

FDIA Localization Detection Based on Intuitionistic Fuzzy Set-multiple Hidden Layer Random Vector Functional Link Network

Lei Xi, Zihao Li, Zongze Li, Hongjun Chen, Fangyan Bai, and Yixiao Wang

Abstract—False data injection attacks seriously threaten the security and stability of the cyber-physical power system. To address the issue that existing attack detection methods struggle to accurately identify and locate the attacked bus, a location detection method based on an intuitionistic fuzzy set and a multiple hidden-layer random vector function link network is proposed in this paper. The proposed method extracts the data feature information of the complete measurement through a random vector function connection network to reduce the missed detection rate and expands it into a multi-hidden layer architecture to prevent overfitting of the detection method effectively. Meanwhile, the intuitionistic fuzzy method is employed to mitigate the impact of measurement noise on detection accuracy, thereby achieving high-precision attack location detection. A large number of experiments are conducted on the test systems of IEEE-14, IEEE-57, and IEEE-118 buses to verify the effectiveness of the proposed method. Compared with various methods, it is verified that the proposed method has better accuracy, precision, F1 value, recall rate and AUC value.

Index Terms—Cyber-physical power system, false data injection attack, intuitionistic fuzzy set, random vector functional-link network.

NOMENCLATURE

z	System measurement vector.
J	Measurement Jacobian matrix.
x	System state vector.
e	Measurement error vector.
r	System residual error.
$\hat{x}$	Best estimate of x .
τ	Residual detection threshold.
a	Attack vector.
z_a	System measurement vector after the attack.
x_a	System state vector after the attack.
d	Deviation of the state vector after the attack.
X_i	Original input feature.
n	Samples number.
L_i	Bus of the hidden layer.

$H(X)$	Activation function.
H	Combination matrix.
w_i	Connection weights in the input layer.
b_i	Bias of the neurons in the hidden layer.
β	Output weight of the output layer.
β^*	Least squares solution to β .
O	Objective function.
$M_T(x_i)$	True membership function.
$M_F(x_i)$	False membership function.
η_i	Evaluation function.
M_T^c and M_F^c	Class centers of the true and false data samples.
M_T^r and M_F^r	Class radii of the true and false data samples.

I. INTRODUCTION

MODERN power systems have gradually evolved into a multi-dimensional complex cyber-physical power system (CPPS) [1]–[5] that integrates computing, network and physical environment, with increasingly powerful information interaction capabilities, and it also exposes information security to the threat of network attacks. As a mainstream attack method, false data injection attacks (FDIA) [6]–[10] can bypass bad data detection (BDD) [11] in state estimation. The mechanism BDD injected attack vectors into the energy management system (EMS) [12] to achieve the purpose of destroying the security and stability of CPPS or obtaining illegal benefits. Therefore, it is crucial to study and deploy effective FDIA detection schemes.

Among the many existing detection methods, they can be divided into two detection schemes: model-driven and data-driven, based on different mechanisms. The detection schemes based on model-driven mainly utilize state estimation methods [12]–[14], such as the paper [15], [16] based on a weighted least squares estimator, assuming that the power system has sufficient redundancy and is in a stable state. FDIA is detected by a static state estimation method. However, considering that electricity demand and power generation change in real time, the paper [17] achieves real-time detection of FDIA by utilizing the time and frequency domain analysis capabilities of cross wavelet transform and a dynamic and static parallel state estimation method based on the time correlation of the dynamic system.

Manuscript received September 3, 2024; revised November 12, 2024; accepted January 14, 2025. Date of online publication January 7, 2026; date of current version March 2, 2026. This work was supported by the National Natural Science Foundation of China under Grant 52277108.

L. Xi, Z. H. Li, Z. Z. Li (corresponding author, email: lizongze0608@163.com), H. J. Chen, F. Y. Bai and Y. X. Wang are with College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China.

DOI: 10.17775/CSEEJPES.2024.06550

However, the above model-driven detection schemes strongly depend on the system model and its operating parameters. As the nonlinear structure of CPPS becomes increasingly complex, it is difficult to establish an accurate model of the system. In addition, the model-driven detection process includes a large number of matrix operations [18], and highly complex matrix operations are prone to producing ill-conditioned matrices, which significantly reduce the detection reliability.

With the continuous development and maturity of artificial intelligence [19], [20], the data-driven FDIA detection scheme has emerged [21], [22]. This detection method does not require the parameter information of the system model. Instead, an artificial intelligence algorithm is used to train and learn from the measurement data in the system, extracting features to identify the outlier characteristics of FDIA and facilitate detection. In [23], principal component analysis is utilized to train the distributed support vector machine (SVM) for binary classification of data. In [24], a symmetrical hybrid deep network model is constructed by combining a long short-term memory network and a convolutional neural network. The self-attention module is used to eliminate the redundancy of mixed features, thereby enabling the detection of FDIA. In [25], a subgrid-oriented microservice framework for FDIA detection in AC model power systems is proposed by integrating a spatial-temporal neural network and simulating the spatial-temporal relationships of the buses within subgrids.

Although the above data-driven detection method can effectively avoid the defects of model-driven detection methods that strongly depend on the system model and its operating parameters, as well as highly complex matrix operations, it can effectively realize the detection of FDIA. However, it ignores the positioning problem of FDIA. The accurate positioning of FDIA is particularly important for the rapid deployment and effective defense of the power grid when FDIA is faced. To realize FDIA location detection, the paper [26] proposes an FDIA location detection method based on an ensemble extreme learning machine (ELM). However, the traditional ELM has limited feature expression capabilities and struggles to detect anomalies in high-dimensional and complex CPPS measurement data accurately. Therefore, the paper [27] extends ELM to a multi-hidden-layer ELM to improve feature extraction ability and introduces the algorithm of Grey Wolf Optimization to optimize its hyperparameters, thereby achieving better localization and detection performance. Although the above detection methods can realize the location detection of FDIA, some feature information will be lost after the original data is processed by neuron dimension reduction, resulting in the missed detection of outliers.

The random vector functional link network (RVFLN) [28] has a similar network update mechanism to ELM. Still, it has additional direct links on the network structure, and the model can capture the complete feature information of the measurement data. In [29], RVFLN is employed for fault diagnosis, demonstrating excellent detection capabilities. However, the model of RVFLN detection with a single hidden layer is prone to overfitting during the early stage of training, which significantly affects the model's interpretability. Therefore,

the hidden layer is extended, and a multi-layer hidden layer RVFLN (RVFLN^{ML}) is constructed in this paper, allowing the model to more fully abstract the characteristics of the data and reduce the risk of overfitting. In addition, the long-term exploration reveals that the randomness of RVFLN^{ML} weights and hidden layer parameters makes it very sensitive to the noise of input data. Considering the real scenario, the measurement data collected by CPPS contains a lot of noise. If RVFLN^{ML} is used solely to locate and detect FDIA, the model will learn misleading signals introduced by noise during training. This can lead to misjudgment when the model detects new data.

The paper [30] examines intuitionistic fuzzy (IF) sets, which offer a flexible and effective mathematical approach for addressing noise problems. A membership degree is assigned to each sample feature in IF to measure its similarity with each class. The fuzzy processing method can tolerate noise and uncertainty in the data, making the classification result more accurate. Therefore, in this paper, a single-layer RVFLN is first stacked into an RVFLN^{ML} with a deep structure to address the problem of insufficient feature extraction capability in the single-layer structure. Then, IF is introduced into the RVFLN^{ML} detection model to reduce the influence of noisy data on detection accuracy.

Meanwhile, an FDIA location detection method based on an intuitionistic fuzzy multiple hidden layer random vector function connection network (IF-RVFLN^{ML}) is proposed. The effectiveness of the proposed method is verified through a large number of simulations on the standard test systems of IEEE-14, IEEE-57, and IEEE-118 buses. The superiority of the proposed method is also demonstrated through a comparative analysis with various positioning detection methods.

II. PROBLEM DESCRIPTION

A. Principle of FDIA

As a key part of CPPS, state estimation filters the measured data collected by the supervisory control and data acquisition (SCADA) system [31] and the power management unit (PMU) [32]. It obtains the state vector of the system operation after processing. The mapping relationship between the state vector and the measurement data in the process of state estimation is as follows:

$$z = Jx + e \quad (1)$$

where z represents the system measurement vector; J is the measurement Jacobian matrix determined by the network structure and linear parameters; x is the system state vector; e represents the measurement error vector. The FDIA for CPPS is shown in Fig. 1.

