Computers and Electrical Engineering 119 (2024) 109572

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

FDIA localization and classification detection in smart grids using multi-modal data and deep learning technique

Jun Wang [a,*], Haoran Chen [a], Yifei Si [a], Yonghai Zhu [b], Tianci Zhu [a], Shanshan Yin [a], Bo Liu [a]

[a] School of Information Engineering, Henan University of Science and Technology, 263 Kaiyuan Avenue, Luolong District, Luoyang, 471000, China

[b] Guangdong Nanhai Power Design & Engineering Co., LTD, No. 33 Jihua East Road, Guicheng Street, Nanhai District, Foshan, 528000, China

ARTICLE INFO

Keywords:
Power system
FDIA
VGAE
TCN-GRU
Multimodal data

ABSTRACT

False data injection attacks (FDIA) exploit the vulnerabilities of bad data detection in energy management systems to maliciously tamper with state estimation results, seriously jeopardizing the safe and reliable operation of power systems. In order to promptly detect FDIA, recent studies have used machine learning techniques to extract attack characteristics and detect FDIA based on changes in these characteristics. Current research predominantly focuses on detecting a single type of FDIA attack model and topology. However, the increasing diversification of FDIA construction methods and the variability of topological structures in power systems can reduce or even invalidate detection effectiveness. To cope with the abovementioned difficulties, this study introduces a multimodal deep learning detection model based on variational graph auto-encoders (VGAE), temporal convolutional networks (TCN), and gated recurrent units (GRU). The topological features of power systems are obtained through VGAE to adapt to the detection needs of various topological structures and performance enhancement. These features are then integrated with preprocessed measurement data via BDD to form multimodal data. Furthermore, this multimodal data is used to locate and classify FDIA with complete information, FDIA with incomplete information, and topology attack after applying a multi-label classification algorithm based on TCN-GRU for temporal feature extraction. Experiments carried out on the IEEE 14 and IEEE 118 bus systems show that the proposed method is robust and has high detection performance. The results indicate that the proposed detection method achieves AUC values that are, on average, 0.085, 0.145, and 0.04 higher than those of CNN, LSTM, and CNN-LSTM, respectively. Moreover, under various noise and attack intensities, recall, precision, F1 score, and RACC remain above 0.859, 0.877, 0.867, and 0.818, respectively, with a classification accuracy greater than 0.912. This study provides a unique perspective on detecting FDIA across various attack models and power system topologies.

1. Introduction

The electric power system is undoubtedly a critical infrastructure ensuring the safety of the national economy and is an essential basis for the development of modern society [1]. The emergence of intelligent devices such as phase measurement units (PMU), remote terminal units (RTU), and smart meters for smart grids provides a new method of collecting, transmitting, and processing

* Corresponding author.
E-mail address: wj@haust.edu.cn (J. Wang).

https://doi.org/10.1016/j.compeleceng.2024.109572
Received 15 January 2024; Received in revised form 14 July 2024; Accepted 15 August 2024
Available online 23 August 2024

Abbreviations	
FDIA	false data injection attacks
VGAE	variational graph auto-encoders
TCN	temporal convolutional networks
GRU	gated recurrent units
PMU	phase measurement units
RTU	remote terminal units
DOS	denial of service
BDD	bad data detection
ICA	independent component analysis
PCA	principal component analysis
ML	machine learning
CVF	collaborative vulnerability factor
CDBN	conditional deep belief network
SVM	support vector machine
GAB	gentle Adaboost
GAN	generative adversarial network
DNNs	deep neural networks
EE-MPQ	early exit mechanism and mixed-precision quantization
CNN	convolutional neural network
LSTM	long and short memory
KNN	k-nearest neighbors
VAE	variational auto-encoder
RNN	recurrent neural network
IoT	internet of things
SCADA	supervisory control and data acquisition
DC	direct current
AC	alternating current
WLS	weighted least-squares
KL	Kullback–Leibler
RACC	row accuracy
AUC	area under the curve
ROC	receiver operating characteristic curve
FPR	false positive rate
TPR	true positive rate

