TYPE Original Research
PUBLISHED 13 November 2023
DOI 10.3389/fenrg.2023.1285317

Check for updates

OPEN ACCESS

EDITED BY
Nantian Huang,
Northeast Electric Power University,
China

REVIEWED BY
Yuanshi Zhang,
Southeast University, China
Tao Zhou,
Nanjing University of Science and
Technology, China

*CORRESPONDENCE
Ren Liu,
✉ liuren248@gmail.com

RECEIVED 29 August 2023
ACCEPTED 27 October 2023
PUBLISHED 13 November 2023

CITATION
Zhu Y, Liu R, Chang D and Guo H (2023), Detection of false data injection attacks on power systems based on measurement-eigenvalue residual similarity test.
Front. Energy Res. 11:1285317.
doi: 10.3389/fenrg.2023.1285317

Detection of false data injection attacks on power systems based on measurement-eigenvalue residual similarity test

Yihua Zhu[1,2], Ren Liu[1,2]*, Dongxu Chang[3,4] and Hengdao Guo[3,4]

[1]State Key Laboratory of HVDC, Electric Power Research Institute, China Southern Power Grid, Guangzhou, China, [2]National Energy Power Grid Technology RD Centre, Guangzhou, China, [3]Guangdong Provincial Key Laboratory of Intelligent Operation and Control for New Energy Power System, Guangzhou, China, [4]CSG Key Laboratory for Power System Simulation, Electric Power Research Institute, China Southern Power Grid, Guangzhou, China

Existing False data injection attack (FDIA) detection methods based on measurement similarity testing have difficulty in distinguishing between actual power grid accidents and FDIAs. Therefore, this paper proposes a detection method called the measurement-eigenvalue residual similarity (MERS) test, which can accurately detect FDIAs in AC state estimationof power system and effectively distinguish them from actual power grid accidents. Simulation results on the IEEE 39-bus system demonstrate that the proposed method achieves higher detection rates and lower false alarm rates than traditional methods under various operation conditions.

KEYWORDS

false data injection attacks, AC state estimation, measurement-eigenvalue residual similarity, bad data detection, cyber security

1 Introduction

State estimationis an important component of supervisory control and data acquisition (SCADA) systems in power systems. It can accurately estimate the true operating conditions of the system based on the real-time operating status information and network topology information collected by remote terminal units (RTUs), providing a foundation for system analysis such as optimal power flow and bad data detection (Ericsson, 2010; Feng et al., 2023).

False data injection attack (FDIA) is a type of network attack on power system state estimation, which was first proposed by Liu et al. (2011). The paper considers two different attack scenarios: one assumes that the attacker has complete knowledge of the system parameters, while the other assumes that the attacker is limited by the resources required to damage certain specific instruments. Subsequently, many scholars have conducted extensive research on FDIA.

In the early stage, to simplify the computational difficulty and improve the accuracy of FDIA detection, scholars usually used FDIA based on the DC power flow model for experiments. However, in practice, the power industry has adopted a nonlinear state estimator with a complete AC power flow model (Abur and Exposito, 2004; Xiahou et al., 2020; Chen et al., 2023), so it cannot be guaranteed that the FDIA detection method based on the DC model has equivalent performance under the nonlinear AC state estimation (Zhao et al., 2018; Tiande et al., 2023). Therefore, to solve the limitations of the DC model, a large number of scholars have started researching on FDIA detection and prevention based on the AC power flow model (Chaojun et al., 2015; He et al., 2017; Zhang Y. et al., 2020). Currently, the research on FDIA detection technology can be roughly divided into three categories: detection methods based on state estimation (Coluccia et al., 2013; Chen et al., 2020; Liu et al., 2022), trajectory prediction (Yang et al., 2019), and artificial intelligence (Jena et al., 2021). FDIA also poses a significant threat to the security of the electricity market, and research has developed detection methods based on the critical load level (CLL) concept to help operators identify risk periods (Zhang Q. et al., 2020). There is also research on the monitoring and response mechanisms in learning-based collaborative FDIA, proposing a real-time FDIA detection method and a re-evaluation mechanism with penalty schemes, which can detect the initiation of learning-based collaborative FDIA (Jiang et al., 2020). In response to the characteristics of power grid systems with large data redundancy, some scholars have proposed a multi-source self-attention data fusion model, which uses time alignment technology to integrate the collected multi-source remote sensing data into the same time dimension. Subsequently, a symmetric hybrid deep network model was established to detect FDIA by symmetrically combining long and short durations (Wu et al., 2023).

However, these methods have limitations. For example, for methods that rely on secure PMUs, the fundamental assumption that a specific set of PMUs is completely invincible may not always hold true (Zhao et al., 2018). In addition, for machine learning-based detection methods, if false data does not significantly deviate from historical trends, the detection performance may not be satisfactory. As for the measurement consistency verification, how to accurately distinguish FDIA from actual power accidents is still an urgent problem to be solved.

In order to deal with the aforementioned problems, this paper proposes a detection method called the measurement-eigenvalue residual similarity (MERS) test. The rest of this paper is structured as follows: Section 2 introduces the construction principles of FDIA and the basis and methods for constructing imperfect FDIA under practical conditions Section 3 elaborates in detail on the FDIA detection method proposed in this article Section 4 conducted simulation verification on the content of Section 2 and Section 3. Section 5 is the conclusion of the paper.

2 FDIA constrction strategies

2.1 Bad data detection module

During the entire process of collection, transmission, and exchange of measurement data, it is impossible to ensure that all data is accurate, but there may be data damage or deviation caused by equipment or network reasons. Therefore, it is necessary to detect and identify measurement data, using a large number of normal redundant measurements and mathematical processing methods to detect defective data. Based on the results of WLS SE, conventional bad data detection methods such as chi-square test, weighted residual (WR) test, and largest normalized residual (LNR) test are commonly used.

