IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 4, 2018, accepted December 15, 2018, date of publication December 19, 2018,
date of current version January 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2888582

False Data Injection Attacks With Incomplete Network Topology Information in Smart Grid

YUANCHENG LI AND YUANYUAN WANG
School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China
Corresponding author: Yuanyuan Wang (yuan1call2me@163.com)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2018ZD06.

ABSTRACT False data injection (FDI) attack causes a disadvantage to the safety of a power system. In the past, the adversary launching FDI attack mostly needed the complete network topology information of the power system and a large amount of measurements. However, the reality is that the adversary has limited resources and can hardly manipulate massive measurements without being detected. This paper proposes a new method to form attack vector in the condition of incomplete system information. The proposed method first maps the restricted data into a new Jacobian matrix by a kernel-independent component analysis, then constructs the undetectable attack model, and finally, designs a less costly attack vector to accomplish the FDI attack. The experimental results on different IEEE buses with Matpower tool illustrate the validity of the proposed method. Additionally, the proposed method shows a relatively high success rate in the cases of different degrees of incomplete information and various quantities of manipulated data in different power systems.

INDEX TERMS False data injection attack, smart grid, attack vector, incomplete information, state estimation.

I. INTRODUCTION

The combination of smart grid and information technology lead to tremendous dangers of hacker attacks. As a result, higher security requirements are necessary to meet the data stable transportation in smart grid. In the tradition power system, Supervisory Control And Data Acquisition (SCADA) is used for real-time monitoring system, and collected measurements by Remote Terminal Units (RTUs) transport to State Estimator. State Estimator filters noise data and estimate the security of current system. Then, the evaluating result is used for Optimal Power Flow, Energy Distribution, Contingency Analysis, Real-time Pricing, etc. Furthermore, PMU (Phasor Measurement Units), wide area measurement system (WAMS) and other physical infrastructures are deployed [1]–[6]. Although those methods can protect the system to some degree, the malicious intruder still can steal and damage measurements in system, meanwhile, passing by the bad data detection (BDD) mechanism [7]–[12].

Attacks that destroy data integrity have been studied in previous researches. In those researches, attack usually has whole knowledge of smart grid topology information and parameter in transmission circuit. Accordingly, it is easier to derivate attack vector, inject error to measurements, change state estimation without being detected by BDD,e.g., Vukovic *et al.* proposed the denial of service attack that disrupt control center of communication infrastructure in connected power system. It lead to system operator strike of each locality [7]. Liu and Li [10] proposed the local load redistribution attack based on incomplete electric network information. In paper [13]–[17], the malicious intruder manipulate measurements and inject artificially generated data into primordial measurements passing by BDD. In general, this type of data integrity attack also known as False data injection (FDI) attack. Besides, Xie *et al.* [11] also proposed that attackers explore the weak point and launch FDI attack in smart grid when they have a small amount of system information. Zong-Han Yu *et al.* proposed that attacker construct attack vector apply with principal component analysis method when they don't have Jacobian matrix and distribution of state variables.

In regard to those threats to security of smart grid, a great deal of research about FDI attack identification, filtering and detection have been done. Such as in paper [18], Sridhar *et al.* introduced the detection of automatic generation control (AGC) anomalous and alleviating loss mechanism. Liu *et al.* [19] constructed detection of FDI attack which

specially exist in smart meters. Chaojun *et al.* [20] detected FDI attack by Kullback-Leibler distance comparing the different between current and history measurements. Chen *et al.* [21] detected attack by Chi-square detector and Cosine Similarity Matching method. Rawat and Bajracharya [22] analyzed the correlation between each part of power system and proposed FDI attack detection method. After all, Detection of FDI attack which possibly exist in the power system can explore the weak point of system. Accordingly, implementing precaution in advance will make smart grid more efficiently defend FDI attack [23]–[28]. In paper [29], the principal component analysis method is adopted at first to reduce dimension of measurements, then separate normal data and abnormal data by clustering. In paper [36], the Markov chain theory and Euclidean distance metric based method is able to detect malicious attack. Some defense mechanisms focus on FDI attack have been studied [37]–[39]. In paper [38], the defender can incur due to a lack of information about the actions of the attackers. In paper [38], a graph-theoretic framework is proposed to generalize the analysis of FDI attacks in smart grid. The proposed algorithm characterizes the critical set of measurements which must be removed along with a certain measurement to make the system unobservable.

As described in paper [30], [31], after selected and falsified the sets of measurements by attack vector, malicious attacker changes the state estimation in system without being detected. In paper [30], a least cost attack method has been constructed. Ozay *et al.* [31] showed an attack mechanism based on Gaussian procedural formation. A. A. Cardenas *et al.* proposed the attack based on game-theoretic. With the aim of stealing some power resources at the least possibility of being detected, attacker gets the balance between cost and power loss.

However, the incomplete topology information of power is rarely taken into count in those existed formation of attack vector [32]. In pater [33], authors adopt iteration to get attack vector for interrupting state estimation. However, those methods have great computational complexity.

In this paper, we present a new method of getting effective FDI attack vector which doesn't need the whole topology information of power system and doesn't have to falsify a large amount of measurements transported in system lines. In first, map the limited system topology information to high dimension by KICA method. After that, get a Jacobian matrix of incomplete topology information which is used for assuming the current state of power system. Next, this paper constructs a new model of FDI attack without being detected by BDD. At last, obtain attack vector by formed Lagrangian function which has light computational cost. In the end of this paper, abundant experiments have been shown. The experimental results prove that attack vector obtained by proposed method not only has fast speed of formation, but also can launch FDI attack with different degree of integrity. Besides, the attack vector also has high success rate when attack only can manipulate finite measurements.

The remainder of this paper is formatted as follows. Section II presents the review of state estimation, BDD and FDI attack. Section III introduces the formation of attack vector, including the Jacobian matrix of incomplete topology information obtained by KICA, the model of FDI attack, and the Lagrangian function for attack vector. Section IV shows the experimental results based on different IEEE standard bus system. Finally, the paper is summarized in section V.

II. PROBLEM

In smart grid, measurements that include load voltage, active power, voltage phase angle et al. are collected in buses, transported to monitor and control center, e.g. SCADA system. State estimator computes if the residual satisfy the threshold to identify and filer error data. There are two models of power flow in state estimation of power system which are AC model and DC model.

