{\Omega}(i,i)}, \quad i = 1,\cdots,m \tag{19}$$

where $\boldsymbol{z}_t$ is the measurement vector at time t, m denotes the measurement number, $\boldsymbol{\Omega} = \boldsymbol{R} - \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T$ is the residual covariance matrix, $\boldsymbol{R}$ is the measurement error covariance matrix, $\boldsymbol{H}$ and $\boldsymbol{G}$ are the Jacobian and gain matrices of TWLS-based SE, respectively. Then, a flag vector $\boldsymbol{\varphi}$ with the same dimension as γ_t is defined. Specifically, the element $\boldsymbol{\varphi}(i)$ is set to 1 when $\gamma_t(i)$ is larger than a given threshold, indicating that $\boldsymbol{z}_t(i)$ is a bad data and cannot be trusted; otherwise, it is set to 0.

Before state replacement, bad measurements due to FDIAs cannot be detected by BDD, since they are well coordinated and satisfy the Kirchhoff's circuit laws. Once some contaminated states are identified and replaced by their forecasts, the well-constructed relationship between states and measurements in FDIAs are damaged, rendering high residual values in BDD. Moreover, because the attack vector generally exhibits high-sparsity [13] and $\boldsymbol{z}_t^{\text{new}}$ cannot be viewed as the measurements before attack, the very small amount of untrusted measurements with flag 1 in $\boldsymbol{\varphi}$ are removed directly instead of replaced by the corresponding derivative elements in $\boldsymbol{z}_t^{\text{new}}$.

C. State Recovery with QN Method

Compared with TWLS-based SE which begins from the flat start condition, state recovery method uses forecasted states obtained from forecasted system loads with power flow analysis as the initial guess. Once the reduced measurement vector $\breve{\boldsymbol{z}}$ becomes available after removing the bad data in $\boldsymbol{z}$, state recovery is formulated by minimizing the following objective.

$$\begin{aligned} &\min \quad J(\boldsymbol{x}) = \frac{1}{2}\left[\breve{\boldsymbol{z}} - \boldsymbol{h}(\boldsymbol{x})\right]^T \boldsymbol{R}^{-1}\left[\breve{\boldsymbol{z}} - \boldsymbol{h}(\boldsymbol{x})\right] = \frac{1}{2}\boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} \\ &s.t.: \quad \breve{\boldsymbol{z}} = h(\boldsymbol{x}) + \boldsymbol{\varepsilon} \end{aligned} \tag{20}$$

where $\boldsymbol{h}(\boldsymbol{x})$ denotes the functions relating the reduced measurement vector $\breve{\boldsymbol{z}}$ to the state vector $\boldsymbol{x}$, $\boldsymbol{\varepsilon}$ is the vector of measurement errors where $\boldsymbol{\varepsilon}_i \sim N(0, \sigma_i)$, $\boldsymbol{r}$ denotes the estimation residual. $\boldsymbol{R}$ represents the measurement error covariance and its reciprocal is regarded as weights in (20).

At the minimum, the first-order optimality conditions will have to be satisfied, which can be expressed as follows:

$$\boldsymbol{J}'(\boldsymbol{x}) = -\boldsymbol{H}^T \boldsymbol{R}^{-1}\left[\breve{\boldsymbol{z}} - \boldsymbol{h}(\boldsymbol{x})\right] = -\sum_i \frac{r_i}{\sigma_i^2}\frac{\partial h_i(\boldsymbol{x})}{\partial \boldsymbol{x}} = 0 \tag{21}$$

where $\boldsymbol{H}$ is the Jacobian matrix of $\boldsymbol{h}(\boldsymbol{x})$. Expanding the above non-linear function $\boldsymbol{J}'(\boldsymbol{x})$ into its Taylor series around the state vector $\boldsymbol{x}$ by omitting higher order terms yields:

$$\boldsymbol{J}'(\boldsymbol{x})\Big|_{(x+\Delta x)} = \boldsymbol{J}'(\boldsymbol{x})\Big|_{(x)} + \boldsymbol{J}''(\boldsymbol{x})\Big|_{(x)} \cdot \Delta \boldsymbol{x} = 0 \tag{22}$$

where $\boldsymbol{J}''(\boldsymbol{x})$ is the Hessian matrix and can be expressed as:

$$\begin{aligned}\boldsymbol{J}''(\boldsymbol{x}) &= \sum_i \frac{1}{\sigma_i^2}\frac{\partial h_i(\boldsymbol{x})}{\partial \boldsymbol{x}}\left(\frac{\partial h_i(\boldsymbol{x})}{\partial \boldsymbol{x}}\right)^T - \sum_i \frac{r_i}{\sigma_i^2}\frac{\partial^2 h_i(\boldsymbol{x})}{\partial \boldsymbol{x}^2} \\ &= \boldsymbol{H}^T \boldsymbol{R}^{-1}\boldsymbol{H} - \sum_i \frac{r_i}{\sigma_i^2}\frac{\partial^2 h_i(\boldsymbol{x})}{\partial \boldsymbol{x}^2}\end{aligned} \tag{23}$$

TWLS method adopts gain matrix $\boldsymbol{G} = \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}$ to replace the Hessian matrix $\boldsymbol{J}''(\boldsymbol{x})$ in (22) for simplicity, which inevitably introduces inaccuracy into state recovery when power system nonlinearity coincides with sudden load changes [48]. In addition, measurement design is generally implemented to maintain a certain level of reliability against branch outages or loss of measurements. However, a possible singular $\boldsymbol{G}$ may occur and TWLS method fails to find the solution of (20) in case of some critical measurements are identified as the bad data and deleted in the previous sub-section.