The LNR test is a widely used method for detecting bad data. The core is to normalize the measurement residuals.

$$r_{Ni} = \frac{r_i}{\sigma_{Ni}} = \frac{|z_i - h_i(\hat{x})|}{\sqrt{\Omega_{ii}}}\tilde{N}(0,1) \tag{1}$$

$$\Omega = \left[I - H\left(H^T R^{-1} H\right)^{-1} H^T R^{-1}\right] R \tag{2}$$

where z_i is the i-th measurement; $h(\cdot)$ is a nonlinear function related to the i-th measurement value and the state vector; r_{Ni} is the i-th normalized residual; Ω_{ii} is the i-th diagonal element of Ω; $I \in R^{m\times m}$ is an identity matrix. If the maximum normalized residual is greater than the threshold ε, the corresponding measurement results will be suspected as bad data.

For the i-th normalized residual $r_{Ni} \sim N(0, \Omega_{ii})$ according to the 3σ rule, there is $P\left\{|r_{Ni}| < 3\sqrt{\Omega_{ii}}\right\} = 0.9975$. If $\sqrt{\Omega_{ii}} \approx 1$, when the maximum normalized residual of the test $|r_{Ni}|$ is greater than the threshold (which is 3 at this time), it is considered that there is defective data in the measurement set, and the probability of false detection is $P_e = 0.0025$.

Compared to the chi-square test, the WR test is not affected by the system size compared to the LNR test. In terms of detection sensitivity, when the redundancy of equivalent measurement is high, both WR test and LNR test have excellent inspection results. Generally, the effectiveness of the LNR test is not inferior to that of the WR test. However, the LNR test requires complex calculations and incurs more computational costs. However, the LNR test method can identify which data is most likely to be bad, laying the foundation for identifying bad data.

The chi-square test, WR test, and LNR test are usually effective for detecting natural defect data, which can cause significant measurement residual disturbances. However, for maliciously designed false data, these methods are often ineffective and only cause small residual disturbances.

2.2 FDIA construction principle

The biggest difference between AC state estimation and DC state estimation is the nonlinearity of the problem. However, based on the AC state estimation model, the conclusion from DC state estimation can still be applied when launching a false data injection attack on a specific state variable-the measurement Jacobian matrix $H(x)$ reflects the electrical connection between the measurement and state variables. If a measurement z_i depends on a state variable x_k, then the element value in the k-th column and i-th row of $H(x)$ must be non-zero.

Therefore, in order to launch a false data injection attack in AC state estimation, the attacker must know the values of other state variables that are related to the target state variable at the current time, in addition to satisfying the two conditions for DC false data injection attacks. This further increases the difficulty of launching AC-based FDIA in practice.

Assuming that the attacker has access to real-time state estimation, network topology, and parameter information, they can launch a perfect FDIA, where the attack vector a is designed as follows

$$z + a - h(\hat{x} + c) = z - h(\hat{x}) \tag{3}$$

$$a = h(\widehat{x} + c) - h(\widehat{x}) \tag{4}$$

where z is the measurement vector; $\widehat{x}$ represents the state estimation vector that does not include the FDIA; c represents the bias vector injected into the state estimation; a denotes the attack vector; $h(\cdot)$ represents the nonlinear relationship between the state vector $\widehat{x}$ and the measurement vector z.

Reference (Liu et al., 2011) has shown that in the case where the state, parameter, and network information is completely leaked, the attack vector a constructed according to Eq. 4 can produce the same measurement residual vector as that without the attack. Therefore, false data can evade conventional detection methods such as chi-square, weighted residual, and likelihood ratio tests. However, it is difficult to achieve a perfect attack vector, and the specific reasons will be discussed in detail in the next section. Nevertheless, if the attacker can create an attack vector that is close enough to Eq. 4, such that the perturbation of the residual vector remains below the threshold of bad data detection, there is still a high chance that imperfect FDIA can bypass bad data detection and cause misjudgment and incorrect actions by the control center.

2.3 Imperfect FDIA construction strategy

Based on the analysis of the FDIA attack principle in the previous section, it can be inferred that it is difficult for network attackers to achieve perfect FDIA against AC state estimation. Therefore, this study is based on the assumption that AC-based FDIA is difficult to implement in reality, especially since attackers are unlikely to obtain accurate real-time state estimation and accurate model parameters, particularly real-time state estimation. The specific assumptions are as follows:

(1) Firstly, network parameters and state estimation are stored in the database of the control center, which is relatively well protected. For attackers, it is more difficult to enter the control center than the substation-level network due to communication or technical limitations.

(2) Even if the attacker gains accurate network parameters by invading the control center, it is still difficult to obtain accurate state estimation. Due to the real-time fluctuations of the load, the power flow is time-varying. Therefore, the latest state estimation is necessary for a perfect AC-based FDIA. However, once the FDIA is started, the state estimation of the control center will be biased.

(3) Attackers also need to consider the cost factor when designing the FDIA vector. Generally, the more perfect the FDIA requires modification of data, the greater the difficulty and cost of injecting the attack vector. Under cost constraints, attackers usually consider the minimum number of measurements to be tampered with to ensure that the attack is not detected, as well as how to choose these measurements. This also makes AC-based FDIA still imperfect.