A. STATE ESTIMATION

The vector $z = [z_1, z_2, z_3, \ldots, z_m]^T \in \mathbb{R}^m$ presents measurements of different buses. m is the total number of measurement vectors, e.g. m is 304 in IEEE-118 system. Besides, m also is the number of rows of state matrix H of power system that mentioned behind. The vector $x = [x_1, x_2, x_3, \ldots, x_n]^T \in \mathbb{R}^n$ presents state variable in current system. n is the number of features of measurements that the n phase angle needed by state estimation. n is also the number of column of state matrix H. Usually $m > 2n + 1$.

The nonlinear relationship between measurements and statements can be formulated by

$$Z = h(x) + e, \tag{1}$$

where vector $e = [e_1, e_2, e_3, \ldots, e_m]^T \in \mathbb{R}^m \sim \mathcal{N}(0_{m\times n}, \Sigma_e)$ presents noise vector which follows Gaussian distribution. Its average value is zero vector $0_{m\times n}$ and covariance matrix is $\Sigma_e = diag(\sigma_{11}, \sigma_{11}, \ldots, \sigma_{mm})$. $H \equiv \frac{\partial H(x)}{\partial x}|_{x=0}$ is the Jacobian matrix decided by topology structure of power system and parameter attribute of transmission circuits [34]. It is also called topology matrix. We simply describe the formation of matrix H. l presents the number of transmission circuits and assume that the transmission direction is arbitrary. $l \times n$-sized matrix L presents the connection of each transmission circuit. i presents one transmission line. $L_{ij} = 1$ means that line i and j connect and current flows form i to j. $L_{ij} = -1$ means that line i and j connect and current flows form means that line i and j connect and current flows form i to j to i. $L_{ij} = 0$ means that line i and j doesn't connect. The direction of power flow is same as the direction of connection. Meanwhile, the $l \times l$-sized diagonal matrix D presents admittance of transmission circuits. Thus, matrix $H = (L^T DL, DL, -DL)^T$. Since $m \gg n$ is assumed, the rank of matrix H is $R(H) = n$. The state estimation in power system has three principles including maximum similarity principle, weighted least square method, and minimized variable. When the measured noises follow the Gaussian distribution that has mean value of $0_{m\times l}$, state value x based on maximum

similarity principle can obtained by Eq (2).

$$\hat{x} = (H^T WH)^{-1} H^T Wz, \tag{2}$$

where matrix W is the transport matrix of $m \times m$-sized variance matrix.

B. BDD

Compare the difference between measurements vector z and estimated measurements vector to obtain the residual vector r. It can be expressed as Eq (3).

$$r = z - H\hat{x}, \tag{3}$$

where the expected value of r is zero and can be expressed as $E[r] = 0_{m\times l}$.

Define covariance matrix as Eq (4).

$$Cov(r) = (I_{m\times m} - G)\Sigma_e (I_{m\times m} - G)^T, \tag{4}$$

where $G = H(H^T WH)^{-1} H^T W$ is $m \times m$-sized. Set threshold value γ based on experiences. Residual detection check whether the inequality $max_{i=1}^{m} |r| > \gamma$ is satisfied or not, in order to modify or delete the bad data caused by attack vector, sensor error or topological disorder. If the equation $max_{i=1}^{m} |r| \leq \gamma$ is satisfied, the measurements don't have bad data and be used for further state estimation.

C. FDI ATTACK

FDIA use loop holes of networks via network devices to attack SCADA communication and control centers to temper data measurements, and hinder the normal operation of state estimator. Having some knowledge of structure and topology of power grid, attacker can implement FDIA without being detected by detection system of state estimator using optimal and least cost attack vector generation method. Attacker, who has access to power grid sensors, can easily perform stealth SDIA attack by skillfully modifying sensors measurements.

Two types of attacks are expected on power grid with SCADA control system depending upon attacker's knowledge. If attacker is familiar with electrical equipment, configuration parameters, topology used and network connectivity, attack is skilled and structured. Else attack is unstructured. These attacks possibly target state estimator, decision system or management system, to steal power grid system's information like configuration parameters, topology used and system state information. Using this information FDIA can partially delete, replace or tamper measured data. If all or majority of measured data is tempered, it can be detected easily by detection system, and operator will get notified for further action. But attacker having power system information can strategically temper or manipulate measured data without being detected by detection system, and leads power system toward instability. For example, an attacker manipulated state estimation output, which lead control center to make decision opposite to actual physical state of power grid system, may results as power supply interruption.

The attack vector is presented as $a = (a_1, a_2, \ldots, a_m)^T$. Eq (5) presents that inject attack vector to original measurements.

$$z_a = z + a, \tag{5}$$

where $z_a = (z_{a1}, z_{a2}, \ldots, z_{am})^T$ presents the measurements injected by attack vector. If the i^{th} value of vector a is non-zero, the i^{th} branch of measurements Z_i is manipulated to $z_i + a_i$ by attack. $\hat{x}_a$ presents the state variable injected by attack vector as follows.

$$\hat{x}_a = \hat{x} + a, \tag{6}$$

Constructed in FDI attack, the attack vector has linear relationship with the state matrix H of power system, as Eq(7). It doesn't change residue r for passing by DBB.

$$a = Hc, \tag{7}$$

where $c = (c_1, c_2, \ldots, c_n)^T$ is $n \times 1$-sized inject vector and arbitrarily non-zero vector.

Proving process is showed as follows.

As we all known, z has to satisfy Eq (8) for passing by BDD.

$$\| z - H\hat{x} \|_2 \leq \tau, \tag{8}$$

Reducing Eq (5), Eq (6), Eq (7), we obtain

$$\begin{aligned} \| \hat{z}_a - H\hat{x}_a \|_2 &= \| z + a - H(\hat{x} + c) \|_2 \\ &= \| z - H\hat{x} + (a - Hc) \|_2 \\ &= \| z - H\hat{x} \|_2 \leq \tau. \end{aligned} \tag{9}$$

As presented in Eq (9), the measurements z_a attacked by FDI attack can pass by the BDD detector without notice if $\ell_2 - norm$ of difference between z_a and z is smaller than residual τ. Meanwhile, the state variables are falsified.