This paper adopts an effective state recovery approach based on QN method which uses a symmetric positive definite matrix $\boldsymbol{D}$ to approximate the inverse the Hessian matrix $(\boldsymbol{J}''(\boldsymbol{x}))^{-1}$ at each iteration. Currently, the broden-Flecher-Goldfarb-Shanno (BFGS) formula [49] is the most widely used QN method due to its great performance for low accuracy line searches [48]. With the help of QN method, matrix inverse calculation for large power systems and a possible singular of gain matrix can be avoided because matrix $\boldsymbol{D}$ is proved to be always positive definite to guarantee a decent search direction. Firstly, a search direction $\boldsymbol{s}_q$ at the qth iteration is determined as:

$$\boldsymbol{s}_q = -\boldsymbol{D}_q \cdot J'(\boldsymbol{x}_q) \tag{24}$$

where $\boldsymbol{D}_1$ is a unit matrix. ALS is implemented to find the smallest nonnegative integer u such that $\kappa_q = \lambda^u$ satisfies:

$$J(\boldsymbol{x}_q + \kappa_q \boldsymbol{s}_q) \le J(\boldsymbol{x}_q) + a_1 \kappa_q (\boldsymbol{s}_q)^T \boldsymbol{J}'(\boldsymbol{x}_q) \tag{25}$$

$$\left|\boldsymbol{J}'(\boldsymbol{x}_q + \kappa_q \boldsymbol{s}_q)^T \cdot \boldsymbol{s}_q\right| \le -a_2 (\boldsymbol{s}_q)^T \boldsymbol{J}'(\boldsymbol{x}_q) \tag{26}$$

where λ, a_1 and a_2 are constants, $0<\lambda<1$, $0<a_1<a_2<1$. The above criteria are called strong Wolfe conditions ensuring the step length κ decreases $\boldsymbol{J}(\boldsymbol{x})$ and $\boldsymbol{J}'(\boldsymbol{x})$ sufficiently to accelerate convergence. The state vector and matrix $\boldsymbol{D}$ can be updated as:

$$\boldsymbol{x}_{q+1} - \boldsymbol{x}_q = \kappa_q \boldsymbol{s}_q \tag{27}$$

$$\boldsymbol{D}_{q+1} - \boldsymbol{D}_q = \left(1 + \frac{(\boldsymbol{p}_q)^T \boldsymbol{D}_q \boldsymbol{p}_q}{(\boldsymbol{l}_q)^T \boldsymbol{p}_q}\right)\frac{\boldsymbol{l}_q(\boldsymbol{l}_q)^T}{(\boldsymbol{l}_q)^T \boldsymbol{p}_q} - \frac{\boldsymbol{l}_q(\boldsymbol{p}_q)^T \boldsymbol{D}_q + \boldsymbol{D}_q \boldsymbol{p}_q (\boldsymbol{l}_q)^T}{(\boldsymbol{l}_q)^T \boldsymbol{p}_q} \tag{28}$$

where $\boldsymbol{l}_q = \boldsymbol{x}_{q+1} - \boldsymbol{x}_q$ and $\boldsymbol{p}_q = \boldsymbol{J}'(\boldsymbol{x}_{q+1}) - \boldsymbol{J}'(\boldsymbol{x}_q)$.

D. Summary of the Whole State Reconstruction Scheme

The whole procedure of SR is described in Fig. 2. Note that the BDD procedure is contained in the traditional hybrid SE, so the bad data caused by sampling and communication errors have been detected and deleted. After these conventional SE processes, the proposed scheme further analyzes the measurements and the estimated system states. First, the anomaly state separation method applies the well-trained ELM ensemble classifier to recognize the contaminated states. In step 2, the estimated $\tilde{P}_{i,t}$ and $\tilde{Q}_{i,t}$ are calculated by using the estimated states $\tilde{\boldsymbol{x}}_t^{\mathrm{SR}}$ (if there is a cyber-attack at time t) or $\tilde{\boldsymbol{x}}_t^{\mathrm{SE}}$ (if there is no cyber-attack at time t). Then, $\tilde{P}_{i,t}$ and $\tilde{Q}_{i,t}$ are regarded as the forecasted $\hat{P}_{i,t}$ and $\hat{Q}_{i,t}$ when entering the load forecasting of the next time slot in (18). The system states obtained via power flow analysis are applied to replace the detected anomaly states and regarded as the initial state values before state recovery in step 3. Finally, the reduced measurement vector is constructed and sent to QN method with ALS to reconstruct system states, and thus to complete the SR of the whole system.