(4) The attacker's state estimation function should be executed at the same time resolution (every 2–3 s) as the SCADA measurement, to determine how to operate the measurement perfectly. However, this is almost an impossible task for attackers because it depends on the availability of real-time measurements and powerful computing capabilities, which are not available to attackers. Note that it does not require an accurate latest estimation to start a perfect DC-based FDIA. However, almost all existing state estimators in the control center are based on AC, and it does require an accurate latest state estimation to start AC-based FDIA.

As it is almost impossible to meet all the above conditions simultaneously, it can be reasonably assumed that AC-based FDIA is imperfect. Consider an imperfect AC-based FDIA as follows:

$$\tilde{a} = \tilde{h}(\widehat{x} + \xi + c) - \tilde{h}(\widehat{x} + \xi) \tag{5}$$

where $\tilde{a}$ is the imperfect attack vector; $\tilde{h}(\cdot)$ is the nonlinear function of inaccurate model parameters; ξ represents the error in the attacker's guessed state estimation vector z.

The derivation of the measurement residual vector based on imperfect AC FDIA is given as follows

$$\begin{aligned} r_{\tilde{a}} &= z + \tilde{a} \\ &= z + \tilde{h}(\widehat{x} + \xi + c) - \tilde{h}(\widehat{x} + \xi) - \tilde{h}(\widehat{x}_{\tilde{a}}) \\ &= r + \tilde{h}(\widehat{x} + \xi + c) - \tilde{h}(\widehat{x} + \xi) + h(\widehat{x}) - \tilde{h}(\widehat{x}_{\tilde{a}}) \end{aligned} \tag{6}$$

where $x_{\tilde{a}}$ represent the state estimation vector; r and $r_{\tilde{a}}$ denote the measurement residual vector before and after FDIA.

The state estimation vector can be represented as

$$x_{\tilde{a}} = \widehat{x} + \eta + c \tag{7}$$

where η is the state estimation error vector of the system operator. Note that η and ξ are different because the model parameters obtained by the attacker are inaccurate.

Using to represent the measurement residual deviation vector caused by FDIA based on imperfect AC as:

$$\mathrm{r}_{\tilde{\mathrm{a}}} = r + \Delta \tag{8}$$

$$\Delta = \tilde{h}(\widehat{x} + \xi + c) - \tilde{h}(\widehat{x} + \xi) + h(\widehat{x}) - \tilde{h}(\widehat{x}_{\tilde{\mathrm{a}}}) \tag{9}$$

In this case, the probability density function of imperfect FDIA measurement residuals is different from that of measurement residuals without FDIA. Such subtle differences may not trigger traditional LNR or chi-square tests, but it is possible to design more specialized and refined tests to discover them.

3 Proposed FDIA detection method

3.1 Measurement value similarity

In this section, a cosine similarity ratio detection method is proposed. Specifically, the proposed method utilizes the ratio of the cosine similarity between the actual measured values and the state estimation values of the power system at two different time instants as the detection index. When the system is operating normally without FDIA, the cosine similarity ratio between the two will remain relatively stable. However, when the system is under FDIA, the cosine similarity ratio will experience a sudden change, which can be used to detect the presence of FDIA.

The FDIA detection method based on measurement values proposed in the article is based on state estimation measurement data. It is not fixed to individual or determined measurement values, but rather focuses on global measurement data, including active and reactive power measurements of power branches, voltage amplitude and phase angle measurements of generators or bus nodes, and generator output force measurements.

Firstly, in the process of state estimation, the estimated measurement value $\hat{z}_{k|k+1}$ at time k are obtained based on the historical state estimation value $\hat{x}_{k-1}$ at the previous time step, and then the actual system measurement value z_k at time k are obtained from the measurement equations. The cosine similarity between the actual and estimated measurement values at time k is

$$cos_k = \frac{z_k \cdot \hat{z}_{k|k-1}}{\|z_k\| \times \left\|\hat{z}_{k|k-1}\right\|} \tag{10}$$

Similarly, the cosine similarity between the actual and estimated measurement values at the initial time instant can be computed as

$$cos_1 = \frac{z_1 \cdot \hat{z}_{1|0}}{\|z_1\| \times \left\|\hat{z}_{1|0}\right\|} \tag{11}$$

where z_1 represents the measurement vector and $\hat{z}_{1|0}$ is the estimated vector at the initial time.

The measurement values at initial time 1 are chosen to be taken when the system reaches a steady-state condition. Under steady-state conditions, the variations in system variables such as voltage and current reach their minimum values, thus providing a more reliable baseline.

TABLE 1 Operation steps of cosine similarity ratio detection method.

Algorithm: Cosine Similarity Ratio Detection method
Input: Measured data available at time k in the system.
Step 1: Initialize the system state variable $\hat{x}_{k-1}$ and covariance matrix R_{k-1}.
Step 2: Perform state estimation.
Step 3: Predict the current state estimation value $\hat{x}_{k|k-1}$ based on the previous time step's state estimation value $\hat{x}_{k-1}$.
Step 4: Substitute the predicted state estimation value into the measurement equation to compute the predicted measurement value $\hat{z}_{k|k-1}$ at the current time step.
Step 5: Calculate the cosine similarity between the actual and predicted measurement values at the current time step and the initial time step.
Step 6: Compute the target detection function by taking the ratio of the cosine similarity values μ_k.
Step 7: Determine whether $\mu_k \notin [\lambda_1, \lambda_2]$ holds or not, where λ_1 and λ_2 are the set detection thresholds.
Step 8: If an FDIA attack is detected, the algorithm ends. Otherwise, $k = k + 1$ and return to Step 2.
Output: Detection of FDIA attack in the system.