The carefully designed attack vector will pass the residual detection of BDD after it is injected into the system. While the false data is not recognized, the measured data of the SCADA system and PMU can be tampered with, which will lead to a misjudgment of the current power grid's status and incorrect decisions by the scheduling center. The residual error when the system is in a normal state is expressed as follows:

$$r = z - J\hat{x} \quad (2)$$

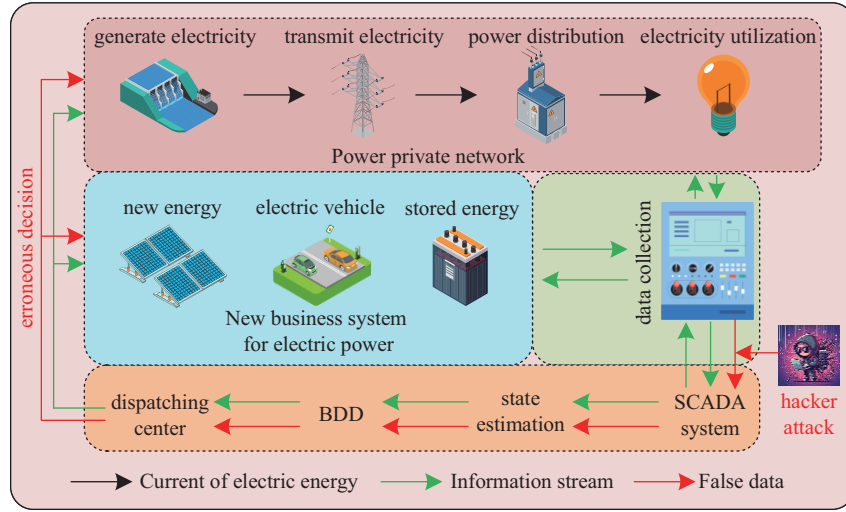


Fig. 1. FDIA schematic in CPPS.

where $\hat{x}$ is the best estimate of x while the convergence condition is satisfied. The data is identified by BDD as bad data whose L2 norm of the residual r exceeds the residual detection threshold τ , that is $\|r\| > \tau$ and it will be eliminate.

In order to evade the BDD mechanism, the attacker injected the attack vector a into the original measurement data z , then the measurement data obtained by the control center became $z_a = z + a$. Meanwhile, the state estimation using z_a would get the wrong state vector $x_a = \hat{x} + d$. At this point, the system residual after the attack is as follows:

$$\begin{aligned}
 \|r_a\| &= \|z_a - Jx_a\| \\
 &= \|z + a - J(\hat{x} + d)\| \\
 &= \|z + a - Jd - J\hat{x}\| \\
 &= \|z - J\hat{x}\| \\
 &= r < \tau
 \end{aligned} \tag{3}$$

where d is the deviation of the state vector after being attacked; As the attack vector satisfies $a = Jd$, the system residual after the attack remains unchanged and it is still below the detection threshold. This makes the traditional BDD mechanism based on residual invalid and unable to detect false data in the measurement data. Thus, FDIA is successfully injected.

B. FDIA Model

While FDIA is constructed, it can tamper with the measured values of the meter if the attacker can obtain all the configuration information of CPPS, to directly construct the above attack vector a and launch the attack successfully. Considering the real-world scenario, it is difficult for the attacker to fully grasp all the information of the power system, so the methods of common FDIA construction have constraints. The attack constraints mainly consider that the attacker can invest the limited attack resources, the attacker has incomplete system information, and some measurement devices of CPPS are protected.

With the increasing complexity of CPPS, the construction method of FDIA, which only targets the SCADA system, is obviously not comprehensive, and more measurement devices

should be considered. In this paper, the modeling method of the hybrid FDIA proposed in [26], which targets SCADA and PMU simultaneously, is adopted to generate the data sets required for simulations. The effectiveness of the attack will be verified in the section on simulations. The test topologies of the IEEE-14, IEEE-57, and IEEE-118 buses are shown in Fig. 2.

III. IF-RVFLN^{ML}

To address the issue of FDIA detection in CPPS, this paper treats FDIA detection as a multi-label binary classification problem [33] and proposes a data-driven method for IF-RVFLN^{ML} detection. In this method, RVFLN^{ML} is utilized as the feature extraction model and classifier of the sample. IF is used to address the uncertainty between training data samples, limiting the influence of noise and outliers on the model, thereby improving the detection accuracy of RVFLN^{ML}. Furthermore, it also achieves the accurate positioning detection of FDIA.

A. RVFLN

Both RVFLN and ELM belong to typical single-hidden-layer feed-forward neural networks. Unlike ELM, the network structure of RVFLN not only has an input layer, a hidden layer, and an output layer, but also features a linear link that directly connects the input layer and the output layer. The direct connection design is well-suited for extracting the potential features of the input data. Moreover, unlike common neural networks, such as convolutional neural networks [34], which require finding the optimal network parameters through iteration, RVFLN randomly generates the weights and biases of neurons in the hidden layer and determines the output weights through matrix operations. In RVFLN, the input to the output layer contains the nonlinear transformation feature H from the hidden layer and the original input feature X . The hidden layer parameters are randomly generated and remain unchanged during the training phase, which allows this network structure to prevent overfitting. The network structures of ELM and RVFLN are shown in Fig. 3.

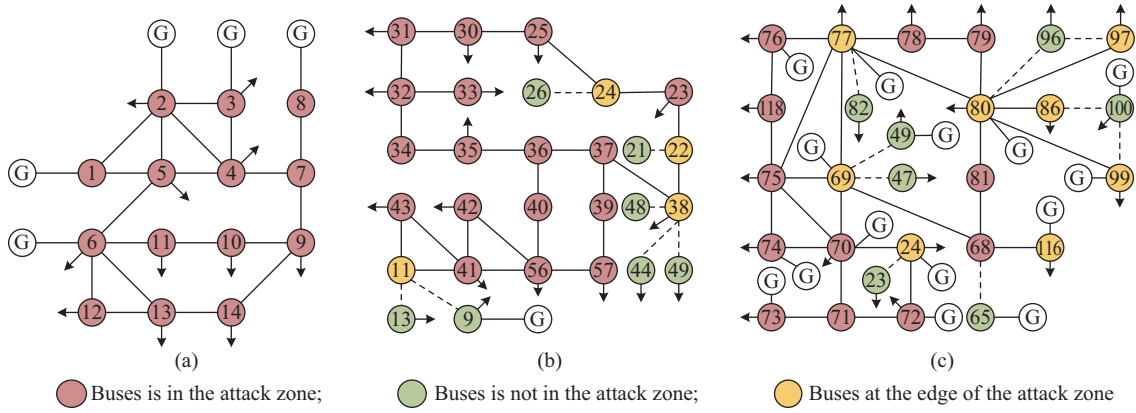


Fig. 2. The topology of the bus test system. (a) IEEE-14 bus test system. (b) IEEE-57 bus test system. (c) IEEE-118 bus test system.

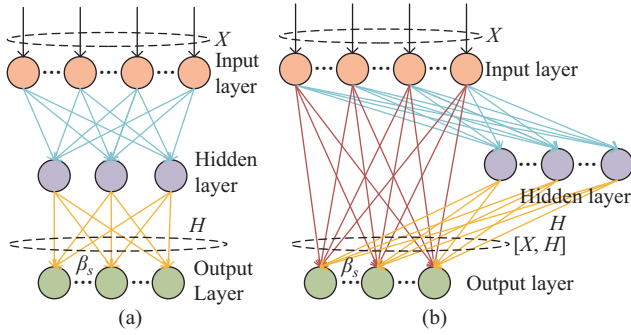


Fig. 3. The network structure of ELM and RVFLN. (a) The network structure of ELM. (b) The network structure of RVFLN.

In Fig. 3, the samples number of the input X_i in the input layer is n , X_i traverses L_i bus of the hidden layer, and nonlinear mapping is performed at the hidden layer. The activation function of the neurons in the hidden layer is $H(X)$, w_i and b_i are the connection weights in the input layer and the bias of the neurons in the hidden layer, respectively. The dimension of the connection weight w_i of each neuron is the same as the dimension of the input X_i . The output function of the RVFLN model can be described as follows:

$$f(x) = \sum_{j=1}^L \beta_j h_j(X_i) + \sum_{j=L+1}^{L+d} \beta_j x_{ij} \quad (4)$$

where $i = 1, 2, \dots, N$, β is the output weight of the output layer. (4) is expressed in matrix form as follows:

$$Y = H\beta \quad (5)$$

where

$$H = \begin{bmatrix} h_1(x_1) & \cdots & h_L(x_1) & x_{11} & \cdots & x_{1d} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_1(x_N) & \cdots & h_L(x_N) & x_{N1} & \cdots & x_{Nd} \end{bmatrix} \quad (6)$$

H is the matrix formed by combining the output values of the neurons in the hidden layer and the sample data in the input layer. The objective function solved by the model of RVFLN is as follows:

$$O = \sum_{j=1}^m \|f_i(x) - y_i\| \quad (7)$$

While the optimal parameters are being solved, the optimal solution of the network parameters is obtained by solving the minimum value of the objective function O :

$$O = \min \|H\beta - Y\| \quad (8)$$

The network parameters w and b are randomly generated after the training data is input. Therefore, the matrix H in (6) is a known quantity after the activation function type is specified by the network. The training process of RVFLN is the process of solving the least square solution β^* of β . The output weight matrix is as follows:

$$\beta^* = H^{-1}Y \quad (9)$$

where H^{-1} is the generalized pseudo-inverse matrix of the matrix H .

$$H^{-1} = (H^T H)^{-1} H^T \quad (10)$$

where H^T is the transpose matrix of H .