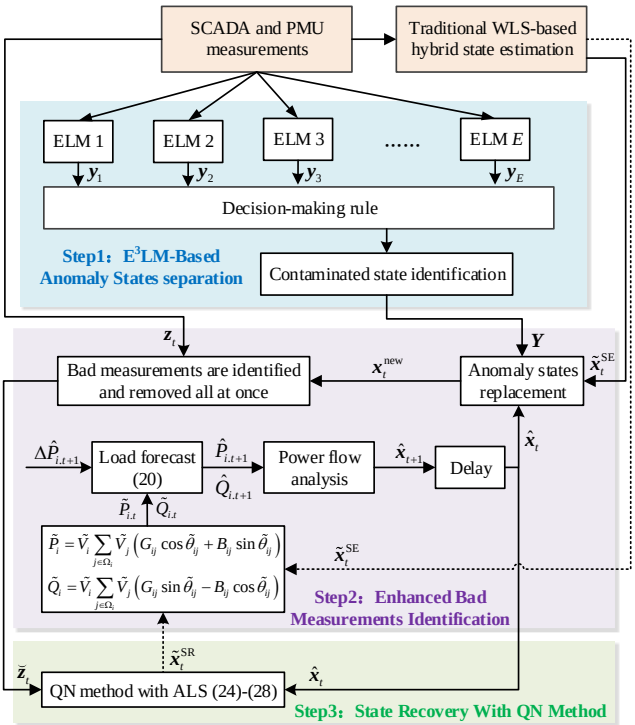


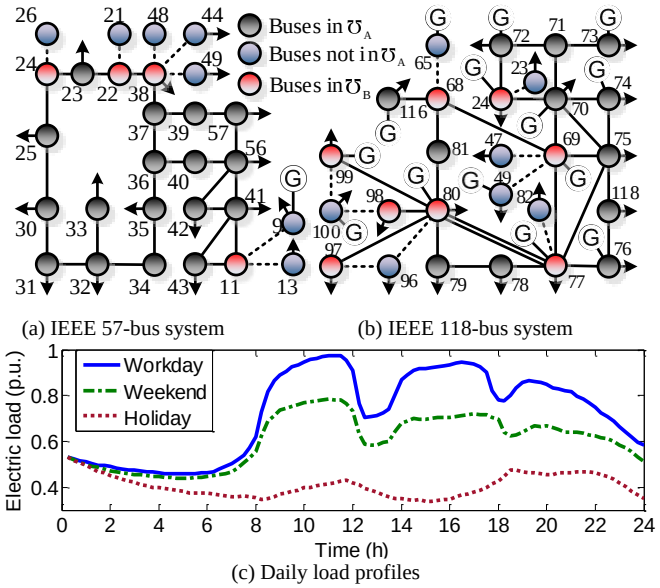
Fig. 2. Structure of the proposed state reconstruction scheme.

IV. Case Studies

In this section, the feasibility and effectiveness of the generic FDIA model and SR mechanism have been extensively tested and benchmarked on the IEEE 14-, 57-, and 118-bus power systems with parameters and network topologies from [50].

A. Investigation on Generic FDIA

1) Experimental settings: The FDIA with complete network information is a special case of proposed generic attack model. In this section, the proposed attack model with complete network information is tested on the IEEE 14-bus system, and the attack model with incomplete information is investigated on the IEEE 57- and 118-bus systems. In these test systems, it is assumed that power injections are taken at all buses, and power flows are taken across all branches ("from" terminal only). Moreover, the PMU placement configurations and parameters adopted in this study are given in [5]. The attacking regions in the 57- and 118-bus systems were arbitrarily chosen and shown in Fig. 3(a) and (b) respectively. We assumed that all load points in the test systems followed a typical daily load curve, e.g., curves from a workday (October 20), a weekend (February 7) and a holiday (January 1), with a 15-min resolution, as depicted in Fig. 3(c). The three daily load curves were collected from Dongguan dispatch center in China 2016. Moreover, the L0-norm minimization is NP-hard and thus very difficult to be solved, therefore, the objective (1) is relaxed to L1-norm minimization and then solved for sparse attack construction with YALMIP [13]. In addition, the attack vector obtained from (1)-(12) may contain many small-valued nonzero elements that were treated as zeros if their values are within the variation tolerances of measurement noise. We implemented the proposed generic FDIA with MATLAB R2014a on a PC with an i7-6700 4.0 GHz CPU and 16 GB of RAM.



(a) IEEE 57-bus system (b) IEEE 118-bus system

(c) Daily load profiles

Fig. 3. Incomplete network topologies for power systems and load profiles.

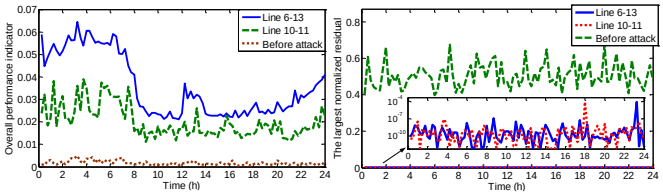
2) Numerical results and analysis: A series of simulations are carried out to demonstrate the feasibility of the proposed attack model. For the IEEE 14-bus system, the attack strategy is designed to overload lines 6-13 or 10-11 under the workday. The overall performance indicator at each time slot, calculated as (29), and the largest normalized residuals before and after attacks are presented in Fig.4.

$$\psi_t = \frac{1}{n-1}\sum_{i=2}^{n}\left|\frac{\tilde{\theta}_{i,t}-\theta_{i,t}^{\text{real}}}{\theta_{i,t}^{\text{real}}}\right| + \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\tilde{V}_{i,t}-V_{i,t}^{\text{real}}}{V_{i,t}^{\text{real}}}\right| \quad (29)$$

where $\theta_{i,t}^{\text{real}}$ and $V_{i,t}^{\text{real}}$ are the real phase angel and voltage magnitude at bus i, time t respectively; $\tilde{\theta}_{i,t}$ and $\tilde{V}_{i,t}$ denote their estimated values before or after attack.

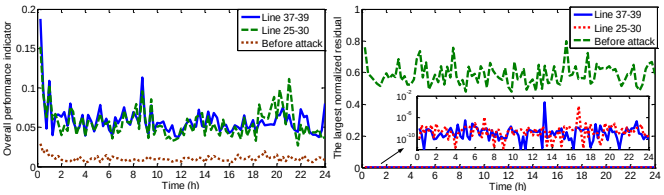
As shown in Fig. 4(a), the SE errors after attacks are much larger than that without attacks if there is no defense mechanism. The biased states may mislead system operator to make false decisions, which may cause catastrophic consequences in power system. In Fig. 4(b), all the largest normalized residuals, no matter it is the residual before or after an attack, are below the threshold of the largest normalized residual test that is generally chosen as 3 for 99.7% confidence level. This implies that the well-constructed FDIAs can effectively bypass BDD in ac SE even the PMUs are installed in the target power system, because the proposed generic attack model can handle both SCADA and PMU measurements and make them satisfying Kirchhoff's circuit laws, rendering all residual-based attack detection methods invalid.