If there are no FDIA in the system, the cosine similarity values should approach 1 due to the stable operation of the system. However, if FDIA occurs at time k in the power system, the injected attack sequence will immediately cause a measurement increment. The size of the actual measurement value z_k will experience a sudden change, which can be denoted by a variable b. Therefore, the new cosine similarity value can be expressed as

$$z_k^b = z_k + b \tag{12}$$

where z_k^b represents the measurement vector after false data injected attack at time k.

The cosine similarity between the actual and estimated measurement values after the FDIA can be calculated as

$$cos_k^b = \frac{(z_k + b) \cdot \hat{z}_{k|k-1}}{\|(z_k + b)\| \times \left\|\hat{z}_{k|k-1}\right\|} \tag{13}$$

where cos_k^b represents the cosine similarity between the actual (after FDIA) and estimated measurement at time k.

The value of μ_k^b caused by the FDIA can be calculated as follows

$$\mu_k^b = \frac{cos_k^b}{\cos_1} \tag{14}$$

where μ_k^b represents the ratio of measurement cosine similarity between time k (after FDIA) and initial time.

Assuming that the system is safe and reliable before time k, when an attack occurs and the attack vector is injected, it will cause a sudden change in the actual measurement value, leading to a change in the value of cos_k^b. Obviously, according to Eq. 14, the value of μ_k^b will deviate significantly from 1, which can be used to detect covert false data injection attacks.

The threshold for detecting FDIA can be set based on a large amount of historical data collected during normal operation of the system. The calculation formula for the threshold can be expressed as follows

$$\begin{cases} \lambda_1 = min\{\mu_1, \ldots, \mu_k, \cdots\} - l \\ \lambda_2 = max\{\mu_1, \ldots, \mu_k, \cdots\} + l \end{cases} \tag{15}$$

where l>0, $[\lambda_1, \lambda_2]$ represents the threshold margin of FDIA. If $\mu_k \notin [\lambda_1, \lambda_2]$, it is considered that the power system has been subjected to FDIA and a warning will be triggered.

The specific steps of the detection algorithm are shown in Table 1.

3.2 Eigenvalue residual similarity test

One of the limitations often raised in the FDIA detection method based on measurement data is that it cannot be distinguished from actual power grid accidents because actual power grid accidents can also cause measurement deviations, which may lead to false positives in the FDIA detection method based on measurement data, resulting in incorrect judgments by the control center.

When a power grid actual accident occurs, the measurement values will be perturbed and deviate from the normal level, changing from z_k to $z_{k'}$. According to Eq. 10, the cosine similarity between the current estimated measurement and the actual measurement is

$$cos_{k'} = \frac{z_{k'} \cdot \hat{z}_{k|k-1}}{\left\| z_{k'} \right\| \times \left\| \hat{z}_{k|k-1} \right\|} \tag{16}$$

$$\mu_{k'} = \frac{\cos_{k'}}{\cos_1} \tag{17}$$

where $cos_{k'}$ represents the cosine similarity between the actual (after actual power accident) and estimated measurement at time k; $\mu_{k'}$ denotes the ratio of measurement cosine similarity between time k (after actual power accident) and initial time.

Assuming that the system is safe and reliable before the power grid failure at time k, i.e., the system state estimation value $\widehat{x}_k$ at time k-1 is normal, and $\hat{z}_{k|k-1}$ is also a normal prediction. However, when an actual power grid failure occurs, it will cause a sudden change in the actual measurement value, and $cos_{k'}$ will also change accordingly. According to Eq. 17, the value of $\mu_{k'}$ will deviate sharply from 1.

Therefore, in Section 4, the cosine similarity ratio detection algorithm for FDIA detection was extended by adding an eigenvalue residual similarity test to distinguish between actual power accidents and FDIA.

As mentioned in Section 3.1, the construction of a perfect AC FDIA based on complete state information and network structure data can achieve high consistency in residual similarity. However, in the construction of imperfect AC FDIA, attackers cannot obtain exact real-time state estimation and model parameter data, resulting in defects in residual similarity of the constructed imperfect AC FDIA vector. However, due to the cost constraints of imperfect FDIA, it will choose to manipulate as little data as possible, thereby maintaining high similarity in some residual characteristic values.

At the same time, the measurement data anomalies caused by actual power grid accidents have randomness and irregularity, and the similarity between the measurement residual data calculated from abnormal measurement data and the measurement residual data without attack is very low, which is opposite to FDIA Therefore, in the FDIA detection scheme designed in the previous section, a n eigenvalue residual similarity test was added to effectively distinguish between FDIA and actual power grid accidents. The detection process is shown in Figure 1.

The data used for the eigenvalue residual similarity test is the largest normalized residual (LNR). The sensitivity of LNR to a