high-precision real-time data, making it more convenient for the control center to grasp the global information for assuring the stability and efficiency of the power system. However, as the smart grid relies heavily on information techniques, the control center of the smart grid is dramatically vulnerable to various cyber-attacks (Denial of service (DOS) attacks [2], data eavesdropping [3], false data injection attacks (FDIA), etc.). Highly stealthy FDIA can bypass inferior detection mechanisms and may lead to power system failures, power outages, system instability, and user data and privacy damage. Compared to other cyber-attacks, FDIA has more of a vicious threat to the security of power systems. Furthermore, due to the advantages of the DC grid, such as high efficiency, stability, reliability, and sustainability, it is widely used in industrial and civil fields [4]. Hence, FDIA detection for the DC grid deserves unique attention.

FDIA, first named by Liu, means that an FDI attacker could systematically generate erroneous measurements after acquiring the power system configuration to avoid the residual-based bad system detection mechanism [5]. However, to launch such an attack, the attacker needs to understand the highly confidential global topology information of the target power system in advance. Therefore, the FDIA method using incomplete information is presented. Namely, an attacker with limited grid configuration details can perform a stealthy FDIA attack that bad data detection (BDD) cannot discover. For instance, Esmalifalak et al. [6] used independent component analysis (ICA) to estimate Jacobi matrices to further generate attack vectors. Yu and Chin [7] applied the principal component analysis (PCA) approximation method to create attack vectors under the condition of incomplete information. Bi and Zhang [8] utilized the min-cut FDIA model to form attack vectors. Nevertheless, these FDIA attacks only falsify the measurement information and do not modify the topology information of the control center. In order to forge topology information, Kim et al. [9] proposed a topology attack combined with the traditional FDIA attack by modifying topology parameters and infiltrating specific measurements, resulting in the incorrect topology information transmitted to the control center not being detected.

Currently, approaches to solving the FDIA threat can be divided into model-based detection algorithms and data-driven detection algorithms [10]. The model-driven methods mainly employ the dynamic data of real-time observation and the static data (physical

characteristics, topology structure, and power flow, etc.) to establish the system response model, and realize the FDIA detection by comparing the dissimilarity between the actual monitored data and the model-predicted data. Moreover, data-driven methods rely on machine-learning (ML)-based means that learn complex features and representations from the vast amounts of collected data and are trained to distinguish between normal operations and potential attacks. Unlike model-driven algorithms, data-driven algorithms do not require statistical assumptions about the system model or predefined attacks and can be more flexible to adapt to various attacks.

In this study, we propose a detection method based on variational graph auto-encoders (VGAE), temporal convolutional networks (TCN), and gated recurrent units (GRU) for detection of three types of FDIAs including FDIA with complete information, FDIA with incomplete information, and topology attack. We consider the FDIA detection as a multi-label classification issue. Specifically, the BDD detector is used to remove low-quality data after being attacked by the three FDIAs. Moreover, the topological features are extracted from the power system under attack using the VGAE, and the metrological data and topological features are integrated into a multimodal dataset. Furthermore, the hyperparameters (Learning rate and batch size) of the TCN-GRU are optimized using the grid search method to locate and classify the FDIA accurately. The main contributions and innovations of this study are as follows:

- Three FDIAs (FDIA with complete information, FDIA with incomplete information, and topology attack) commonly performed in power systems are assessed to develop the composite detection model to enhance the generalization of the FDIA detection method.
- Considering the changeable topology structure, using VGAE to extract topological features of the power grid combined with the measurement data effectively improves the detection capability and can be adjusted to diverse topologies.
- The learning rate and batch size of VGAE-TCN-GRU are optimized by comparing different number of hidden layers and using the grid search method to promote the accuracy of the FDIA localization and classification detection algorithm.

2. Related work

FDIA obliterates the normal operation of the power system by manipulating or insinuating false data. Recently, FDIA detection has become an essential mission in protecting the security of the power system. Appropriate countermeasures can be promptly conducted by accurately uncovering the attack source and type to prevent the attack from spreading further.