For the IEEE 57-bus system, the FDIAs are launched under the holiday to overload lines 37-39 or 25-30. The SE errors and corresponding residuals are illustrated in Fig. 5. In order to demonstrate the scalability of the proposed attack strategy, we also implement the FDIAs on the IEEE 118-bus system on the weekend to overload the 71-73 or 70-75 lines. The corresponding results are presented in Fig. 6. As expected, these constructed FDIAs could stealthily circumvent traditional residual test and significantly contaminate system states, even if the attacker can only access limited local network information.



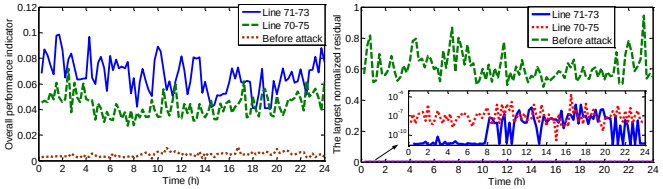
(a) Overall performance indicator ψ_t. (b) The largest normalized residuals

Fig. 4. Daily performance and residuals before and after the attack for the IEEE 14-bus system under a workday.



(a) Overall performance indicator ψ_t. (b) The largest normalized residuals

Fig. 5. Daily performance and residuals before and after the attack for the IEEE 57-bus system under a holiday.



(a) Overall performance indicator ψ_t. (b) The largest normalized residuals

Fig. 6. Daily performance and residuals before and after the attack for the IEEE 118-bus system under a weekend.

The above FDIAs are designed to overload a single line, which may be of very limited threat to power system due to the N-1 principle. Table I lists the statistical results of ψ_t (averaged over 10 FDIAs with different operational scenarios and target lines for a given number of overloaded lines) in IEEE 14-, 57-, and 118-bus systems to demonstrate the extensive applications of the proposed attack model for overloading multiple lines. It shows that the estimated error increases greatly with the number of overloaded lines, which means the more the overloaded lines, the larger will be the threaten of FDIA. However, it is harder to launch a stealthy attack with more overloaded lines due to the reduced sparsity of attack vector. Thus, the attacker should carefully choose the number of overloaded lines to make sure a stealthy FDIA with effective threaten. Moreover, we checked all the residuals for the FDIAs in Table I and found that all attacks would bypass the traditional residual test, proving the feasibility of the proposed attack strategy.

TABLE I
THE AVERAGE VALUES OF ψ_t FOR OVERLOADING MULTIPLE LINES

Systems	Number of attacked transmission lines				
	1	2	3	4	5
14-bus	0.052	0.061	0.086	0.222	0.446
57-bus	0.056	0.143	0.162	0.191	0.235
118-bus	0.072	0.165	0.192	0.235	0.379

In addition, to evaluate the computational efficiency and complexity of our attack strategy, Table II illustrates the statistics of the elapsed time required for solving (1)-(12). For each test system, ten independent FDIAs are considered. The elapsed time is obtained by calling the solvertime function in YALMIP. Table II shows that it takes an average time of 1.08s with a minimum of 0.79s and a maximum of 1.67s to solve the proposed attack strategy implemented in IEEE 14-bus system. The time expenses for attack construction in IEEE 57- and 118-bus systems amount to 5.68s and 16.63s, respectively. It can be seen that the required time for launching an attack is extended when system gets larger, since there are more measurements and constraints for attack formulation in more complex grids. Moreover, it is obvious that the attack construction can be finished in an extremely short period of time. This is because the NP-hard L0-norm optimization is relaxed into a L1-norm problem, which can be efficiently solved.

TABLE II
THE ELAPSED TIME STATISTICS FOR SOLVING FDIA MODEL

Systems	Max	Min	Average
14-bus	1.67s	0.79s	1.08s
57-bus	7.37s	3.87s	5.68s
118-bus	26.13s	10.45s	16.63s

B. Investigation on State Reconstruction

1) Experimental settings: The developed SR scheme is tested on IEEE 14-, 57-, and 118-bus systems. For each test system, 2480 instances are generated and used to train and test the ELM-based classifier. First, the system generation/load patterns are randomly varied within 80~120% level of the nominal value. As a result, 480 base operating conditions are generated. For each operating condition, the real system states are obtained from the solution of power flow analysis and the accurate SCADA and PMU measurements can be calculated by using $\boldsymbol{h}(\boldsymbol{x})$. The real measurements, constituted by adding Gaussian measurement errors to all the accurate measurements, are regarded as the input of ELM ensemble classifier. And the corresponding desired output is a zero vector with dimension of $2n$-1 because there is no contaminated state derived from hybrid state estimator without FDIA. Besides the above 480 normal operational cases, another set of 2000 attack-cases are constructed by applying the proposed generic FDIA model. For each operating condition, an FDIA is constructed by randomly choosing the attacking region and the overloaded lines. And the desired output vector can be derived based on the result of comparing the estimated states obtained from SE before and after FDIA. For cross validation, the total instances in the three test systems are randomly divided into a training set and a testing set by 3:1 ratio. To verify the effectiveness of the proposed E^{3}LM approach, the comparative study is carried out against BE^{2}LM and the performance is evaluated in terms of classification accuracy, training and detection time. For both approaches, 200 ELMs are employed as base learners for the ensemble learning. In addition, 90% of the total training instances are randomly chosen to train each ELM. The activation function and the hidden node number can be optimally determined by a pre-tuning procedure [41]. The forecasted load increment is assumed to follow normal distribution with both mean and magnitude of variance equal to $\Delta\hat{P}_{i,t+1}$ [48]. Other settings were the same as in section IV-A.