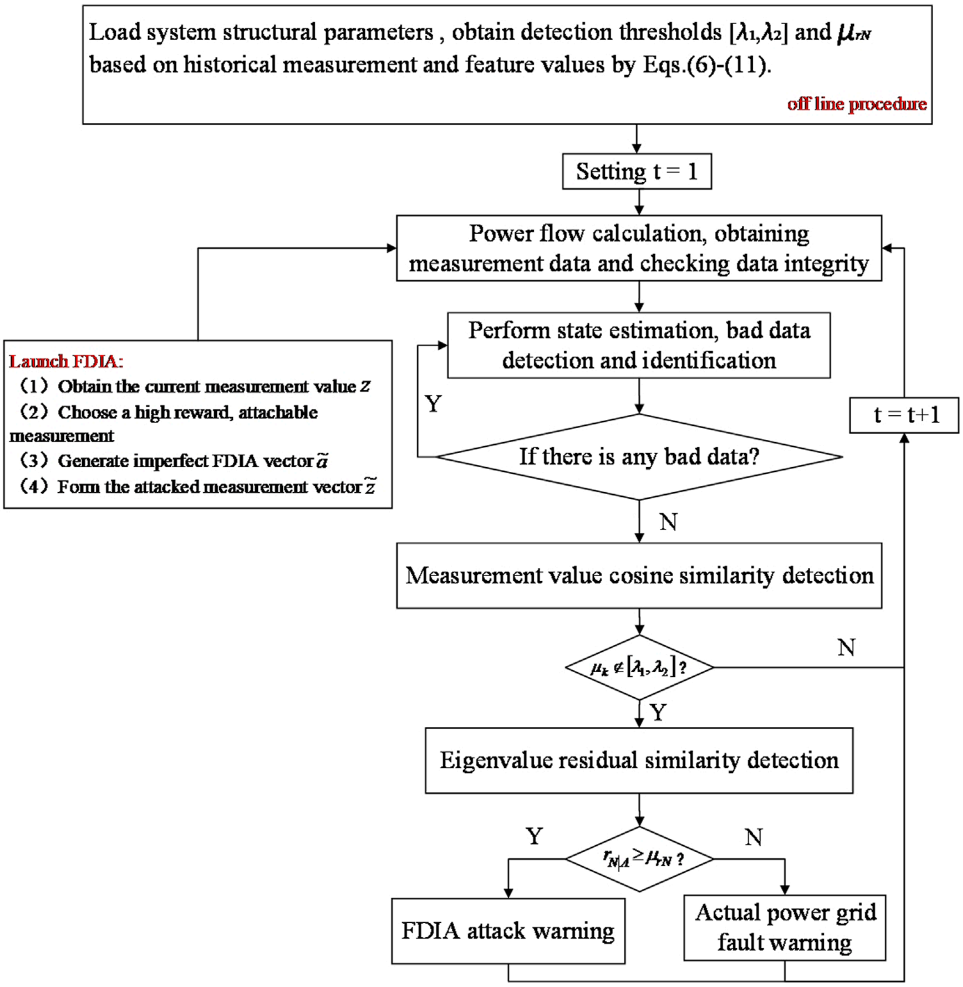
Load system structural parameters , obtain detection thresholds [λ1,λ2] and μrN based on historical measurement and feature values by Eqs.(6)-(11).
off line procedure
Setting t = 1
Power flow calculation, obtaining measurement data and checking data integrity
Perform state estimation, bad data detection and identification
Y
If there is any bad data?
N
t = t+1
Launch FDIA:
（1）Obtain the current measurement value z
（2）Choose a high reward, attachable measurement
（3）Generate imperfect FDIA vector ã
（4）Form the attacked measurement vector z̃
Measurement value cosine similarity detection
μk ∉ [λ1, λ2]?
N
Y
Eigenvalue residual similarity detection
Y
rN|A ≥ μrN ?
N
FDIA attack warning
Actual power grid fault warning

FIGURE 1
Flow chart of the MERS test.

single residual perturbation provides a good foundation for the identification of FDIA. The similarity algorithm used in the test is still the cosine similarity algorithm used in the detection algorithm, which is given as follows

$$cos_{r_N} = \frac{r_N \cdot r_{N|A}}{\|r_N\| \times \|r_{N|A}\|} \tag{18}$$

where r_N represents the eigenvalue LNR before the attack, $r_{N|a}$ represents the eigenvalue LNR after the attack, and cos_{rN} represents the similarity of the eigenvalue residual.

When an actual power grid failure occurs, due to the irregular and abnormal fluctuations of the measurement, the overall similarity of the characteristic values of $r_{N|a}$ and r_N is low, and cos_{rN} will deviate significantly from 1. However, when an FDIA occurs, some of the characteristic values of $r_{N|a}$ and r_N will have high similarity, and cos_{rN} will be very close to 1, which can serve as a metric for comparison with actual power grid accident.

4 Simulation results

4.1 Effects of false data injection attacks

To verify the effectiveness of the proposed detection algorithm proposed above, in this section, the matpower 7.0 simulation package is used to perform power flow calculation on the load data of the IEEE-39 bus standard test system. The single-line diagram of the test system is shown in Figure 2.

The data is processed as follows: First, perform power flow calculation on the existing system data to obtain the various measurement data of the current system. Secondly, simulate PMU measurement data, and add additive noise following a Gaussian distribution with a mean of 0 and a variance of 0.01 during the measurement process to simulate the influence of measurement random errors on the detection process. Finally, at a specified time, generate the attack vector for the specified measurement value using the method described in Section 2, and add it to the sensor measurement.

Firstly, the state estimation algorithm of the IEEE-39 bus system under normal operation should be validated. Taking the generator voltage amplitude and phase angle data as an example, the comparison between the estimated values and the actual values is shown in Figure 3. It can be seen that the state estimation algorithm has high accuracy in estimation performance.