D. FDI ATTACK IN AC POWER FLOW MODEL

According to Eq (1), the power flows are nonlinearly dependent on voltage magnitudes and angles, in the AC state estimation, and the system state variable includes not only the bus phase angle, but also the bus voltage amplitude. In the case of AC, malicious data injected to the original measurements can evade detection if the Eq (10) are met:

$$a = h(\hat{x} + c) - h(\hat{x}), \tag{10}$$

When an attacker attempts to inject false data into the system, there are usually two targets: one is to tamper with a particular state variable; another is to tamper with a particular measurement unit.

1) Tamper with a particular state variable: In the AC state estimation, there are two types of state variables: phase angle (θ) and voltage amplitude (V). If an attacker attempts to attack a particular state variable, all measurement units associated with that state variable are affected. The relationship between measurements and state variables is determined by the following equation:

Active and reactive power injection of bus i:

$$P_i = V_i \sum_{j=1}^{n+1} V_j(G_{ij}cos\theta_{ij} - B_{ij}sin(\theta_{ij}), \quad (11)$$

$$Q_{ij} = V_i \sum_{j=1}^{n+1} V_j(G_{ij}sin\theta_{ij} - B_{ij}cos(\theta_{ij}), \quad (12)$$

The branch active and reactive power flows between bus i and bus j:

$$P_{ij} = V_i^2(g_{si} + g_{ij}) - V_iV_j(g_{ij}cos\theta_{ij} + b_{ij}sin\theta_{ij}), \quad (13)$$

$$Q_{ij} = -V_i^2(b_{si} + b_{ij}) - V_iV_j(g_{ij}sin\theta_{ij} - b_{ij}cos\theta_{ij}). \quad (14)$$

where, V_i is the voltage amplitude of the bus i. θ_i is the phase angle of the bus i. $\theta_{ij} = \theta_i - \theta_j$ is phase angle difference. $G_{ij} + jB_{ij}$ is the susceptance of bus i and bus j, and $g_{ij} + jb_{ij}$ is the susceptance of the parallel branch of bus i.

From Eq (11) - Eq (14)), it is clear that: When attacking a state variable, such as V_i, the measurement units that need to be modified include P, Q, P_{ij} and Q_{ij}, where $j \in Q$. Of course, if attacker attempting to attack multiple state variables simultaneously, there will be more measuring units to be tampered with.

2) Tamper with a particular measurement unit: A specific measurement unit whose measurements are related to the system structure and are associated with at least two state variables. In order to change a particular measurement unit, the attacker needs to change at least one state variable that controls the measurement unit. In order to evade detection, an attacker needs to tamper with all measurement units affected by the state variable.

III. FDI ATTACK WITH INCOMPLETE TOPOLOGICAL INFORMATION

Considering the strict protection of control center and harsh geographic environment of RTUs in reality, it is difficult for attacker to acquire the whole topological information of power system or falsify a large amount of measurements. Assume the detector of stator vector estimator has been a certain degree of understanding by attacker. However, attacker doesn't know the threshold γ. Standing on the position of attacker, attack vector should have worst impact on measurements with least attack cost.

According to the hypothesis in [35], FDI attack doesn't depend on the whole information of matrix H. The author proposed the construction of attack vector based on linear independent component analysis (ICA) which can launch a FDI attack passing by traditional BDD. Since Kernel Independent Component Analysis (KICA) introduces the kernel function on the basis of ICA, the linear method is extended by nonlinear weights, and the nonlinear connection processing of the data is more capable. Based on the ability of ICA retaining the original linear data information, nonlinear relationship in data is further retained. However, in the practical applications, there are a large number of nonlinear relationships in the measurement dataset. ICA has a well effect when

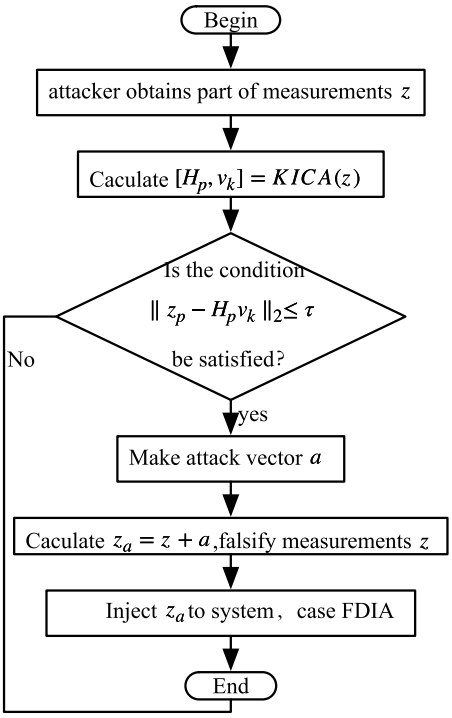

FIGURE 1. The process of FDI attack based on KICA.

the state variables obey the Gaussian distribution, while in other distributions, the extraction effect is not as good as other methods. The ICA algorithm is a linear algorithm that has poor processing capabilities for nonlinear connections. As a nonlinear algorithm, KICA can better preserve the essential relationship in nonlinear data, and KICA is not limited to the distribution type of state variables. Its generalization ability and processing ability are better, which is suitable for solving this problem. The KICA method is adopted in this paper to construction the attack vector in the condition of incomplete information of matrix H. It remodels $z = Hx$ to $z_p = HAv_k$, where A presents matrix of unknown impurity values, and $v_k, k = 1, 2, \ldots, p$ presents kernel independent components.

The measurements z_p is mapped to feature space in high dimensionality by the KICA method. Moreover, the eigenvalues $\tilde{K}$ of kernel independent components can be deduced. $z_p = \sqrt{n}\Lambda^T\tilde{K}$, where $v = [v_1, v_2, \ldots, v_p]$, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$.

$H_p = HA$ presents the matrix with parts of information unknown. Then, $z = Hx + e = H_px + e$. H_p has same number of rows with H. The columns of H_p are kernel independent components of v_k. If the dynamic maintains slight in state estimation, $\| z_p - H_pv_k \|_2 \leq \tau$. In this case, even though the whole information of Jacobian matrix H is unknown, the attacker can construct the attack vector by the deduced matrix H_p to launch a successful FDI attack without being detected. It causes a serious threat to the security of smart grid, and provides the alternative path to the malicious attacker who only has parts of measurements. The process FDI attack is shown as Figure. 1.