2) Numerical results and analysis: To demonstrate the feasibility of the proposed E^{3}LM-based classifier, three different test systems are assessed, and the numerical results are illustrated in Table III. It is obvious that E^{3}LM performs much better than BE^{2}LM in terms of classification accuracy, indicating that the proposed weight initialization with the assist of GRD and LHS techniques can create more diversity for boosting the performance of ensemble learning. The instances that can be classified by the proposed contaminated state separation method with 100% accuracy take up 86.45%, 83.71%, and 81.94% of the total instances in the testing set for IEEE 14-, 57-, and 118-bus systems, respectively. For most of the remaining instances, i.e., about 80 out of 84 instances in the 14-bus system, 95 out of 101 instances in the 57-bus system, and 105 out of 112 instances in the 118-bus system, the classification accuracy is more than 80%. These results demonstrate the satisfactory generalization capability of the proposed ELM ensemble classifier.

In addition, the training time (TT) and average detection time (ADT) of both approaches are also presented in Table III. Due to the high computational efficiency of the ELM learning process, the training and testing processes of both ensemble classifiers can be completed very quickly for three test systems. Compared to BE^{2}LM, E^{3}LM consumes more time for training the ensemble model due to the adopted novel weight initialization. Since the weight initialization is performed offline in the training stage, there is no significant difference in terms of ADT for two evaluated methods. The computational complexity of the detection process can be calculated as $O(EdKcM)$, where E, d, K, c, and M denote the number of base ELMs for ensemble learning, the input dimension of base ELMs (number of measurements), the number of hidden nodes for base ELMs,

the output dimension, and the number of types of each system state for classification, respectively. For these three test systems, the values of E and M are set to be 200 and 2 respectively. The values of (d, c) are set to (90, 27), (372, 113) and (726, 235) for IEEE 14-, 57- and 118-bus systems, respectively. In addition, the hidden node number, K, is set as 200, 300 and 400 for IEEE 14-, 57- and 118-bus systems, respectively. Therefore, the computational complexity of the detection process is considered very low, which guarantees the scalability of the algorithm for large-scale systems.

TABLE III
PERFORMANCE OF CONTAMINATED STATE SEPARATION METHOD

System		14-bus		57-bus		118-bus	
Approach		BE^{2}LM	E^{3}LM	BE^{2}LM	E^{3}LM	BE^{2}LM	E^{3}LM
CA*	100%	85.48%	86.45%	82.26%	83.71%	80.32%	81.94%
	Acc1	10.65%	10.48%	12.26%	12.58%	13.39%	13.87%
	Acc2	2.90%	2.42%	3.87%	2.74%	4.35%	3.06%
	Acc3	0.97%	0.65%	1.61%	0.97%	1.94%	1.13%
	Acc4	0	0	0	0	0	0
TT (s)		6.74	7.39	15.26	24.35	27.20	42.58
ADT (ms)		1.47	1.49	5.78	5.96	9.77	9.64

*Classification Accuracy; Acc1∈[90%, 100%); Acc2∈[80%, 90%); Acc3∈[70%, 80%); Acc4∈[0, 70%)

For the worst case in the testing set for IEEE 14-bus system, i.e., 77.8% classification accuracy (five false positives and one false negative), the absolute incremental changes in measurements for FDIA construction and the flag vector $\boldsymbol{\varphi}$ obtained by the proposed bad data identification method are plotted in Fig. 7, in which the elements of $\boldsymbol{\varphi}$ with value 1 are denoted by the gray bar. It shows that the undiscovered contaminated state (false negative) has no influence on the bad data identification. This is because the well-constructed relationships between states and measurements in FDIAs are damaged once partial contaminated states are detected by E^{3}LM-based classifier, rendering high residual values. Moreover, it can be observed that the false positives during contaminated state separation process lead to the deletion of very few un-attacked measurements, which thanks to the acceptable forecasting accuracy of system states.

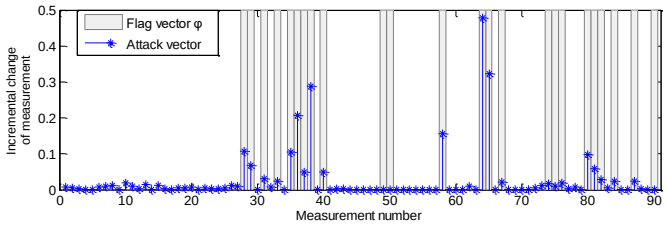
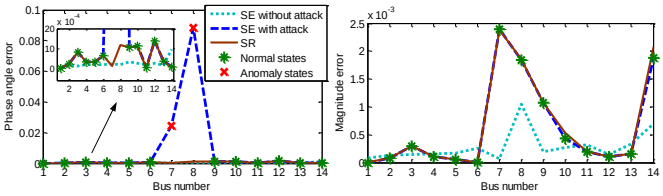


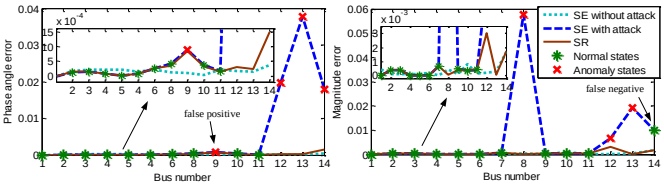
Fig. 7. Performance of enhanced BDD for the worst case in 14-bus system.