Figure 4 shows the results of the bad data detection based on the LNR residual detection method. When the significance level is 0.05 and the degree of freedom is $K = m - n = 122 - 77 = 45$ (n is the number of states, m is the number of observations), the test p-value is 2.81. The blue line represents the residual distribution under normal conditions, the yellow line represents the residual

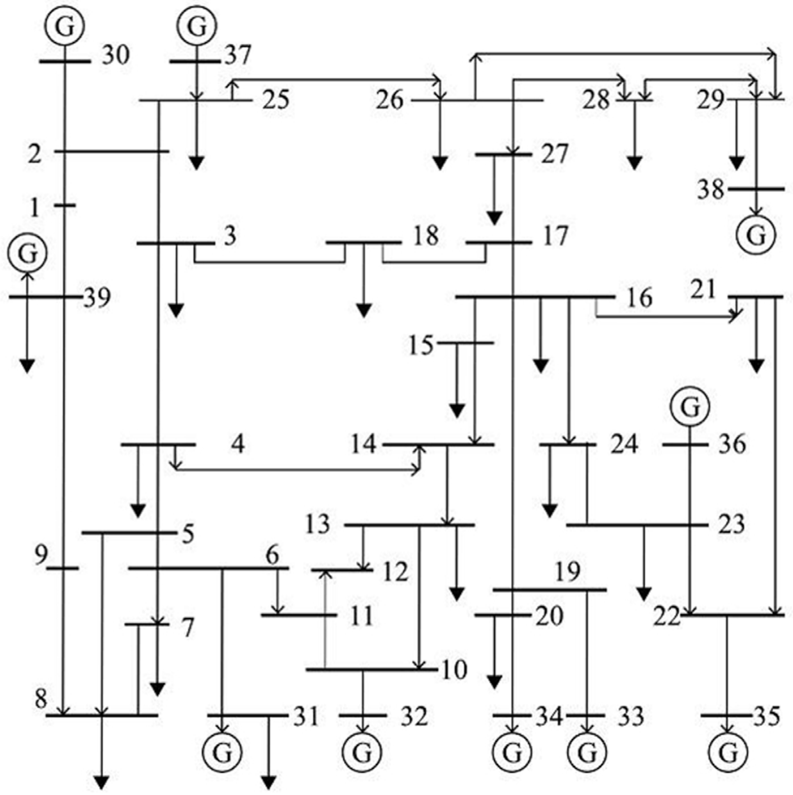

FIGURE 2
IEEE-39 bus test system single line diagram.

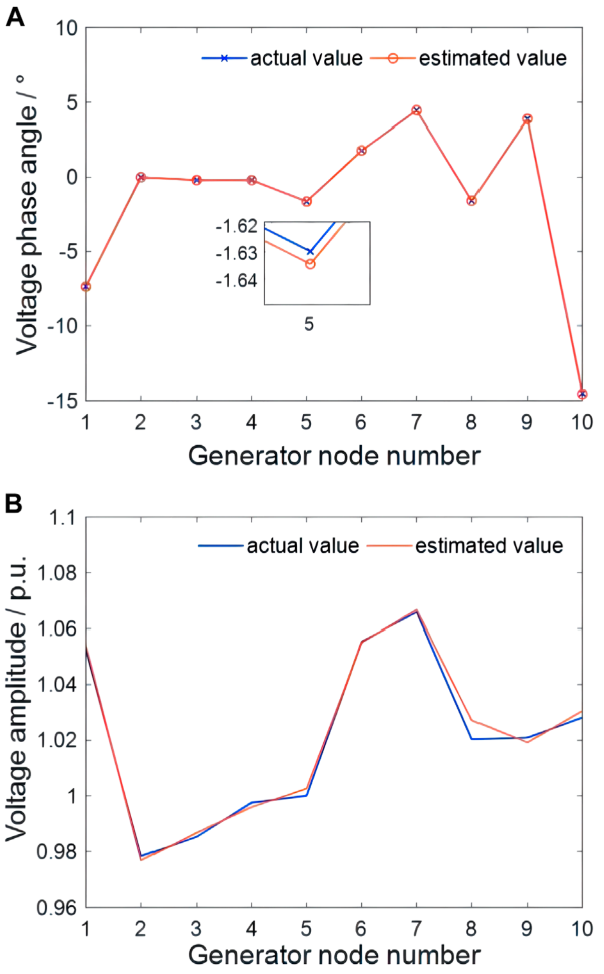

FIGURE 3
Accuracy test of **(A)** generator node voltage phase angle estimation and **(B)** voltage amplitude estimation.

distribution under the FDIA, and the red line represents the bad data detection threshold. It can be seen that although the carefully designed FDIA causes changes in the residual within a certain range, the fluctuation amplitude is not large and cannot exceed the test *p*-value. Therefore, the FDIA can bypass the bad data detection and identification process with a certain amount of redundancy, which indicates the reasonableness and feasibility of the FDIA setting in section III and provides assurance for subsequent research.

4.2 Effectiveness of proposed methods

Before verifying the effectiveness of the detection algorithm, in order to better approximate the actual operating state of the power grid, load fluctuation simulation was added to the experiment. The load fluctuated with a random probability of ±1% in each simulation, with the aim of improving the accuracy and robustness and making the detection algorithm more universal.

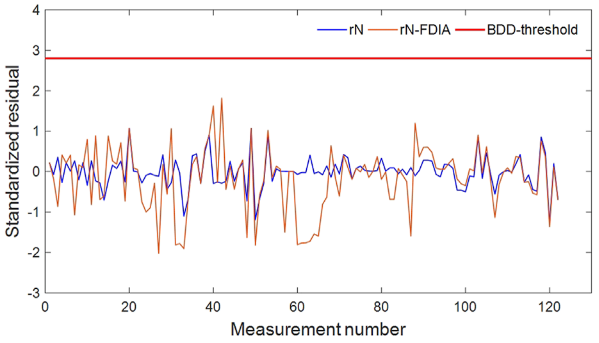

FIGURE 4
Simulation of bad data detection for FDIA.

In this section, the performance of the cosine similarity ratio detection algorithm based on measurement values is validated. As mentioned earlier, the detection indicator is the ratio of cosine similarity between the actual and estimated measurement values at two time instances. This experiment takes state estimation, bad data detection, and FDIA detection as a time cycle (approximately 2–3 s), eliminating subsequent steps such as local protection and recovery control, focusing only on the FDIA detection process and simplifying the overall process. Therefore, a 50 min simulation experiment is conducted on the system, and three different types of FDIA are launched in the 5th, 10th, and 15th minutes. Actual power accident events are added as references in the 35th, 40th, and 45th minutes. The specific event types are shown in Table 2 and Figure 5 shows the simulation results of cosine similarity ratio detection.