The training process of RVFLN can avoid complex iterative operations, and the network parameters can be quickly adjusted by solving the output weight matrix. Consequently, the entire network model can also be determined.

B. RVFLN^{ML}

With the increasing complexity of the CPPS structure and the diversity of attack characteristics, the data of the power grid is becoming more complex and multi-dimensional. As an algorithm of a single-hidden-layer feed-forward neural network, RVFLN is prone to overfitting in the early stages of training, particularly when dealing with complex data, which significantly affects the model's interpretability. Study [35] demonstrates that RVFLN can be expanded into a deep network by stacking, which significantly enhances the model's feature learning and extraction capabilities. Meanwhile, its training time and computational complexity are also smaller than those of the traditional deep network.

Therefore, RVFLN is extended to RVFLN^{ML} in this paper, and the network structure is illustrated in Fig. 4. Its hidden layer consists of several stacked single hidden layers, with the input of each hidden layer being the output of the previous layer. Furthermore, the weights and biases between the hidden layers are randomly generated, similar to RVFLN, and do not

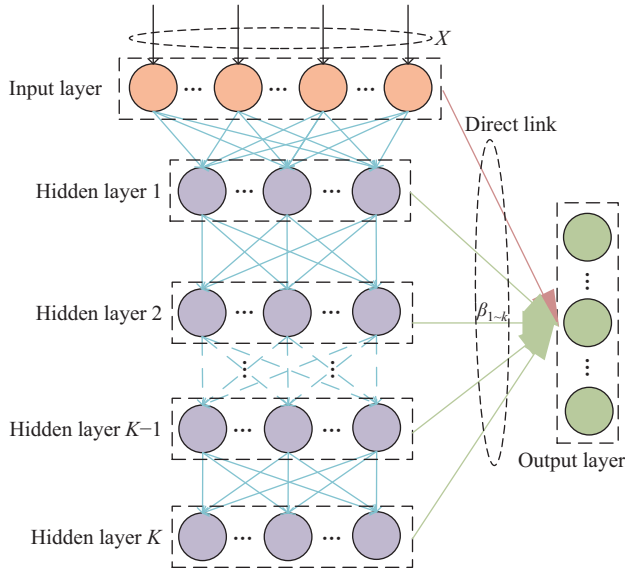


Fig. 4. The network structure of RVFLN^{ML}.

change during the training process. Only the output weight β needs to be computed.

The number of hidden layers and the number of neurons in the hidden layer of the RVFLN^{ML} network can be flexibly adjusted according to the specific problems being solved. However, for the convenience of formula description, the number of hidden layers is K , each hidden layer has the same n neurons, and the bias term is not considered. The output of the first hidden layer is defined as follows:

$$h_1 = f(\mathbf{X}w_1) \quad (11)$$

The output of the k th hidden layer is:

$$h_k = f(\mathbf{H}_{k-1}w_l) \quad (12)$$

where w is the weight matrix of each layer, $f(\cdot)$ is a nonlinear activation function. The input to the output layer is the following:

$$\mathbf{H} = [h_1, h_2, \dots, h_{l-1}, h_l, \mathbf{X}] \quad (13)$$

The transmission mode is similar to the single-hidden-layer RVFLN. The input of the output layer consists of the nonlinear features of each hidden layer as well as the original features. The output of the output layer is defined as follows:

$$\mathbf{Y} = \mathbf{H}\beta_{1 \sim k} \quad (14)$$

The linear combination between the features and the output layer weight matrix in RVFLN^{ML} enhances the model's ability to express the deep features of the measurement data, compensating for the single hidden layer RVFLN's tendency to overfit.

C. IF-RVFLN^{ML}

The algorithm RVFLN^{ML} has the ability to express strong feature expression and generalization. However, it is found to be very sensitive to noise in the measurement data collected by SCADA and PMU. Therefore, IF is introduced to deal

with the noise in the measurement data, and the RVFLN based on constrained optimization minimization is solved by adding the intuitionistic fuzzy membership degree in each training sample. The intuitionistic fuzzy set includes a true membership degree and a false membership degree, which has a stronger performance ability in dealing with noisy data. It describes the nature of the data in detail, allowing the model to accurately extract the required features. The true membership function $M_T(x_i)$, the false membership function $M_F(x_i)$ and the evaluation function η_i can be formulated by the class centers M_T^c and M_F^c of the true and false data samples. It can also be formulated by the class radii M_T^r and M_F^r of the true and false data samples.

The class centers of real and false data samples are defined as the following:

$$\begin{aligned} M_T^c &= \frac{1}{r_1} \sum_{y_i=+1} \phi(x_i) \\ M_F^c &= \frac{1}{r_2} \sum_{y_i=-1} \phi(x_i) \end{aligned} \quad (15)$$

The class radii of real and false data samples are defined as follows:

$$\begin{aligned} M_T^r &= \max_{y_i=+1} \|\phi(x_i) - M_T^c\| \\ M_F^r &= \max_{y_i=-1} \|\phi(x_i) - M_F^c\| \end{aligned} \quad (16)$$

The true membership function $M_T(x_i)$ for each point is expressed as follows:

$$M_T(x_i) = \begin{cases} 1 - \frac{\|\phi(x_i) - M_F^c\|}{M_F^r + \zeta}, & \text{if } y_i = 1 \\ 1 - \frac{\|\phi(x_i) - M_T^c\|}{M_T^r + \zeta}, & \text{if } y_i = 0 \end{cases}, \quad i = 1, \dots, r \quad (17)$$

The false membership function for each point is expressed as follows:

$$M_F(x_i) = (1 - M_T(x_i)) \frac{|\{x_j : \|x_i - x_j\| \leq \zeta, y_i \neq y_j\}|}{|\{x_j : \|x_i - x_j\| \leq \zeta\}|} \quad (18)$$

In (15) and (16), $\phi(x_i)$ is the feature map of the sample data, ζ is a positive parameter, $|\cdot|$ represents the elements number in the set and it satisfies the constraints $0 \leq M_T(x_i) + M_F(x_i) \leq 1$.

The evaluation function evaluates each data sample and classifies it by the evaluation value. The evaluation function is defined as follows:

$$\begin{aligned} \eta_i &= (M_T(x_i), M_F(x_i)) \\ &= \begin{cases} (M_T)_i, & \text{if } (M_F)_i = 0 \\ 0, & \text{if } (M_T)_i \leq (M_F)_i \\ \frac{1 - (M_F)_i}{2 - (M_T)_i - (M_F)_i}, & \text{otherwise} \end{cases} \end{aligned} \quad (19)$$

where $(M_T)_i$ and $(M_F)_i$ are the weighted values of intuitionistic fuzzy.

According to the rating value, the noisy data can be handled effectively and support vectors can be distinguished. The differentiated ratings are as follows:

- 1) While the false membership degree of the sample $M_F(x_i) = 0$, it means that the sample has no neighborhood of other classes, that is, the membership rating is $\eta_i = M_T(x_i)$.
- 2) While $M_T(x_i) > M_F(x_i)$ of the sample, $\eta_i = 0$.
- 3) While $M_T(x_i) > M_F(x_i)$ and $M_F(x_i) \neq 0$, few samples are in its neighborhood. They are treated as support vectors, then $\eta_i = 1 - (M_F)_i/2 - (M_T)_i - (M_F)_i$.

The network of IF-RVFLN^{ML} can be defined as follows:

$$\min_{\beta} \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^r \eta_i \|y_i\|^2, \quad (20)$$

The objective function solved by IF-RVFLN^{ML} is as follows:

$$\min \left(\sum [M_i, X] \beta - y_i \right) \quad (21)$$

where β is the weight vector of the output layer, C is the regularization parameter to prevent the model overfitting, η_i is the evaluation function of intuitionistic fuzzy, M_i is the output value of the neurons in the hidden layer after fuzzy intuitionistic processing.