Considering the IEEE 14-bus system, a stealthy attack with complete network information was launched at 5:00 AM to overload line 7-8. The absolute SR and SE errors are shown in Fig. 8, in which red multiplication and green asterisk symbols represent the anomaly and normal states identified by ELM ensemble classifier respectively. It is evident that the classifier can exactly identify the anomaly states for this attack, and the negligible deviations between SR errors and SE errors without attack demonstrate the proposed scheme can effectively reconstruct states. We also present the SR performance in Fig. 9 for an attack strategy designed to overload two independent lines, i.e., 7-8 and 6-13, at 11:00 PM. It indicates the proposed scheme is capable of recovering states for an attack with multiple overloaded lines even the false positive and false negative issues exist in contaminated state separation procedure.



(a) Voltage phase angle error (b) Voltage magnitude error

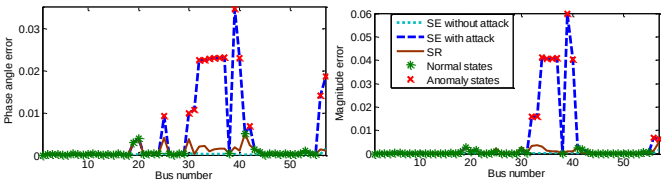
Fig. 8. Absolute estimation/reconstruction errors for IEEE 14-bus system with an attack on line 7-8 at 5:00 AM.



(a) Voltage phase angle error (b) Voltage magnitude error

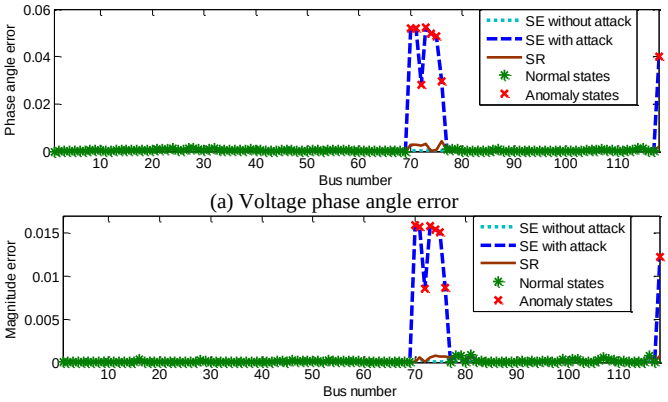
Fig. 9. Absolute estimation/reconstruction errors for IEEE 14-bus system with an attack on lines 7-8 and 6-13 at 11:00 PM.

Moreover, an FDIA was implemented on the subnetwork of the IEEE 57-bus system, shown in Fig. 3(a), at 7:00 AM, with reconstruction result presented in Fig. 10. To validate the scalability of the proposed SR scheme, we considered an FDIA launched on the subnetwork of the IEEE 118-bus system, depicted in Fig. 3(b), to overload lines 69-70 and 69-75 at 8:15 AM, and Fig. 11 illustrates its SR results. It is evident that the contaminated system states due to cyber-attack can be recovered with high accuracy by using the proposed SR scheme, irrespective of network topology. The great performance demonstrates the scalability of our proposed scheme for solving a large-scale SR problem.



(a) Voltage phase angle error (b) Voltage magnitude error

Fig. 10. Absolute estimation/reconstruction errors for IEEE 57-bus system with an attack on lines 36-40 and 37-39 at 7:00 AM.



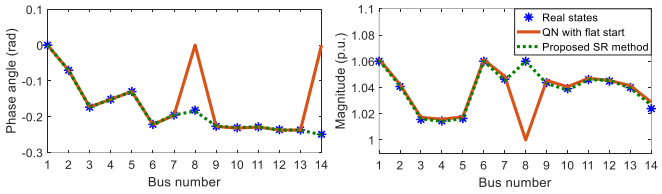
(a) Voltage phase angle error

(b) Voltage magnitude error

Fig. 11. Absolute estimation/reconstruction errors for IEEE 118-bus system with an attack on lines 69-70 and 69-75 at 8:15 AM.

Generally, the measurement redundancy is large enough for SR to converge when bad measurements are excluded. However, in some rare cases, the FDIAs are imposed on the system's critical measurements and the removal of attacked critical measurements causes the system to become unobservable. To emphasize the performance of state recovery method, an FDIA was launched in IEEE 14-bus system to cause three states, i.e., voltage magnitude of bus 8 and voltage phase angles of buses 8 and 14, to be unobservable. In this case, the matrix $\boldsymbol{G} = \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}$ becomes singular and TWLS will fail to reconstruct the system states. The SR results obtained from the proposed method and QN method with flat start are depicted in Fig. 12.

It is worth noting that both the two methods can be normally implemented because the matrix $\boldsymbol{D}$ in QN method is kept positive definite to simulate $(\boldsymbol{J}''(\boldsymbol{x}))^{-1}$ so that singular $\boldsymbol{J}''(\boldsymbol{x})$ is avoided. Moreover, in QN method, the unobservable states use the corresponding initial values as their SR results. Therefore, the magnitude 1 for bus 8 and phase angle 0 for buses 8 and 14 are obtained by using the QN method with flat start. Due to the acceptable forecasting accuracy of system states and the use of forecasted states as the initial guess, the proposed SR scheme is able to effectively solve the problem of insufficient observability caused by the removal of attacked critical measurements.



(a) Voltage phase angle (b) Voltage magnitude

Fig. 12. Performance of the proposed SR method and the QN method with flat start when critical measurements are compromised.