A. ESTABLISHMENT OF ATTACK MODE

In reality, Jacobian matrix can be presented as H_p when parts of topological information is unknown.

Define M is all the attributes of measurements, $A \subseteq M$ is all the attributes relate to FDI attack, and $S \subseteq M$ is the rest of attributes. Therefore, $S = \bar{A}, A \cup S = M$. Let $a_i \neq 0, \forall i \in A$ presents the attacking elements of vector a, accordingly, let $a_i = 0, \forall i \in S$ presents the safe elements of vector a. Assume the i^{th} element of vector n as follows:

$$\begin{cases} a_i = 1, & i \in A \\ a_i = 0, & i \in S \end{cases} \quad i = 1, 2, \ldots, \mathfrak{C} \tag{15}$$

where $\mathfrak{C}$ is cardinality of set S.

The objective function of attack vector a is modeled as follows:

$$U = pn^T a + q\frac{2}{exp(r^T Dr)/\lambda + 1}, \tag{16}$$

where U is the probability that attack is undetected. p and q are preference parameters. λ is scale parameter. $n^T a$ presents all falsified values by attack, and it also indicates the profit of FDI attack. D is $\mathfrak{C} \times \mathfrak{C}$-sized symmetric matrix. D_{ii} indicates the vulnerability of i^{th} measurement. L is the real measurements in system, and then $r = a./L$ is the proportion of attacked vectors to measurements. $./$ presents the division operation in vector. The first part of the equation (16) presents profits of attack, and the second part of the equation (16) presents the probability of the FDI attack detected after the measurements are falsified.

Define the elements in diagonal matrix N are $N_{ii} = n_i$, and the elements in diagonal matrix $\widetilde{N}$ are $\widetilde{N}_{ii} = 1 - n_i$. The aim of attack is transformed from equation (16) to the optimized model as follows.

$$\begin{aligned} &\underset{a}{maximize}\ U \\ &subject\ to \begin{cases} \| a - H_p c \| \leq \tau, \\ N(a + L) \preceq 0, \\ a^T \widetilde{N} a = 0. \end{cases} \end{aligned} \tag{17}$$

Thus, the measurements falsified by FDI attack can still satisfy the threshold τ of BDD in state estimation.

B. CONSTRUCTION OF ATTACK VECTOR

In order to achieve the objective of obtain largest profit without being detected, the optimal value of model (17) should be calculated.

Firstly, the Lagrangian function is formed as follows.

$$\begin{aligned} &\mathcal{L}(a, \lambda_1, \lambda_2, \lambda_3) = U(a) + \lambda_1^T h_1(a) + \lambda_2^T h_2(a) + \lambda_3^T h_3(a), \\ &where \begin{cases} h_1 = parallela - H_p c \| - \tau \\ h_2 = N(a + L) \\ g(a) = a^T \widetilde{N} M a \end{cases} \end{aligned} \tag{18}$$

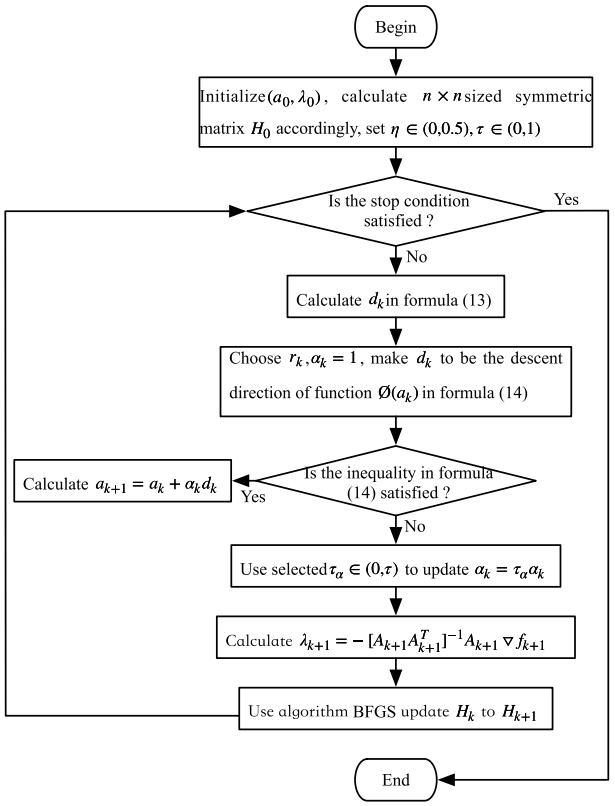

FIGURE 2. The process of constructing attack vector.

The function for calculating optimal value can be reformed as follows.

$$\begin{aligned} &\max_{d_k}\ \nabla U(a_k)^T d_k + \frac{1}{2} d_k^T H_k d_k \\ &subject\ to \begin{cases} h_1(a_k) + \nabla h_1(a_k)^T d_k \leq 0, \\ h_2(a_k) + \nabla h_2(a_k)^T d_k \leq 0, \\ g(a_k) + \nabla g(a_k)^T d_k = 0. \end{cases} \end{aligned} \tag{19}$$

where $H_k = \nabla_{aa}^2 \mathcal{L}(a_k, \lambda_k)$, $\lambda_k = \{\lambda_1^k, \lambda_2^k, \lambda_3^k\}$ presents Hessian of the Lagrangian. $\lambda_1^k, \lambda_2^k, \lambda_3^k$ are the values of $\lambda_1, \lambda_2, \lambda_3$ at k^{th} iteration respectively. d_k is used for updating attack vector in each iteration. α_k is step length. The merit function $\varnothing(a_k)$ is used for adjusting step length. The updating process of attack can presents as follows.

$$\begin{aligned} &a_{k+1} = a_k + \alpha_k d_k, \\ &where \begin{cases} \varnothing(a_k) = U(a_k) + r_1 h_1(a_k) \\ + max(0, r_3 h_3(a_k)) + r_2 h_2(a_k) \\ r_i = \max_i \{\lambda_i, \dfrac{\lambda_i + (r_k)_i}{2}\}, \quad i = 1, 2, 3 \\ \varnothing(a_k + \alpha_k d_k) \leq \varnothing(a_k) + \eta \alpha_k D_{d_k} \varnothing(a_k). \end{cases} \end{aligned} \tag{20}$$

The α_k satisfied the merit function $\varnothing(a_k)$ descend sufficiently. η is chosen parameter, and set as $\eta \in (0, 0.5)$ in this paper. $\alpha_k D_{d_k} \varnothing(a_k)$ is the directional derivative of function $\varnothing(a_k)$ on the direction d_k. r_1, r_2, r_3 are penalty factors.