By assigning a membership degree to each sample feature to measure its similarity with each class, IF reduces the sensitivity of the RVFLN^{ML} model to noise and improves the accuracy of the classification results. The specific process of FDIA positioning detection in the proposed IF-RVFLN^{ML} algorithm is as follows:

- 1) The samples are divided into a training set and a test set at a ratio of 4:1, and the data of the training set are input to the model.
- 2) The degree of membership and non-membership of the data in the training set are calculated according to (15) and (16).
- 3) The value of intuitionistic fuzzy evaluation is determined according to (17).
- 4) Assign a weighted fuzzy number to each training feature according to (a-c).
- 5) The optimal classifier is determined by solving the objective function of IF-RVFLN^{ML} using equations (18) and (19).
- 6) The data from the test set are input into the trained IF-RVFLN^{ML} algorithm for testing.
- 7) If the output is greater than 0.5, it is determined that the bus is attacked.
- 8) The data from real-time measurements is imported into the model to detect whether FDIA is present in the system or not and locate the affected bus.

IV. SIMULATION ANALYSIS

In this paper, the test systems of IEEE-14, IEEE-57 and IEEE-118 buses are employed as the test environment for the experiments. It is assumed that the attacker has complete information on the power grid topology in the IEEE-14 bus, but only partial information on the power grid topology in the IEEE-57 and IEEE-118 buses. In the test environment,

the problem of FDIA location detection is regarded as a multi-label binary classification problem. Each output label of the detection model corresponds to a state quantity. Each label is represented by a binary value of 0 or 1, where 0 indicates that the state quantity is normal, and 1 indicates that the state quantity is under attack, thereby verifying the detection performance of the proposed method. The equipment parameters are set as follows: 1) CPU: AMD Ryzen 5 5600 6-Core Processor; 2) RAM: 32 GB; 3) GPU: NVIDIA GeForce GTX 1080 Ti (11.0 GB).

A. Attack Model

To verify the generality of the proposed attack model, this paper utilizes load data collected from a local dispatch center to randomly select a working day for attacking the bus test system. The samples comprise 96 points within 24 hours, with a 15-minute sampling interval. The measured data in both scenarios include bus voltage, active and reactive power injection, active and reactive power flow in the branch, and voltage and current vector from the PMU. The 2480 data sets generated by each system are adopted as input data sets for the detection algorithm in the two sets of example simulation studies. The attack utilizes the method of overloading the power line. For SCADA, the bus voltage amplitude, line power and bus power need to be constrained. For PMU, the voltage amplitude and phase Angle of the bus where the PMU is located, as well as the line current amplitude and phase Angle related to the bus where the PMU is located, are constrained. The Newton-Raphson method is used to calculate the power flow. The test system of the IEEE-14 bus assumed that the attacker had complete information about the power grid to launch an attack on the power grid and overload lines 6–13. However, considering that the attacker has incomplete information about the power grid in the test systems of IEEE-57 and IEEE-118 buses, the line is randomly selected to overload in the attack area. After the overload attack, only the relevant measurements of the branches and bus directly connected to the overload point change, and the overall deviation of the other buses and branches not directly connected is very small. The measurement changes of the bus test system before and after the attack are shown below.

To study the impact of FDIA on the system, this paper analyzes the difference between the measurement data before and after the attack. Each measurement data set includes the bus voltage amplitude measured by the SCADA system, the bus active and reactive power, the active and reactive power of each branch, the bus injection power, the amplitude and phase angle of the voltage, and the amplitude and phase angle of the current measured by the PMU device. It shows in Fig. 5 that the active and reactive power of the attacked branch have a huge deviation compared with that before the attack, and the other measurement data with certain deviation are directly or indirectly related to the bus connected to the attacked line, which successfully overloads the attacked branch and achieves the purpose of the attack. The changes of voltage phase angle and voltage amplitude in the bus test system before and after the attack are shown in Figs. 6 and 7.

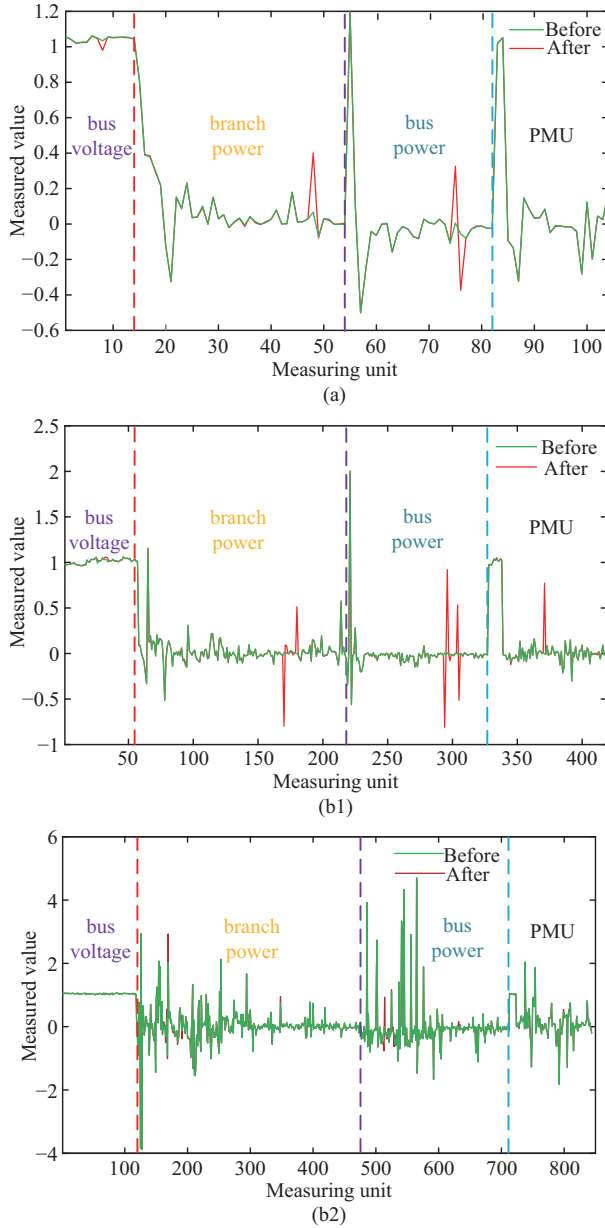


Fig. 5. Value measured before and after. (a) Attack with complete grid information on IEEE-14 bus system. (b1) Attack with incomplete grid information IEEE-57 bus system. (b2) Attack with incomplete grid information IEEE-118 bus system.

After the overload attack, the above measurements will change. However, the mechanism BDD only detects the residual error. As long as the residual error is still within the allowable range of the threshold after the attack, the attack will be considered successful. To verify that the proposed model of general FDIA can successfully evade the BDD mechanism, as shown in Fig. 8, simulations are carried out, and a comparison of the maximum normalized residual before and after the attack is obtained.

The results show that the system residuals of the IEEE-14, IEEE-57, and IEEE-118 buses after the attack are significantly smaller than the residuals before the attack, which proves that the proposed attack model can bypass the BDD detection mechanism, indicating that the attack is successful.

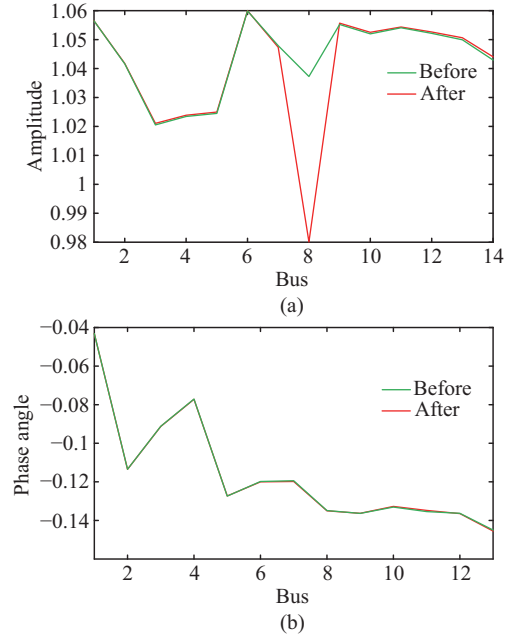
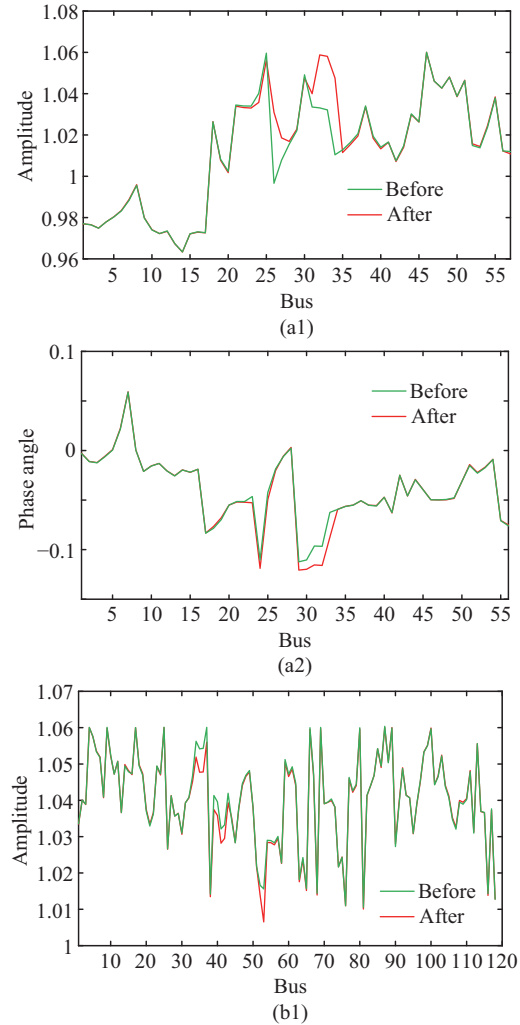


Fig. 6. Bus status under attack with complete grid information. (a) Voltage amplitude of IEEE 14 bus system before and after attack. (b) Voltage phase Angle of IEEE 14 bus system before and after attack.