To further demonstrate the robustness of our scheme, a series of SR processes for the three test systems were carried out. For each test system, 20 independent attacks, ten designed to overload one line, and the other ten to overload two lines, were launched on various operating snapshots of the system. The statistical results of the overall performance indicator $\boldsymbol{\psi}_t$ for SE and SR are tabulated in Table IV. It is evident that our scheme exhibits robust, stable and appealing performance for state construction, irrespective of network topology, operating snapshots, attack types and objectives. It is observed that the indicator $\boldsymbol{\psi}_t$ of SE sharply increase after launching FDIAs, indicating the proposed attack strategy can successfully cause the state estimator to output erroneous values to system operator, and thus make either physical or economic impacts on the power system. Moreover, the reconstruction error varies from 0.0026 to 0.0157, from 0.0262 to 0.0777, and from 0.0127 to 0.0899 for IEEE 14-, 57-, and 118-bus systems, respectively, which demonstrates the reconstructed states stay very close to their real values. From above results, we can conclude that our scheme exhibits robust, stable and appealing performance for SR, irrespective of network topology, operating snapshots, attack types and objectives.

TABLE IV

THE STATISTICAL RESULTS OF ψ_t FOR EACH TEST SYSTEM

System	NoAL‡		Average	STD	Maximum	Minimum
14-bus	1	SE$^{\times}$	0.0020	0.0010	0.0036	0.0018
		SE†	0.0260	0.0187	0.0522	0.0110
		SR	0.0056	0.0019	0.0085	0.0026
	2	SE$^{\times}$	0.0013	0.0003	0.0017	0.0010
		SE†	0.1014	0.0414	0.1324	0.0472
		SR	0.0119	0.0036	0.0157	0.0070
57-bus	1	SE$^{\times}$	0.0153	0.0043	0.0184	0.0063
		SE†	0.0530	0.0053	0.0606	0.0448
		SR	0.0230	0.0075	0.0441	0.0262
	2	SE$^{\times}$	0.0162	0.0048	0.0222	0.0090
		SE†	0.1032	0.0122	0.1273	0.0875
		SR	0.0484	0.0181	0.0777	0.0282
118-bus	1	SE$^{\times}$	0.0093	0.0042	0.0161	0.0049
		SE†	0.0665	0.0213	0.0993	0.0338
		SR	0.0378	0.0249	0.0860	0.0127
	2	SE$^{\times}$	0.0119	0.0021	0.0155	0.0093
		SE†	0.1537	0.0197	0.1748	0.1136
		SR	0.0676	0.0138	0.0899	0.0531

‡number of attacked lines; $^{\times}$SE without attack; †SE with attack

V. CONCLUSION

In this paper, a generic FDIA model is proposed to handle measurements from both SCADA and PMU. Then, a novel SR scheme, consisting of E^{3}LM-based classifier, enhanced bad data identifier and state recovery approach, is developed to detect the possible data manipulation and recovery the system states. The feasibility of the attack model and SR scheme have been demonstrated on IEEE 14-, 57- and 118-bus systems. The numerical results validate that the proposed FDIAs can bypass the traditional BDD procedure and thus severely threaten the security of the power system, which reveals the vulnerability of CPPS and further emphasizes the urgency of updating the traditional SE. Meanwhile, the case studies also show that the various profiles of the SR errors from our scheme almost overlap with that of the SE errors without attacks in most simulation cases, which means our scheme can recovery system states with promising performance, strong robustness, and high stability, regardless of the system topologies, operating conditions, attack types and target lines.

REFERENCES

[1] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Trans. Ind. Inf.*, vol. 11, no. 5, pp. 1-12, Oct. 2015.

[2] H. Ye, Q. Mou, and Y. Liu, "Calculation of critical oscillation modes for large delayed cyber-physical power system using pseudo-spectral discretization of solution operator," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4464-4476, Nov. 2017.

[3] H. Wang, J. Ruan, B. Zhou, C. Li, Q. Wu, M. Q. Raza, and G. Cao, "Dynamic data injection attack detection of cyber-physical power systems with uncertainties," *IEEE Trans. Ind. Inf.*, pp. 1-10, Feb. 2019.

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

[5] T. Wu, C. Y. Chung, and I. Kamwa, "A fast state estimator for systems including limited number of PMUs," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4329-4339, Nov. 2017.

[6] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Inf.*, vol. 13, no. 2, pp. 411-423, Apr. 2017.

[7] J. Zhang, and L. Sankar, "Physical system consequences of unobservable state-and-topology cyber-physical attacks," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2016-2025, Jul. 2016.

[8] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang, "Power system reliability evaluation considering load redistribution attacks," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 889-901, Mar. 2017.

[9] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "A system for denial-of-service attack detection based on multivariate correlation analysis," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 2, pp. 447-456, Feb. 2014.

[10] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645-658, Dec. 2011.

[11] A. Abur, and A. G. Expósito, *Power System State Estimation Theory and Implementation*, New York, NY, USA: Marcel Dekker, 2004.

[12] G. Hug, and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362-1370, Sep. 2012.

[13] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Trans. Ind. Inf.*, vol. 14, no. 11, pp. 4766-4778, Nov. 2018.

[14] G. Valverde, S. Chakrabarti, E. Kyriakides, and V. Terzija, "A constrained formulation for hybrid state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1102-1109, Aug. 2011.

[15] Q. Yang, L. Jiang, W. Hao, B. Zhou, P. Yang, and Z. Lv, "PMU placement in electric transmission networks for reliable state estimation against false data injection attacks," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1978-1986, Dec. 2017.

[16] S. Pal, B. Sikdar, and J. H. Chow, "Classification and detection of PMU data manipulation attacks using transmission line parameters," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5057-5066, Sep. 2018.