In summary, the attack vector a is constructed. The algorithm is shown as Figure. 2.

1) Initialize (a_0, λ_0), calculate the $n \times n$-sized symmetric matrix H_0 accordingly, and set the parameter $\eta \in (0, 0.5)$, $\tau \in (0, 1)$.
2) If the condition is satisfied, stop. If not, goto step 3 by $k = 0, 1, 2 \ldots$
3) Calculate d_k in (18)
4) Choose r_k, $\alpha_k = 1$. Set d_k is the directional derivative of function $\varnothing(a_k)$ in (20)
5) If the inequality in (20) is satisfied, goto step 6. If not, update $\alpha_k = \tau_\alpha \alpha_k$ by the chosen $\tau_\alpha \in (0, \tau)$
6) Calculate $a_{k+1} = a_k + \alpha_k d_k$
7) Calculate $\lambda_{k+1} = -[A_{k+1} A_{k+1}^T]^{-1} A_{k+1} \nabla f_{k+1}$
8) Update H_k to H_{k+1} by BFGS algorithm

IV. SIMULATION RESULT

For proving the feasibility and validity of the proposed FDI attack with the incompletely topological information, we did the experiences to construct the attack vector on IEEE 9-bus, IEEE 14-bus, IEEE 30-bus, IEEE 57-bus, IEEE 118-bus, and IEEE 300-bus standard system used Matpower tool. The experiences run on the computer installed Matlab software and having Window 10 system, dual-core CUPs with 3.9 GHz speed, 16G memory capacity.

A. CONSTRUCTION OF ATTACK VECTOR

Firstly, we run 100 times to construct the valid attack vector with a certainly incomplete matrix H on six different standard IEEE bus system. The results of experience are shown as Fig. 3. As Fig. 3.(a) indicated, the valid FDI attack vector a can be constructed rapidly, and the time cost increases with the scale of system becoming bigger. Meanwhile, the volatility of constructing time in IEEE 9, 14, 30 bus system is more drastic than in IEEE 57, 118, 303 bus system.

Then, we record the time of constructing attack vector a on different IEEE bus system when the topological information is 20%, 35%, 50% complete. From results of experience shown in Fig. 4., we can see that the valid FDI attack vector a can be constructed rapidly and there are narrowly distance of the constructing time between different degree of incompletion. It proves that the constructing algorithm proposed in this paper has exceptional timeliness. The time cost almost is same as the condition that attack has the whole topological information of power system. Meanwhile, the constructing time is stable, e.g. the time cost basically maintain between 0.008s to 0.011s in IEEE 14 bus system, and in between 0.32s to 0.35s in IEEE 300 bus system with three differently incomplete degree of sparse matrix H.

B. THE ANALYSIS OF VALIDITY OF ATTACK VECTOR

The incomplete rate of topological information is set as the parameter in the interval [0,1]. A successful FDI attack means that construct the attack vector with different incomplete degree, then inject the vector to measurements for

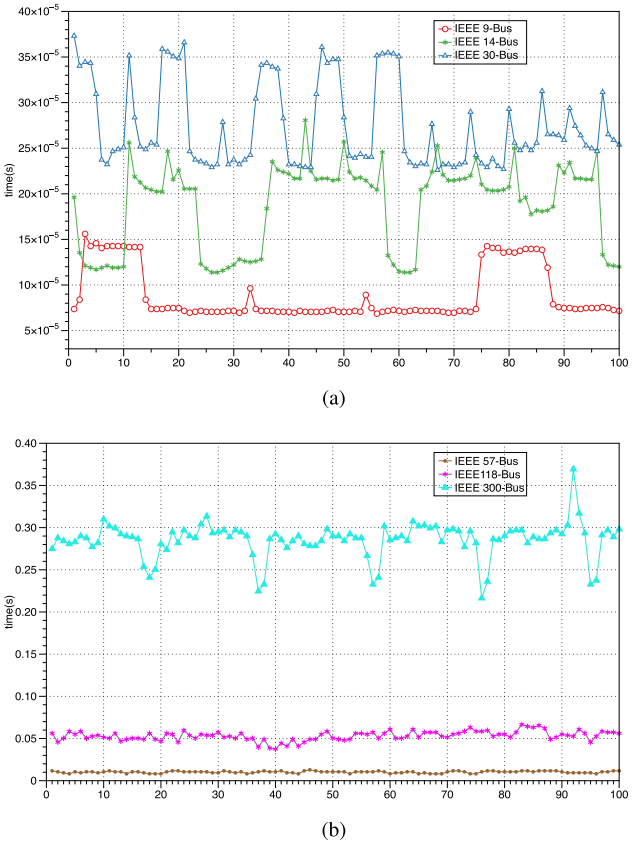

FIGURE 3. The time of constructing attack vector on different IEEE bus. (a) IEEE 9, 14, 30 bus. (b) IEEE 57, 118, 300 bus.

profit, and pass by the detector without alarm. The success rate of FDI attack acceding to different incomplete rate on IEEE 14, 118 bus system is shown in Fig. 5. The success rate becomes higher with the number of attacked lines increasing. However, The success rate becomes lower with the incomplete rate increasing, e. g. The success rate is below 50% when the incomplete rate greater than 0.09 in IEEE 14 bus system. Besides, the distance of success rate according to different amount of attacked line in IEEE 118 bus is bigger than in IEEE 14 bus. The former still have the possibility of successful FDI attack when the incomplete rate is 0.14 and the amount of attacked line is 3.

It should be Considered that the finite number of sensor that the attacker can obtain and manipulate when the validity of attack vector is evaluated. With a certain incomplete degree of topological information, the success rate of FDI attack acceding to different amount of tempered measurements on IEEE 14, 118 bus system is shown in Fig. 6. The success rate becomes higher with the amount of tempered measurements increasing.