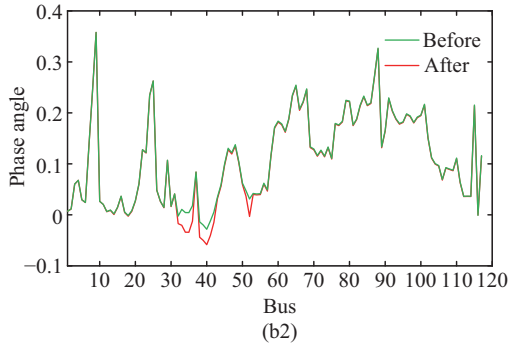


Fig. 7. Bus status under attack with incomplete information. (a1) Voltage amplitude of IEEE 57 bus system before and after attack. (a2) Voltage phase Angle of IEEE 57 bus system before and after attack. (b1) Voltage amplitude of IEEE 118 bus system before and after attack. (b2) Voltage phase Angle of IEEE 118 bus system before and after attack.

B. FDIA Localization Detection and Performance Evaluation Metrics

The measurement data generated by the FDIA model is input into the classifier for binary classification processing. Fig. 9 shows the mapping relationship between the input and output of the FDIA positioning detection. If the output is greater than 0.5, it is judged as “1”; otherwise it is “0”.

To verify the feasibility and effectiveness of the proposed method, this paper utilizes a confusion matrix to visually represent the classification results in the model, providing an intuitive understanding of the differences between the model’s classification results and the actual values. The confusion matrix structure is in the paper [27].

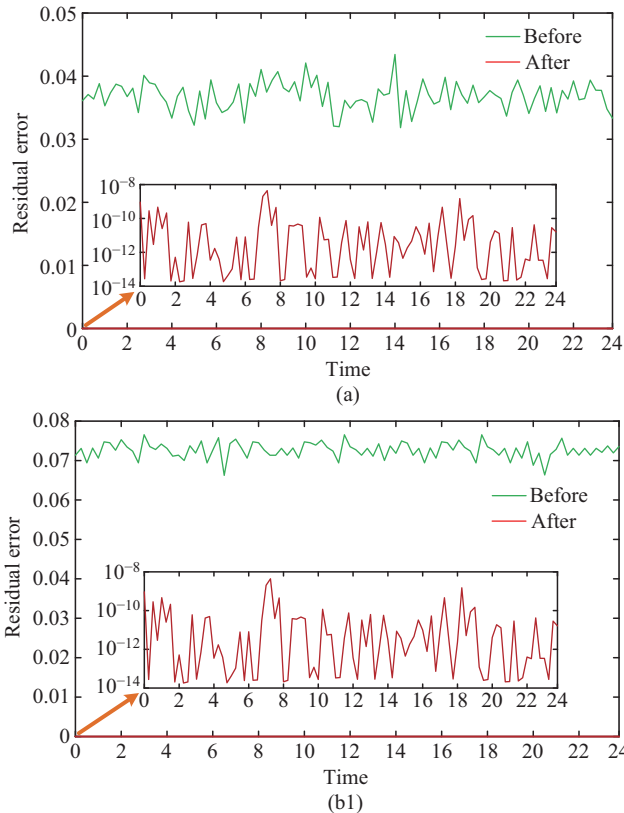


Fig. 8. The largest normalized residual before and after. (a) The largest normalized residual before and after the attack on the 14bus system. (b1) The largest normalized residual before and after the attack on the 57bus system. (b2) The largest normalized residual before and after the attack on the 118 bus system.

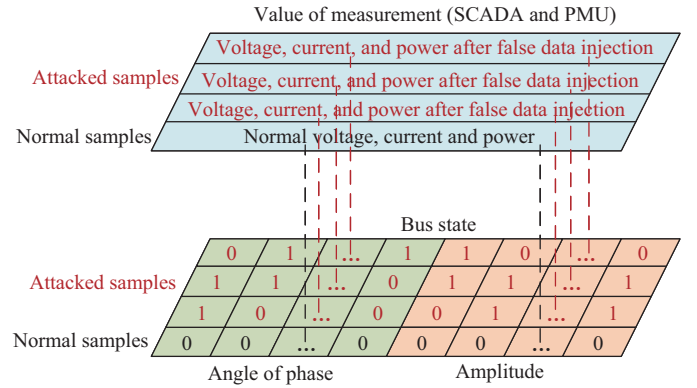


Fig. 9. Principle of FDIA localization detection.

C. Example I: AC Attack Detection for the Test System of the IEEE-14 Bus with Complete Information

Firstly, the performance of the proposed algorithm for FDIA in special cases is examined by considering an attack on the power grid, where the attacker has complete information.

By implementing the overload attack on the IEEE-14 bus system, the detection performance of RVFLN^{ML}, RVFLN, ELM, and SVM is compared, as shown in Fig. 10. The abscissa represents the state variables, arranged in phase angle order first and amplitude order after bus order. Because the amount of the state of the reference bus remains unchanged before and after the attack, the two labels corresponding to the reference bus are deleted, totaling 26 labels. It can be seen that the detection accuracy of SVM is the lowest and the positioning ability fluctuates greatly. Meanwhile, the lowest accuracy is only 80.85%. The detection accuracy of ELM is improved compared with SVM, and the average accuracy reaches 90.27%. However, it also exhibits significant fluctuations, with the lowest accuracy reaching only 84.68%.

The input layer and the output layer are directly connected, and the original features are also used as the input of the output layer. Due to the addition of the linear link processing and the consideration of the original features’ weights in the classification task, the detection accuracy of RVFLN is signif-

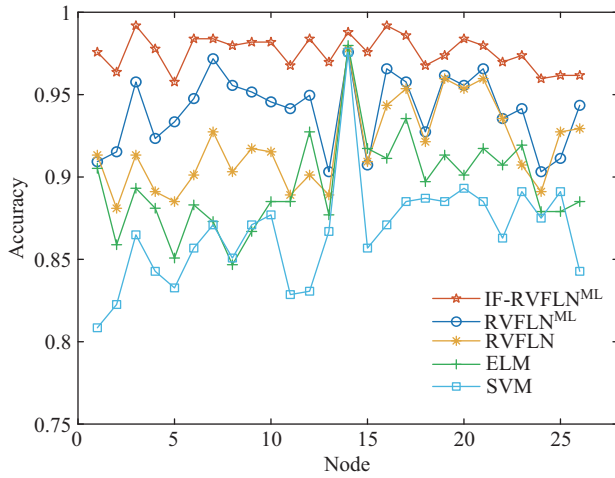


Fig. 10. Location detection accuracy for each bus.

icantly improved. The average accuracy reaches 92.49%, and the fluctuation of the detection rate is significantly reduced. After RVFLN is expanded into a multi-layer deep network, the detection accuracy is significantly improved due to its stronger feature extraction ability, and the average detection accuracy of $RVFLN^{ML}$ reaches 94.49%. After the IF is introduced into $RVFLN^{ML}$, it significantly enhances the ability to deal with uncertain data and compensates for the lack of sensitivity to noise data in $RVFLN^{ML}$. The average detection accuracy of $IF-RVFLN^{ML}$ reaches 97.76%, and the detection accuracy for each bus state remains stable, exceeding 95.77%.

Additionally, the method for calculating Accuracy, precision, F1 score, recall, and area under the curve (AUC) is described in the paper [27]. The values and results of each metric are shown in Fig. 11.