In Figure 5, in the FDIA experiment launched within the first 30 min, it can be clearly seen that before the injection of false data, the value of μ_k tends to remain close to 1. Among them, $min\{\mu_1\cdots,\mu_k\cdots\} = 0.99$ and $max\{\mu_1\cdots,\mu_k\cdots\} = 1.01$. Therefore, it can be concluded that when the system is not under FDIA, $\mu_k \in [0.99, 1.01]$. After the attack occurs, the value of μ_k quickly deviates from the normal fluctuation range. After conducting multiple experiments on attacks on different nodes and based on a large amount of historical data, it was found that the detection effect was optimal when the detection threshold margin $\ell = 0.01$ was set. That is, the range of the threshold for determining the presence of an attack is $\mu_k \in [0.98, 1.02]$.

However, in the FDIA experiment of the last 20 min, the occurrence of the actual power accident will also lead to a significant decline in the similarity ratio of measurement, which is also one of the common problems of the detection algorithm based on the power grid data mentioned above.

In summary, after injecting FDIA, the similarity immediately significantly decreases and falls below the detection threshold, which can trigger an attack alert. However, data anomalies caused by actual power accidents can also lead to this. Therefore, the complete MERS algorithm will demonstrate in the next section how to distinguish FDIA from actual power accidents.

TABLE 2 List of events at different times in simulation.

Time (min)	Events description
5	FDIA for active power measurements
10	FDIA for generator power measurements
15	FDIA for frequency measurements
35	Disconnected of Branch between Node 5 and Node 6
40	Disconnected of Generator connected on Node 34
45	Node 21 splitting

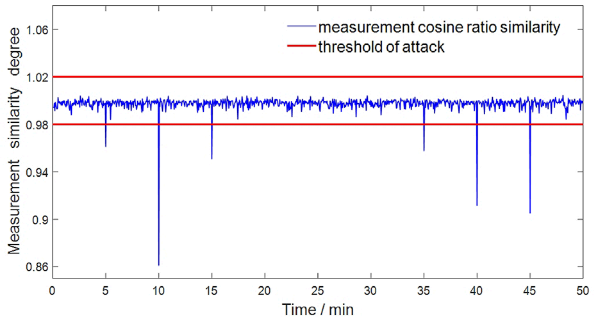

FIGURE 5
The proposed FDIA detection method effectiveness test.

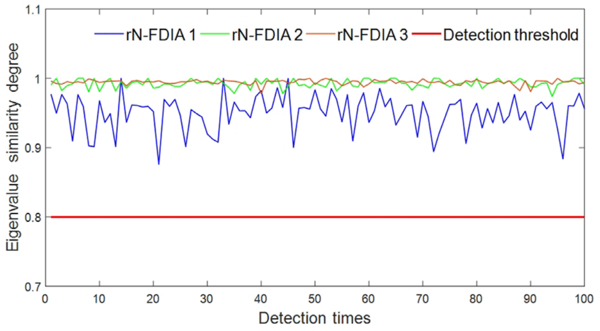

FIGURE 6
The proposed MERS test of FDIA.

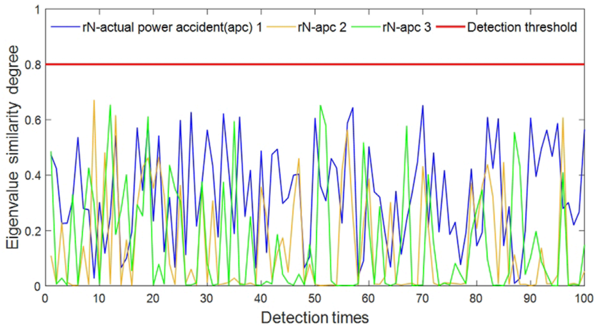

FIGURE 7
The proposed MERS test of actual power accident.

4.3 Distinguishing from actual power accident

The proposed method for distinguishing between FDIA and actual power accidents based on MERS test was validated for its effectiveness in this section.

The selected eigenvalue residuals were obtained based on the principle of constructing imperfect FDIA, and the attack vector of imperfect FDIA could be obtained using Eq. 7. From this principle, the unmodified measurement part could be deduced. In this experiment, three different types of FDIAs were designed, and the specific types are shown in Table 2. Taking the first type of FDIA, which injects false data of branch active power, as an example, Eq. 8 indicates that this attack mainly causes active power deviation by modifying the reactive power and voltage phase angle of the line, and the generator output power can be used for eigenvalue measurement. The complete detection system contains a complete program that includes eigenvalue checks for various types of FDIA.

Figures 6, 7 are the results of MERS tests for the selected eigenvalue. Three types of FDIA and three types of actual power accidents correspond to the events in Table 2. It is apparent that the attack warning caused by FDIA is much higher than that caused by actual power accidents in terms of eigenvalue residual similarity. From Figure 6, it can be seen that the similarity differences caused by FDIA launched for different measurement data are significantly different. The tampering attack against active measurement values needs to be modified to more necessary measurement values, resulting in a much lower similarity compared to the other two FDIAs.

According to the above construction principle of imperfect FDIA, in order to bypass the bad data detection, the attack vector will maintain the high consistency of most eigenvalues and only modify the necessary residual values required by the attack, which will maintain the similarity of residual eigenvalues before and after the attack at a high threshold. On the contrary, the occurrence of actual power grid accidents is random, resulting in low similarity in residual eigenvalues.