Although the amount of tempered measurements is limited, the attack vector still has exceptional success rate when falsified data more than a certain proportion, e.g. the success rate is over 80% when the number of falsified data is 34 and 127 in IEEE 14 and 118 bus respectively.

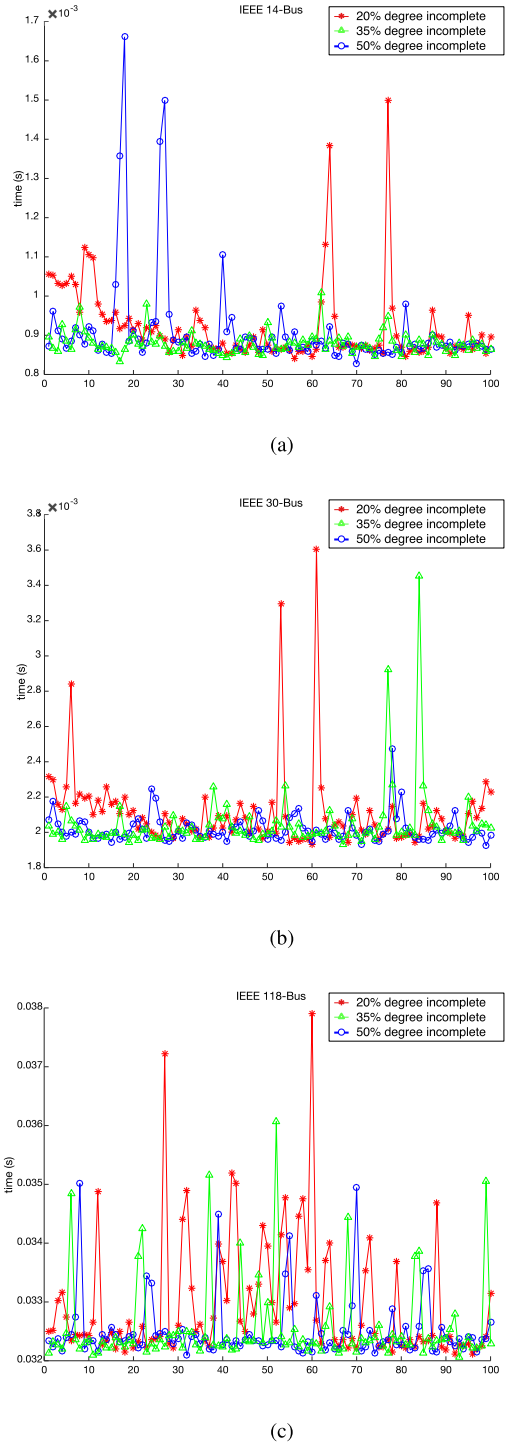

FIGURE 4. The time of constructing vector with different degree of incomplete information. (a) IEEE 14 bus. (b) IEEE 30 bus. (c) IEEE 118 bus.

In addition, we use three different detectors to verify the success rate of attacks by using attack vectors generated by different degrees of incomplete information on 14-bus and 118-bus. The three detectors are residual based detector BDD, extreme learning machine (ELM) based detector [40] and deep belief networks (DBNs) based detector [41]. In our previous work [40], [41], a detection platform based on ELM,

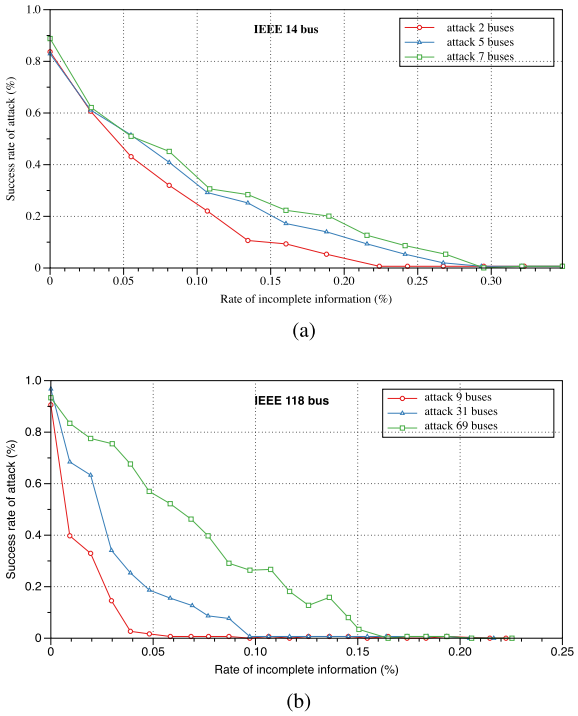

FIGURE 5. The success rate of attack according to different incomplete rate of topological information. (a) IEEE 14 bus. (b) IEEE 118 bus.

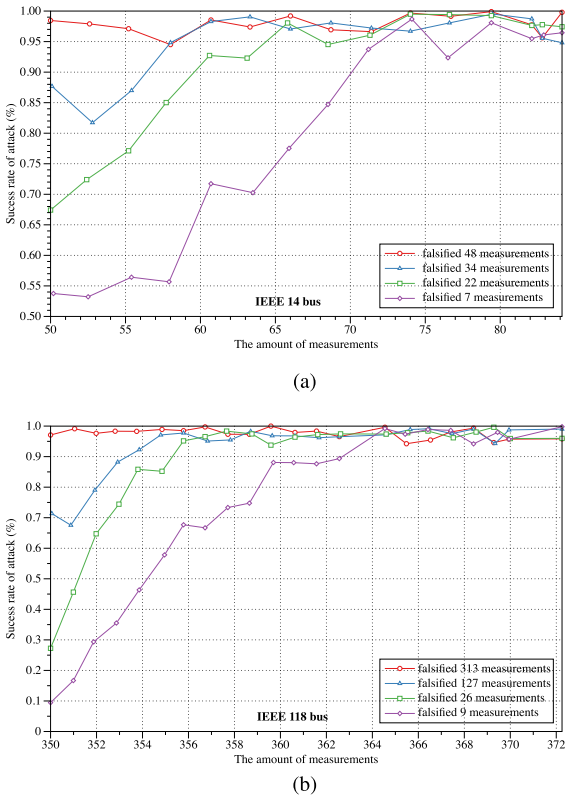

FIGURE 6. The success rate of attack according to different number of falsified data. (a) IEEE 14 bus. (b) IEEE 118 bus.