D. Example II: AC attack detection for the test systems of IEEE-57 and IEEE-118 buses with partial information

Because the complexity of the network structure in the test systems of IEEE-57 and IEEE-118 buses is much larger than that of the IEEE-14 bus system, it is more realistic for the attacker to grasp the global information of the network. Therefore, the example studies the attacker in FDIA detection with incomplete information about the power grid and selects a specific attack area to overload the power grid and selects a specific attack area to overload the line in the test systems of IEEE-57 and IEEE-118 buses.

It shows the accuracy of each bus detection in the 57-bus system in Fig. 12. The results show that the average accuracy of the five algorithms in the system of IEEE-57 bus is 96.10%, 93.83%, 91.70%, 86.15% and 84.89%, respectively, and the advantage of $IF-RVFLN^{ML}$ is more obvious. The reason is that there is a direct connection between the input layer and the output layer in the network structure, which prevents the loss of original feature information after other detection algorithms reduce the dimension of the original feature multiple times through the hidden layer neurons when dealing with high-dimensional data. Compared with the large fluctuations in the detection accuracy of other detection algorithms for different buses, the detection accuracy of $IF-RVFLN^{ML}$ for each bus

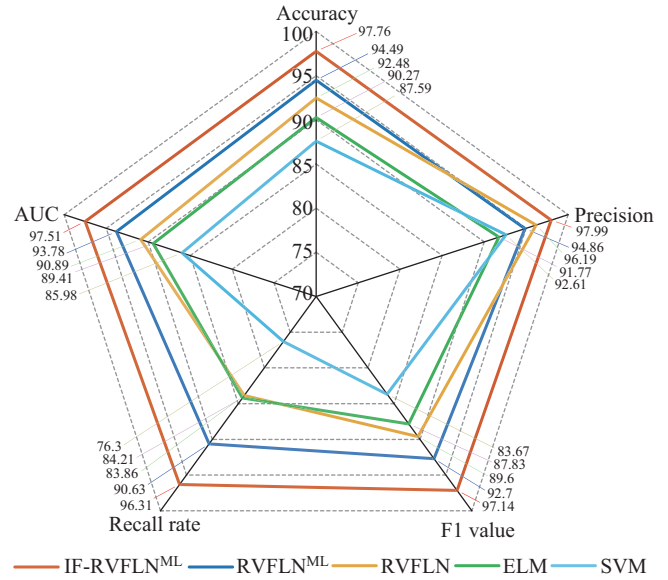


Fig. 11. Location detection index results.

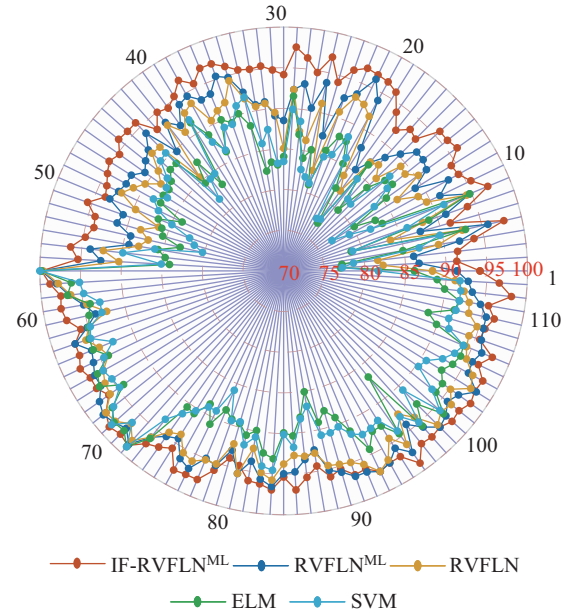


Fig. 12. Location detection accuracy of the IEEE-57.

exceeds 90%, which significantly reduces the influence of noise data on the detection results.

The results of the IEEE-118 bus test system are shown in Fig. 13, and the average detection accuracy of the five algorithms is 99.50%, 99.05%, 98.84%, 97.92%, and 97.64%, respectively, all of which are above 97%. This is because the network structure of the IEEE-118 bus is larger, and the selected line in a specific area makes it overloaded. The bus with low correlation to the overloaded line is almost unaffected by it, compared to the global impact of the network, and the number of unpolluted buses is far greater than that of polluted buses. The undetected attack samples will continue to expand the influence of FDIA on the power grid and even cause cascading failures. The false detection of normal samples will

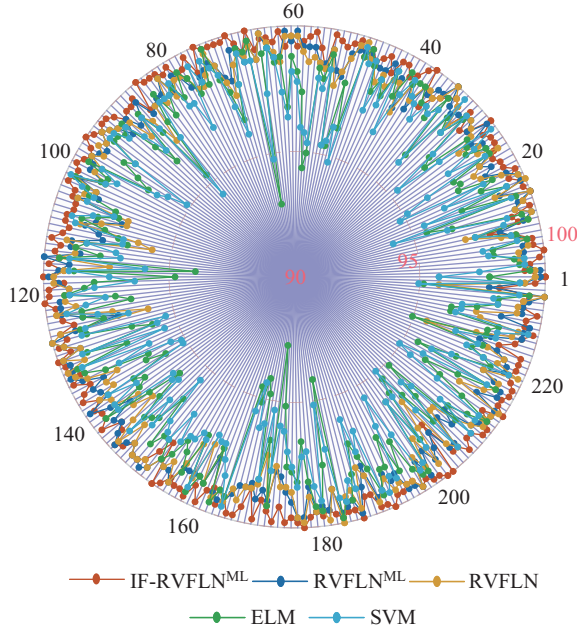


Fig. 13. Location detection accuracy of the IEEE-118.

eliminate the normal lines and disrupt the normal operation of society. The accuracy rate and recall rate focus on false detection and missing detection respectively. The F1 value is the weighted average of the two, and cases of missed detection and false detection can be considered simultaneously. The larger the F1 value is, the fewer the cases of missed detection and correct detection. Therefore, in the face of the unbalanced sample of 118 buses, the F1 value is a more appropriate evaluation index.

The heat map of each evaluation index in the test system of IEEE-57 bus is shown in Fig. 14. The recall rate of the IF-RVFLN^{ML} method is 5.97%, 11.26%, 22.76% and 25.63% higher than that of RVFLN^{ML}, RVFLN, ELM and SVM, respectively. It can be seen that the proposed method can better identify and accurately locate the contaminated state quantity, thereby enabling the precise positioning of FDIA.

The detection evaluation of the different algorithm metrics in the IEEE-118 bus test system is shown in Fig. 15. Notably, the recall rate of SVM, which has the worst detection effect, is only 70.71%, a 22.67% difference from IF-RVFLN^{ML}. Meanwhile, the detection accuracy of SVM at each bus fluctuates more obviously, indicating that the algorithm has poor learning ability for the positive samples of the IEEE-118 bus. Compared to ELM with RVFLN, the indices of the RVFLN^{ML} structure are greatly improved. Still, the latter is sensitive to noise data, and the classification ability of the algorithm can be significantly improved with the powerful ability of IF to handle noise data. The results show that IF-RVFLN^{ML} has the most stable accuracy in positioning detection, and its precision, F1 value, and recall rate are better than those of other algorithms, reaching 96.87%, 95.09%, and 93.38%, respectively. Compared to the value before integrating IF, the F1 value has increased by more than 4%.

The curve of the Receiver Operating Characteristic (ROC) represents the relative relationship between the true positive

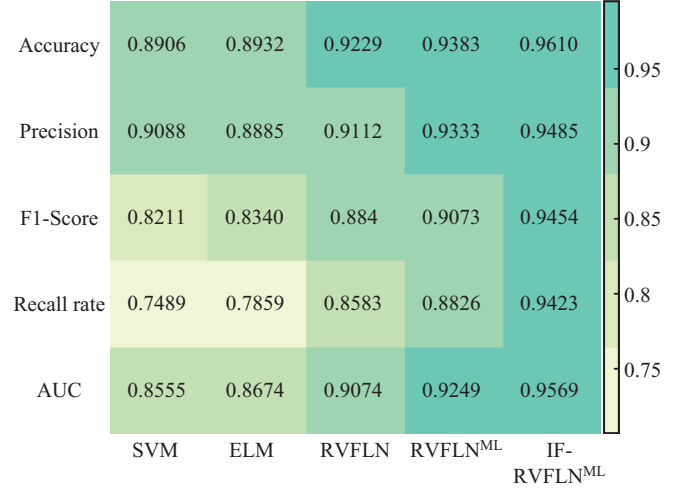


Fig. 14. Location detection index results of IEEE-57.