[17] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi, "Joint-transformation-based detection of false data injection attacks in smart grid," *IEEE Trans. Ind. Inf.*, vol. 14, no. 1, pp. 89-97, Jan. 2018.

[18] O. A. Beg, T. T. Johnson, and A. Davoudi, "Detection of false-data injection attacks in cyber-physical DC microgrids," *IEEE Trans. Ind. Inf.*, vol. 13, no. 5, pp. 2693-2703, Oct. 2017.

[19] J. J. Q. Yu, Y. Hou, and V. O. K. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Trans. Ind. Inf.*, vol. 14, no. 7, pp. 3271-3280, Jul. 2018.

[20] J. Zhao, and L. Mili, "Sparse state recovery versus generalized maximum-likelihood estimator of a power system," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1104-1106, Jan. 2018.

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

[22] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033-1044, May 2014.

[23] Y. Chen, E. Yao, and A. Basu, "A 128-channel extreme learning machine-based neural decoder for brain machine interfaces," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 3, pp. 679-692, Jun. 2016.

[24] H. Li, H. Zhao, and H. Li, "Neural-response-based extreme learning machine for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 539-552, Feb. 2019.

[25] Z. Chen, C. Jiang, and L. Xie, "A novel ensemble ELM for human activity recognition using smartphone sensors," *IEEE Trans. Ind. Inf.*, vol. 15, no. 5, pp. 2691-2699, May 2019.

[26] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, "E^2LMs: Ensemble extreme larning machines for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1060-1069, Apr. 2014.

[27] S. Chai, Z. Xu, and Y. Jia, "Conditional density forecast of electricity price based on ensemble ELM and logistic EMOS," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3031-3043, May 2019.

[28] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Security*, Chicago, IL, USA, Nov. 2009, pp. 21-32.

[29] X. Liu, and Z. Li, "Local load redistribution attacks in power systems with incomplete network information," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1665-1676, Jul. 2014.

[30] X. Liu, Z. Bao, D. Lu, and Z. Li, "Modeling of local false data injection attacks with reduced network information," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1686-1696, Jul. 2015.

[31] X. Liu, and Z. Li, "False data attacks against AC state estimation with incomplete network information," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2239-2248, Sep. 2017.

[32] R. Deng, and H. Liang, "False data injection attacks with limited susceptance information and new countermeasures in smart grid," *IEEE Trans. Ind. Inf.*, vol. 15, no. 3, pp. 1619-1628, Mar. 2019.

[33] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1731-1738, Sep. 2012.

[34] X. Liu, and Z. Li, "Trilevel modeling of cyber attacks on transmission lines," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 720-729, Mar. 2017.

[35] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659-666, Dec. 2011.

[36] D. Choi, and L. Xie, "Ramp-induced data attacks on look-ahead dispatch in real-time power markets," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1235-1243, Sep. 2013.

[37] M. R. Mengis, and A. Tajer, "Data injection attacks on electricity markets by limited adversaries: Worst-case robustness," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5710-5720, Nov. 2018.

[38] P. Zhuang, R. Deng, and H. Liang, "False data injection attacks against state estimation in multiphase and unbalanced smart distribution systems," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6000-6013, Nov. 2019.

[39] R. Tan, H. H. Nguyen, E. Y. S. Foo, D. K. Y. Yau, Z. Kalbarczyk, R. K. Iyer, and H. B. Gooi, "Modeling and mitigating impact of false data injection attacks on automatic generation control," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1609-1624, Jul. 2017.

[40] Y. Tan, Y. Li, Y. Cao, and M. Shahidehpour, "Cyber-attack on overloading multiple lines: A bilevel mixed-integer linear programming model," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1534-1536, Mar. 2018.

[41] Y. Xu, Z. Y. Dong, J. H. Zhao, P. Zhang, and K. P. Wong, "A reliable intelligent system for real-time dynamic security assessment of power systems," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1253-1263, Aug. 2012.

[42] A. Khamis, Y. Xu, Z. Y. Dong, and R. Zhang, "Faster detection of microgrid islanding events using an adaptive ensemble classifier," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1889-1899, May 2018.

[43] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 41-53, Feb. 2016.

[44] G. D. Wyss, and K. H. Jorgensen, A User's Guide to LHS: Sandia's Latin Hypercube Sampling Software, Sandia National Laboratories, Albuquerque, NM, USA, Tech. Rep. SAND98-0210, 1998.

[45] J. J. Q. Yu, A. Y. S. Lam, D. J. Hill, Y. Hou, and V. O. K. Li, "Delay aware power system synchrophasor recovery and prediction framework," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3732-3742, Jul. 2019.

[46] C. Gu, and P. Jirutitijaroen, "Dynamic state estimation under communication failure using Kriging based bus load forecasting," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2831-2840, Nov. 2015.

[47] M. Nejati, N. Amjady, and H. Zareipour, "A new stochastic search technique combined with scenario approach for dynamic state estimation of power systems," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2093-2105, Nov. 2012.

[48] Y. Nie, C. Y. Chung, and N. Z. Xu, "System state estimation considering EV penetration with unknown behavior using quasi-Newton method," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4605-4615, Nov. 2016.

[49] R. Fletcher, "Unconstrained Optimization," in *Practical Methods of Optimization*, 2nd ed. New York, NY, USA: Wiley, 1987.

[50] R. Zimmerman, and D. Gan. "MATPOWER: A Matlab Power System Simulation Package," [Online]. Available: http://www/pserc.cornell.edu. mapower, Dec. 2016.