DBN and MatPower was constructed to study the security of power systems. Here, we enhance them with additional programs emulating detection of the FDI attacks. For the

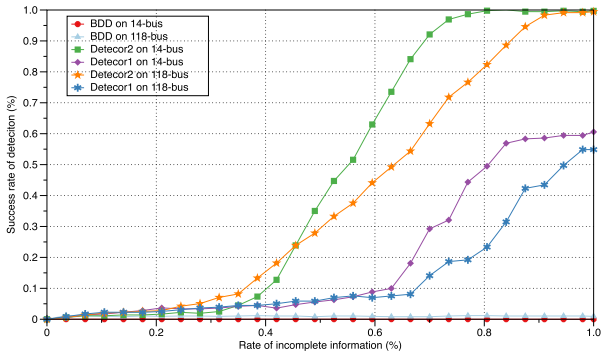

FIGURE 7. The different detector on attack vector.

sake of simplicity, in DBN based detector, we set that each hidden layer has the 64 nodes, the number of DBN layers is 4, the learning rate of pre-training and fine-tuning are both 0.05, and the size of mini-batch is 100. In ELM based detector, we randomly generates input weights and hidden-layer bias, and the output weights are obtained by analysis and calculation. Additionally, in hidden layer of ELM, we choose ''sigmoid'' function as the activation function, and the output layer choose linear activation function as the activation function. The experimental results are shown in Fig. 7, the FDI attack vector generated by the proposed method can effectively avoid the BDD detection and can successfully pass the ELM based detector 1 when the information incompleteness is less than 64%. It also can successfully pass the DBN based detector 2 when the information incompleteness is less than 38%.

V. CONCLUSION

In this paper, we present a new method of constructing FDI attack vector which not only can effectively launch a attack under the condition of uncompleted topology information and limitedly operable measurements, but also have less computational cost and faster formational speed. We disrupted the basic theory of FDI attack and considered the limited measurements obtained and falsified by attacker who has finite knowledge of topology information of current power system. We proposed the formation method of attack vector with those limitations which can also successfully launch FDI attack without being detected. In experimental results showed the rate of successful attack under different degree of topology integrity and different number of operable measurements based on various IEEE standard bus system. Its performance proved this construction of attack vector have lower requirements for the intruder of topology information and reliance of the quantity of measurements. Based on the method of attack, a potential avenue for future work is to detect the FDI attack in smart grid and to develop a forecast-based protection strategy.

REFERENCES

[1] Y. Weng, R. Negi, C. Faloutsos, and M. Ilic, ''Robust data-driven state estimation for smart grid,'' *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1956–1967, Jul. 2017, doi: 10.1109/TSG.2015.2512925.

[2] S. Maharjan, J. C.-H. Peng, J. E. Martinez, W. Xiao, P.-H. Huang, and J. L. Kirtley, ''Improved sample value adjustment for synchrophasor estimation at off-nominal power system conditions,'' *IEEE Trans. Power Del.*, vol. 32, no. 1, pp. 33–44, Feb. 2017, doi: 10.1109/TPWRD.2016.2586946.

[3] 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, doi: 10.1109/TSG.2015.2492827.

[4] S. Choi and A. P. S. Meliopoulos, ''Effective real-time operation and protection scheme of microgrids using distributed dynamic state estimation,'' *IEEE Trans. Power Del.*, vol. 32, no. 1, pp. 504–514, Feb. 2017, doi: 10.1109/TPWRD.2016.2580638.

[5] C. Murphy and A. Keane, ''Local and remote estimations using fitted polynomials in distribution systems,'' *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3185–3194, Jul. 2017, doi: 10.1109/TPWRS.2016.2630743.

[6] S. Bi and Y. J. A. Zhang, ''Graph-based cyber security analysis of state estimation in smart power grid,'' *IEEE Commun. Mag.*, vol. 55, no. 4, pp. 176–183, Apr. 2017, doi: 10.1109/MCOM.2017.1600210C.

[7] O. Vuković and G. Dán, ''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, doi: 10.1109/JSAC.2014.2332106.

[8] Y. Yamaguchi, A. Ogawa, A. Takeda, and S. Iwata, ''Cyber security analysis of power networks by hypergraph cut algorithms,'' *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2189–2199, Sep. 2015, doi: 10.1109/SmartGridComm.2014.7007750.

[9] Y. Zhang, L. Wang, Y. Xiang, and C. W. Ten, ''Power system reliability evaluation with SCADA cybersecurity considerations,'' *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1707–1721, Jul. 2015, doi: 10.1109/TSG.2015.2396994.

[10] 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, doi: 10.1109/TSG.2013.2291661.

[11] L. Xie, Y. Mo, and B. Sinopoli, ''False data injection attacks in electricity markets,'' in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Gaithersburg, MD, USA, Oct. 2010, pp. 226–231, doi: 10.1109/SMARTGRID.2010.5622048.

[12] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, ''Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures,'' in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Gaithersburg, MD, USA, Oct. 2010, pp. 220–225, doi: 10.1109/SMARTGRID.2010.5622045.

[13] M. A. Rahman and H. Mohsenian-Rad, ''False data injection attacks with incomplete information against smart power grids,'' in *Proc. IEEE Global Commun. Conf.*, Anaheim, CA, USA, Dec. 2012, pp. 3153–3158, doi: 10.1109/GLOCOM.2012.6503599.

[14] 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, doi: 10.1109/TSG.2011.2163807.

[15] S. Cui, Z. Han, S. Kar, T. T. Kim, H. Poor, and A. Tajer, ''Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions,'' *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 106–115, Sep. 2012, doi: 10.1109/MSP.2012.2185911.

[16] 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, doi: 10.1109/TSG.2015.2394358.

[17] Z. H. Yu and W. L. Chin, ''Blind false data injection attack using PCA approximation method in smart grid,'' *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015, doi: 10.1109/TSG.2014.2382714.