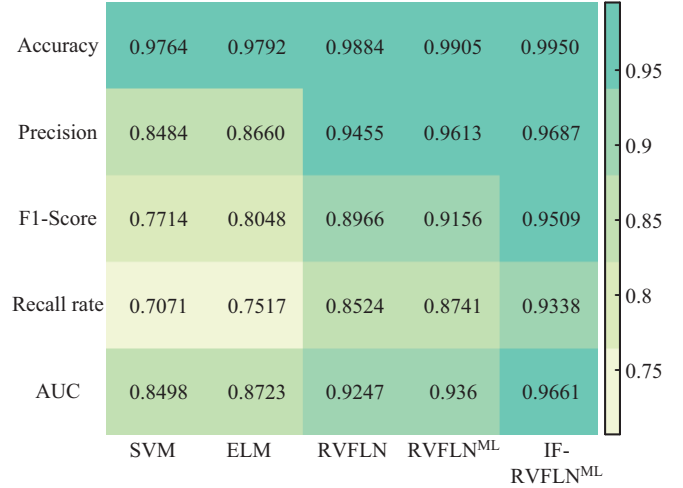


Fig. 15. Location detection index results of IEEE-118.

and false positive rates. The AUC is the area enclosed by the ROC curve and the coordinate axis, and the closer the AUC value is to 1, the higher the detection rate of the attack point.

The ROC curves of the IEEE-57 and IEEE-118 bus test systems are shown in Figs. 16 and 17. The AUC values of the five algorithms in the IEEE-57 bus test system are 0.9569, 0.9249, 0.9074, 0.8674 and 0.8555, respectively. The AUC values of the IEEE-118 bus test system are 0.9661, 0.9360, 0.9247, 0.8723 and 0.8498, respectively. The algorithm IF-RVFLN^{ML} can accurately identify the contaminated state quantity in the case of a low false alarm rate, which verifies the superiority of its FDIA positioning detection.

E. Algorithm Robustness Analysis

In order to further verify the robustness advantage of IF-RVFLN^{ML}, IEEE-14 and IEEE-57 bus test systems are selected in this paper (the 118-bus test system in this paper uses extremely unbalanced measurement data, and the experimental results are not highly comparable, so the selection is not done). Five standard noise levels of 5%, 10%, 15%, 20%, and 25% were added to the measured data to analyze and test the anti-

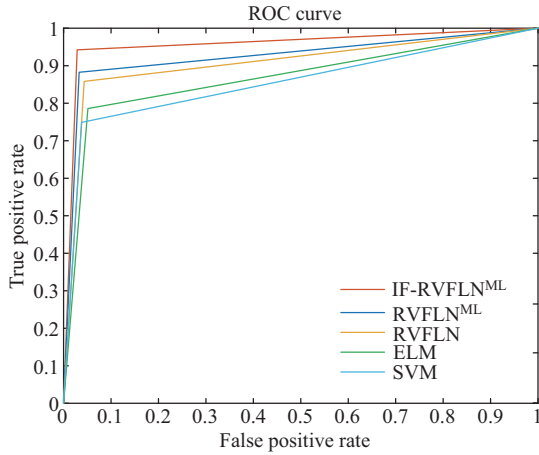


Fig. 16. ROC curve of the IEEE-57.

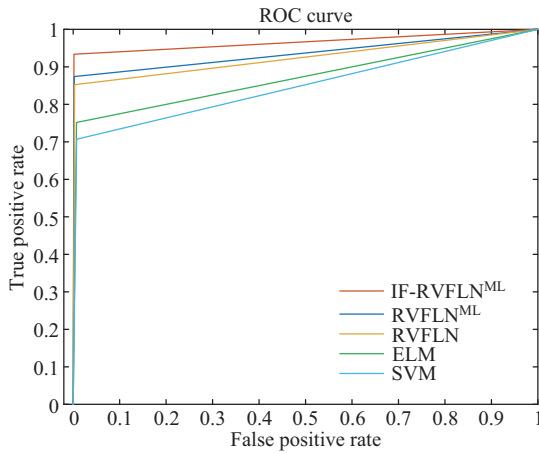


Fig. 17. ROC curve of the IEEE-118.

interference ability of the algorithm.

The results are shown in Figs. 18 and 19. In the IEEE-14 bus test system, the false positive rate of IF-RVFLN^{ML} increases slowly with the increase in noise level. Even when the noise intensity reaches 25%, the false positive rate is still lower than 0.03. When other algorithms add measurement noise, the false detection rate is greatly affected. Taking ELM as an example, when the noise standard is 5%, the attack false positive rate is 0.0297. As the noise level increases, the false positive rate increases significantly, reaching a peak of 0.0736. With increasing noise levels, ELM struggles to differentiate between normal and abnormal measurements after interference compared to IF-RVFLN^{ML}. Consequently, a higher proportion of attacked data is misclassified as normal, and some attacked data remains undetected, thereby significantly reducing FDIA's efficiency in detecting and locating attacks, which poses a greater threat to power systems. In the false positive rate test of the IEEE-57 bus test system, the performance of ELM and SVM algorithms deteriorates further when measuring more complex bus test systems due to their simple structures. However, under different levels of standard noise interference, IF-RVFLN^{ML} can still maintain a lower false positive rate compared to other algorithms, which verifies its superior noise resistance capability and stronger positioning performance.

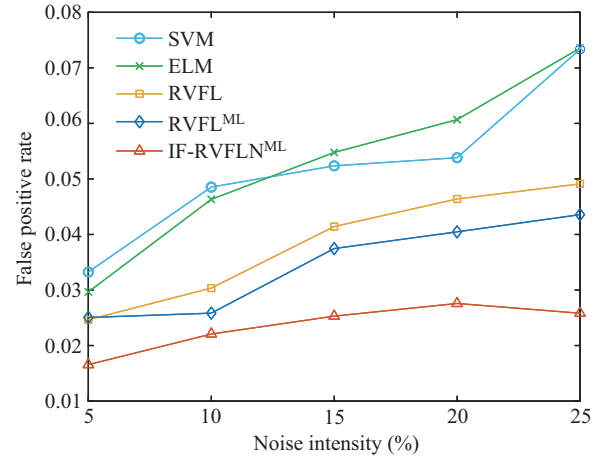


Fig. 18. False positive rate curve of the IEEE-14.

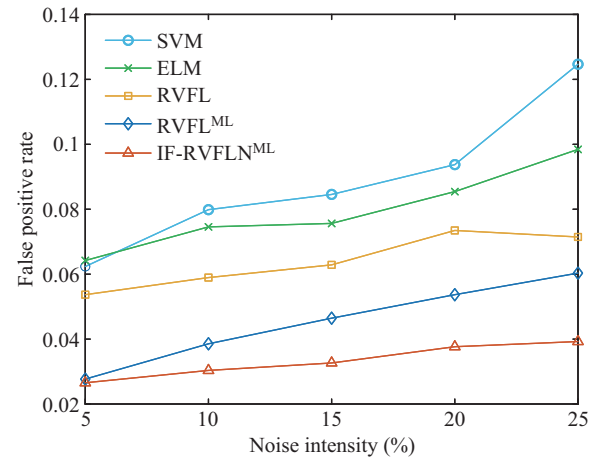


Fig. 19. False positive rate curve of the IEEE-57.

V. CONCLUSION

To address the limitation of existing FDIA detection methods in effectively locating FDIA in CPPS, a novel FDIA detection method based on the IF-RVFLN^{ML} algorithm is proposed in this paper. By extending the single-hidden-layer RVFLN to multiple hidden layers, the RVFLN^{ML} is constructed for feature extraction and classification training, which addresses the problem of its limited feature expression ability. The noise in the measurement data will significantly reduce the localization detection effect, so IF is introduced to mitigate the impact of noise on the detection performance of RVFLN^{ML}. Finally, IF-RVFLN^{ML} is constructed to locate and detect the attacked state.

A significant number of simulations are conducted on the test systems of IEEE-14, IEEE-57 and IEEE-118 buses to analyze the detection performance of the proposed method under the different attack scenarios, and the effectiveness of the proposed method is verified. Compared to the algorithms of RVFLN^{ML}, RVFLN, ELM, and SVM, it is verified that the performance of IF-RVFLN^{ML} is significantly better than that of other algorithms in terms of accuracy, precision, and recall rate. Moreover, it has a more accurate positioning detection performance. The proposed method belongs to the detection

scheme based on data-driven, which can automatically extract sample features to detect the abnormal data without obtaining the model and parameters of the system, and it has strong scalability.