After conducting a large number of experiments, the threshold $\mu_{rN} \approx 0.8$ was determined, i.e., when the eigenvalue residual similarity is greater than or equal to 0.8, the system can be considered to have been invaded by FDIA, and when the eigenvalue residual similarity is less than 0.8, the system is considered to have experienced an actual power accident.

5 Conclusion

Compared with existing testing methods, the detection method proposed in this paper combines measurement similarity testing and eigenvalue residual similarity testing to achieve high-precision detection of FDIA. Specifically, under normal operating conditions or physical power accidents, it does not generate a high false alarm rate, while in the presence of FDIA, it achieves a high detection rate.

The proposed similarity testing method for residual measurement of eigenvalues only requires system state measurements from measurement equipment and measurement residuals from state estimation as inputs, making it very suitable for modern energy management systems. The localization and elimination of false data will be the focus of future research directions.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

YZ: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing–original draft. RL: Conceptualization, Formal Analysis, Methodology, Software, Writing–original draft. DC: Validation, Writing–review and editing. HG: Validation, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project is Supported by SEPRI High Potential Program (SEPRI-K213014).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abur, A., and Exposito, A. G. (2004). *Power system state estimation: theory and implementation*. CRC Press.

Chaojun, G., Jirutitijaroen, P., and Motani, M. (2015). Detecting false data injection attacks in ac state estimation. *IEEE Trans. Smart Grid* 6, 2476–2483. doi:10.1109/tsg.2015.2388545

Chen, B., Wu, Q., Li, M., and Xiahou, K. (2023). Detection of false data injection attacks on power systems using graph edge-conditioned convolutional networks. *Prot. Control Mod. Power Syst.* 8, 16–12. doi:10.1186/s41601-023-00287-w

Chen, L., Xu, A., Jiang, Y., Yang, H., Lu, H., Kuang, X., et al. (2020). Attack pattern recognition algorithm of power information network based on dynamic incremental cluster analysis. *South. Power Syst. Technol.* 14, 25–32. doi:10.13648/j.cnki.issn1674-0629.2020.08.004

Coluccia, A., D'Alconzo, A., and Ricciato, F. (2013). Distribution-based anomaly detection via generalized likelihood ratio test: a general maximum entropy approach. *Comput. Netw.* 57, 3446–3462. doi:10.1016/j.comnet.2013.07.028

Ericsson, G. N. (2010). Cyber security and power system communication—essential parts of a smart grid infrastructure. *IEEE Trans. Power Deliv.* 25, 1501–1507. doi:10.1109/tpwrd.2010.2046654

Feng, L., Yong, M., and Yihua, Z. (2023). Overview of the overall influence of cyber attack on typical scenarios of power systems. *South. Power Syst. Technol.*, 1–15.

He, Y., Mendis, G. J., and Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism. *IEEE Trans. Smart Grid* 8, 2505–2516. doi:10.1109/tsg.2017.2703842

Jena, P. K., Ghosh, S., Koley, E., and Manohar, M. (2021). An ensemble classifier based scheme for detection of false data attacks aiming at disruption of electricity market operation. *J. Netw. Syst. Manag.* 29, 43. doi:10.1007/s10922-021-09610-y

Jiang, Q., Chen, H., Xie, L., and Wang, K. (2020). Learning-based cooperative false data injection attack and its mitigation techniques in consensus-based distributed estimation. *IEEE Access* 8, 166852–166869. doi:10.1109/access.2020.3023117

Liu, R., Mustafa, H. M., Nie, Z., and Srivastava, A. K. (2022). Reachability-based false data injection attacks and defence mechanisms for cyberpower system. *Energies* 15, 1754. doi:10.3390/en15051754

Liu, Y., Ning, P., and Reiter, M. K. (2011). False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur. (TISSEC)* 14, 1–33. doi:10.1145/1952982.1952995

Tiande, G., Pingping, L., Jikeng, L., et al. (2023). Leverage measurement attack of power system considering unknown state variates. *South. Power Syst. Technol.* 17, 56–64.

Wu, Y., Wang, Q., Guo, N., Tian, Y., Li, F., and Su, X. (2023). Efficient multi-source self-attention data fusion for fdia detection in smart grid. *Symmetry* 15, 1019. doi:10.3390/sym15051019

Xiahou, K., Liu, Y., and Wu, Q. (2020). Robust load frequency control of power systems against random time-delay attacks. *IEEE Trans. Smart Grid* 12, 909–911. doi:10.1109/tsg.2020.3018635

Yang, H., Zhang, W., Shi, F., Xie, J., and Ju, W. (2019). Pmu-based model-free method for transient instability prediction and emergency generator-shedding control. *Int. J. Electr. Power & Energy Syst.* 105, 381–393. doi:10.1016/j.ijepes.2018.08.031

Zhang, Q., Li, F., Cui, H., Bo, R., and Ren, L. (2020a). Market-level defense against fdia and a new lmp-disguising attack strategy in real-time market operations. *IEEE Trans. Power Syst.* 36, 1419–1431. doi:10.1109/tpwrs.2020.3020870

Zhang, Y., Wang, J., and Chen, B. (2020b). Detecting false data injection attacks in smart grids: a semi-supervised deep learning approach. *IEEE Trans. Smart Grid* 12, 623–634. doi:10.1109/tsg.2020.3010510

Zhao, J., Mili, L., and Wang, M. (2018). A generalized false data injection attacks against power system nonlinear state estimator and countermeasures. *IEEE Trans. Power Syst.* 33, 4868–4877. doi:10.1109/tpwrs.2018.2794468