[18] S. Sridhar and M. Govindarasu, ''Model-based attack detection and mitigation for automatic generation control,'' *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 580–591, Mar. 2014, doi: 10.1109/TSG.2014.2298195.

[19] X. Liu, P. Zhu, Y. Zhang, and K. Chen, ''A collaborative intrusion detection mechanism against false data injection attack in advanced metering infrastructure,'' *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2435–2443, Sep. 2015, doi: 10.1109/TSG.2015.2418280.

[20] G. Chaojun, P. Jirutitijaroen, and M. Motani, ''Detecting false data injection attacks in AC state estimation,'' *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2476–2483, Sep. 2015, doi: 10.1109/TSG.2015.2388545.

[21] P. Y. Chen, S. Yang, J. A. McCann, J. Lin, and X. Yang, ''Detection of false data injection attacks in smart-grid systems,'' *IEEE Commun. Mag.*, vol. 53, no. 2, pp. 206–213, Feb. 2015, doi: 10.1109/TSG.2015.2388545.

[22] D. B. Rawat and C. Bajracharya, ''Detection of false data injection attacks in smart grid communication systems,'' *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1652–1656, Oct. 2015, doi: 10.1109/LSP.2015.2421935.

[23] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using Kalman filter," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 4, pp. 370–379, Dec. 2014, doi: 10.1109/TCNS.2014.2357531.
[24] 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.
[25] S. Bi and Y. J. Zhang, "Graphical methods for defense against false-data injection attacks on power system state estimation," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1216–1227, May 2014, doi: 10.1109/TSG.2013.2294966.
[26] S. Bi and Y. J. Zhang, "Using covert topological information for defense against malicious attacks on DC state estimation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1471–1485, Jun. 2014, doi: 10.1109/JSAC.2014.2332051.
[27] J. M. Hendrickx, K. H. Johansson, R. M. Jungers, H. Sandberg, and K. C. Sou, "Efficient computations of a security index for false data attacks in power networks," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3194–3208, Dec. 2014, doi: 10.1109/TAC.2014.2351625.
[28] H. Sedghi and E. Jonckheere, "Statistical structure learning to ensure data integrity in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1924–1933, Jul. 2015, doi: 10.1109/TSG.2015.2403329.
[29] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, "On false data-injection attacks against power system state estimation: Modeling and countermeasures," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 717–729, Mar. 2014, doi: 10.1109/TPDS.2013.92.
[30] M. Mohammadpourfard, A. Sami, and A. R. Seifi, "A statistical unsupervised method against false data injection attacks: A visualization-based approach," *Expert Syst. Appl.*, vol. 84, pp. 242–261, Oct. 2017, doi: 10.1016/j.eswa.2017.05.013.
[31] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, Aug. 2016, doi: 10.1109/TNNLS.2015.2404803.
[32] A. A. Cárdenas, S. Amin, G. Schwartz, R. Dong, and S. Sastry, "A game theory model for electricity theft detection and privacy-aware control in AMI systems," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2012, pp. 1830–1837, doi: 10.1109/Allerton.2012.6483444.
[33] 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, doi: 10.1109/TSG.2016.2521178.
[34] Y. Jiang and Z.-P. Jiang, "Global adaptive dynamic programming for continuous-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2917–2929, Nov. 2015, doi: 10.1109/TAC.2015.2414811.
[35] B. Tang, J. Yan, S. Kay, and H. He, "Detection of false data injection attacks in smart grid under colored Gaussian noise," in *Proc. IEEE Conf. Commun. Netw. Secur.*, Philadelphia, PA, USA, Oct. 2016, pp. 172–179, doi: 10.1109/CNS.2016.7860483.
[36] H. Karimipour and V. Dinavahi, "Robust massively parallel dynamic state estimation of power systems against cyber-attack," *IEEE Access*, vol. 6, pp. 2984–2995, 2017, doi: 10.1109/ACCESS.2017.2786584.
[37] A. Sanjab and W. Saad, "On bounded rationality in cyber-physical systems security: Game-theoretic analysis with application to smart grid protection," in *Proc. Joint Workshop Cyber-Phys. Secur. Resilience Smart Grids*, Vienna, Austria, Apr. 2016, pp. 1–6, doi: 0.1109/CPSRSG.2016.7684101.
[38] A. Sanjab and W. Saad, "Data injection attacks on smart grids with multiple adversaries: A game-theoretic perspective," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2038–2049, Jul. 2016, doi: 10.1109/TSG.2016.2550218.
[39] A. Sanjab, W. Saad, and T. Başar. (2018). "Graph-theoretic framework for unified analysis of observability and data injection attacks in the smart grid." [Online]. Available: https://arxiv.org/abs/1801.08951
[40] L. Yang, Y. Li, and Z. Li, "Improved-ELM method for detecting false data attack in smart grid," *Int. J. Elect. Power Energy Syst.*, vol. 91, pp. 183–191, Oct. 2017, doi: 10.1016/j.ijepes.2017.03.011.
[41] Y. Li, X. Nie, and R. Huang, "Web spam classification method based on deep belief networks," *Expert Syst. Appl.*, vol. 96, pp. 261–270, Apr. 2018, doi: 10.1016/j.eswa.2017.12.016.

YUANCHENG LI received the Ph.D. degree from the University of Science and Technology of China, Hefei, China, in 2003. From 2004 to 2005, he was a Post-Doctoral Research Fellow with the Digital Media Lab, Beihang University, Beijing, China. Since 2005, he has been with North China Electric Power University, where he is currently a Professor and the Dean of the Institute of Smart Grid and Information Security. From 2009 to 2010, he was a Postdoctoral Research Fellow with the Cyber Security Lab, College of Information Science and Technology, Pennsylvania State University, State College, PA, USA. He has hosted and participated in several research projects for the National Natural Science Foundation of China and National 863 Plan projects. He has authored more than 70 articles and more than 10 inventions. His research interests include power grid security, information security, cloud computing, big data security, and cloud security.

YUANYUAN WANG received the M.S. degree in software engineering from North China Electric Power University, Beijing, China, in 2016, where she is currently pursuing the Ph.D. degree with the Institute of Smart Grid and Information Security, School of Control and Computer Engineering. Her research interests include smart grid security, electricity market, data mining, and its applications.