REFERENCES

- [1] M. Ni, M. L. Li, J. Li, Y. J. Wu, and Q. Wang, "Concept and research framework for coordinated situation awareness and active defense of cyber-physical power systems against cyber-attacks," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 3, pp. 477–484, May 2021.
- [2] Y. J. Wu, H. Xu, and M. Ni, "Defensive resource allocation method for improving survivability of communication and information system in CPPS against cyber-attacks," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 750–759, Jul. 2020.
- [3] H. Z. Wang, J. Q. Ruan, B. Zhou, C. B. Li, Q. W. Wu, M. Q. Raza, and G. Z. Cao, "Dynamic data injection attack detection of cyber physical power systems with uncertainties," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 10, pp. 5505–5518, Oct. 2019.
- [4] Z. C. Dong, M. Tian, and L. Ding, "A framework for modeling and structural vulnerability analysis of spatial cyber-physical power systems from an attack–defense perspective," *IEEE Systems Journal*, vol. 15, no. 1, pp. 1369–1380, Mar. 2021.
- [5] P. Y. Li, J. Fu, K. G. Xie, B. Hu, Y. Wang, C. Z. Shao, Y. Sun, and W. Huang, "A defense planning model for a power system against coordinated cyber-physical attack," *Protection and Control of Modern Power Systems*, vol. 9, no. 5, pp. 84–95, Sep. 2024.
- [6] M. Tian, Z. C. Dong, and X. P. Wang, "Analysis of false data injection attacks in power systems: a dynamic Bayesian game-theoretic approach," *ISA Transactions*, vol. 115, pp. 108–123, Sep. 2021.
- [7] Z. J. Liu, Q. Wang, Y. J. Ye, and Y. Tang, "A GAN-based data injection attack method on data-driven strategies in power systems," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 3203–3213, Jul. 2022.
- [8] N. Yi, Q. Wang, L. M. Yan, Y. Tang, and J. J. Xu, "A multi-stage game model for the false data injection attack from attacker's perspective," *Sustainable Energy, Grids and Networks*, vol. 28, pp. 100541, Dec. 2021.
- [9] Y. J. Wu, Y. T. Ru, Z. W. Lin, C. J. Liu, T. Xue, X. Zhao, and J. F. Chen, "Research on cyber attacks and defensive measures of power communication network," *IEEE Internet of Things Journal*, vol. 10, no. 9, pp. 7613–7635, May 2023.
- [10] B. J. Yan, Z. Z. Jiang, P. C. Yao, Q. Yang, W. Li, and A. Y. Zomaya, "Game theory based optimal defensive resources allocation with incomplete information in cyber-physical power systems against false data injection attacks," *Protection and Control of Modern Power Systems*, vol. 9, no. 2, pp. 115–127, Mar. 2024.
- [11] F. Aciad, W. Z. Gao, and J. Momoh, "Bad data detection for smart grid state estimation," in *2016 North American Power Symposium (NAPS)*, 2016, pp. 1–6.
- [12] H. Z. Wang, J. Q. Ruan, G. B. Wang, B. Zhou, Y. T. Liu, X. Q. Fu, and J. C. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4766–4778, Nov. 2018.
- [13] J. Q. Yan, F. H. Guo, and C. Y. Wen, "False data injection against state estimation in power systems with multiple cooperative attackers," *ISA Transactions*, vol. 101, pp. 225–233, Jun. 2020.
- [14] H. Y. Wu, Z. Xu, J. Q. Ruan, and X. Z. Sun, "PFL-DSSE: a personalized federated learning approach for distribution system state estimation," *CSEE Journal of Power and Energy Systems*, vol. 10, no. 5, pp. 2265–2270, Sep. 2024.
- [15] Y. Chen, S. W. Huang, F. Liu, Z. S. Wang, and X. W. Sun, "Evaluation of reinforcement learning-based false data injection attack to automatic voltage control," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2158–2169, Mar. 2019.
- [16] A. S. Musleh, G. Chen, and Z. Y. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2218–2234, May 2020.
- [17] B. Y. Chen, H. B. Li, and B. Zhou, "Real-time identification of false data injection attacks: a novel dynamic-static parallel state estimation based mechanism," *IEEE Access*, vol. 7, pp. 95812–95824, Jul. 2019.
- [18] Y. Z. Yang, W. X. Liu, C. Z. Li, G. M. Liu, S. Zhang, and Y. W. Zhang, "Review of FDIA detection methods for electric power SCADA system," *Proceedings of the CSEE*, vol. 43, no. 22, pp. 8602–8621, Nov. 2023.
- [19] Z. H. Liu, L. Xi, Y. Quan, and C. Cheng, "Brain-inspired collaborative power dynamic dispatch for the digital twin smart generation system," *IEEE Transactions on Industry Applications*, vol. 61, no. 1, pp. 1420–1430, Jan./Feb. 2025.
- [20] Z. H. Liu, L. Xi, Y. Quan, C. Cheng, and J. Z. Zhu, "Brain-inspired collaborative automatic generation control with large-scale electric vehicles integration," *IEEE Transactions on Sustainable Energy*, vol. 16, no. 2, pp. 880–891, Apr. 2025.
- [21] Y. B. Zhou, J. An, G. Mu, and Y. Shi, "An improved constraint-inference approach for causality exploration of power system transient stability," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 59, Nov. 2023.
- [22] B. R. Chen, Q. H. Wu, M. S. Li, and K. S. Xiahou, "Detection of false data injection attacks on power systems using graph edge-conditioned convolutional networks," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 16, Apr. 2023.
- [23] M. Esmalifalak, L. C. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.
- [24] Y. Wu, Q. K. Wang, N. W. Guo, Y. J. Tian, F. Y. Li, and X. J. Su, "Efficient multi-source self-attention data fusion for FDIA detection in smart grid," *Symmetry*, vol. 15, no. 5, pp. 1019, May 2023.
- [25] X. F. Yin, Y. M. Zhu, and J. K. Hu, "A subgrid-oriented privacy-preserving microservice framework based on deep neural network for false data injection attack detection in smart grids," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1957–1967, Mar. 2022.
- [26] T. Wu, W. L. Xue, H. Z. Wang, C. Y. Chung, G. B. Wang, J. C. Peng, and Q. Yang, "Extreme learning machine-based state reconstruction for automatic attack filtering in cyber physical power system," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1892–1904, Mar. 2021.
- [27] L. Xi, M. He, B. Q. Zhou, and Y. Y. Li, "Research on false data injection attack detection in power system based on improved multi layer extreme learning machine," *Acta Automatica Sinica*, vol. 49, no. 4, pp. 881–890, Apr. 2023.
- [28] P. Zhou, Y. Jiang, C. Y. Wen, and X. W. Dai, "Improved incremental RVFL with compact structure and its application in quality prediction of blast furnace," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8324–8334, Dec. 2021.
- [29] S. P. Guan, H. H. Huang, and T. Y. Guan, "Quality evaluation of single-crystal wafer X-ray detection using random vector functional-link network," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 6010411, Sep. 2023.
- [30] H. Aggarwal, H. D. Arora, and V. Kumar, "A decision-making problem as an applications of intuitionistic fuzzy set," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 2, pp. 5259–5261, Dec. 2019.
- [31] R. Fu, X. J. Huang, Y. S. Xue, Y. J. Wu, Y. Tang, and D. Yue, "Security assessment for cyber physical distribution power system under intrusion attacks," *IEEE Access*, vol. 7, pp. 75615–75628, Jul. 2019.
- [32] Q. Yang, L. Jiang, W. J. Hao, B. Zhou, P. Yang, and Z. H. Lv, "PMU placement in electric transmission networks for reliable state estimation against false data injection attacks," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1978–1986, Dec. 2017.
- [33] S. Y. Wang, S. Z. Bi, and Y. J. A. Zhang, "Locational detection of the false data injection attack in a smart grid: a multilabel classification approach," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8218–8227, Sep. 2020.
- [34] B. R. Chen, Q. H. Wu, M. S. Li, and K. S. Xiahou, "Detection of false data injection attacks on power systems using graph edge-conditioned convolutional networks," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 16, Apr. 2023.
- [35] Q. S. Shi, R. Katuwal, P. N. Suganthan, and M. Tanveer, "Random vector functional link neural network based ensemble deep learning," *Pattern Recognition*, vol. 117, pp. 107978, Sep. 2021.



Lei Xi received the M.S. degree in Control Theory and Control Engineering from Harbin University of Science and Technology, and the Ph.D. degree in Electrical Engineering from South China University of Technology, China, in 2016. Currently, he is a professor in the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include cyber physical security of smart grid, load frequency control, artificial intelligence techniques and automatic generation control.



Zihao Li is currently pursuing his M.S. degree at the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include cyber physical security of smart grid and artificial intelligence techniques.



Fangyan Bai is currently pursuing his M.S. degree at the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include cyber physical security of smart grid and artificial intelligence techniques.



Zongze Li is currently pursuing his M.S. degree at the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include cyber physical security of smart grid and artificial intelligence techniques.



Yixiao Wang is currently pursuing her M.S. degree at the College of Electrical Engineering and New Energy, China Three Gorges University. Her research interests include cyber physical security of smart grid and artificial intelligence techniques.



Hongjun Chen is currently pursuing his M.S. degree at the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include cyber physical security of smart grid and artificial intelligence techniques.