The model-driven-based detection method is one of the methods usually used to detect FDIA. The model-driven-based detection method can be specifically categorized into state-estimation-based detection and direct computational methods. State-estimation-based detection method detects possible FDIA by estimating the power system state and comparing the discrepancies between the measured data and the predicted values. Manandhar et al. [11] proposed a model-driven estimation method using Kalman filter to discover FDIA in the power grid. Gu et al. [12] addressed a detection method based on distributed state estimation to locate the corresponding subsystem of bad data. In addition, the direct computational approach depends on the system measurements and parameters to discover the FDIA and does not require any estimation process. Sahoo et al. [13] introduced a collaborative vulnerability factor (CVF) framework to determine whether a micro-grid control system suffers from FDIA by probing the secondary output convergence of voltage controllers. Li et al. [14] presented an FDIA detection method based on the low-rank property of the measurement matrix and the sparsity of the attack matrix to separate the measurement data from the attack data.

Model-driven approaches have several advantages, such as interpretable parameters, reduced time costs in model training, and lowered storage requirements. However, it is essential to update or adjust the parameters of the detection algorithms to deal with emerging, unexplored attack patterns when attackers change their strategies or use unknown tactics. Furthermore, the detection rate of model-driven detection algorithms depends on diverse factors such as the attack, the proportion of measurements affected, the number of attack samples, the state estimation detection threshold, and various model parameters. For instance, decreasing the state estimation detection threshold can increase the detection rate. Nevertheless, it can also raise the false alarm rate, heightening the risk of erroneous actions in power system dispatching.

With the rapid development of machine learning and deep learning technologies, many studies have attempted to apply these methods to FDIA detection. These methods model and train historical data from power systems to identify anomalies or false data by use of ML algorithms. For example, He et al. [15] integrated a standard deep belief network and a conditional Gauss-Bernoulli constrained Boltzmann machine to obtain a conditional deep belief network (CDBN) for distinguishing covert FDIA, with a more than 93% accuracy. Xiong et al. [16] combined the support vector machine (SVM) algorithm into the gentle Adaboost (GAB) framework to recognize clandestine FDIA, achieving more than 88% accuracy under various attack intensities. Zhang et al. [17] incorporated an autoencoder into a generative adversarial network (GAN) framework to discriminate FDIA with an accuracy of over 96%. James et al. [18] used discrete wavelet transform to reveal spatial data features and grabbed temporal correlation using deep neural networks (DNNs) to detect FDIA with over 90% accuracy.

Machine learning detection methods require manual feature extraction, while deep learning detection methods can automatically learn these features without additional human intervention. Meanwhile, when extracting attack characteristics from voltage or current signals in power systems, the significant computational complexity can substantially degrade the accuracy and efficiency of attack detection by machine learning methods. In addition, although deep learning methods lessen computational complexity, they are primarily used to identify the existence of FDIA and are still insufficient for locating the specific attack.

To solve these predicaments, researchers have applied ML models as multi-label classifiers to obtain the location of FDIA. For instance, Zhu et al. [19] proposed a fast and lightweight FDIA localization detection framework combined with an early exit mechanism and mixed-precision quantization (EE-MPQ), resulting in a detection precision, recall, F1 score, and RACC higher than

99.39%, 98.1%, 98.99%, and 91.25%, respectively. Zia et al. [20] used binary relevance and classifier chaining methods for position detection of FDIA in smart grids. Wang et al. [21] suggested a multi-label classification method based on convolutional neural network (CNN) architecture for localization detection in FDIA, with more than 99% in terms of precision, recall, F1 score, and RACC. Mukherjee et al. [22] applied CNN framework, convolutional neural networks with long and short memory (CNN-LSTM), convolutional neural networks with gated recurrent units (CNN-GRU), and k-nearest neighbors (KNN) to gain the location of FDIA, respectively.

Although previous literature has employed multi-label classification models for the localization detection of FDIA, it often overlooks the potential impacts of variations in power network topology and diversity in attack patterns on detection performance. Changes in the power network topology, such as maintenance, upgrades, or accidental events that cause disconnections or reconnections, can significantly affect system behavior. If the detection models cannot adapt to these topological changes, normal data may be mistakenly identified as anomalous due to modifications in the topology, leading to an increase in false positive rates. Additionally, the detection system is designed for a single FDIA attack pattern. In this scenario, the existing detection model may not be able to identify and respond to new types of attacks quickly, resulting in underreporting.

VGAE is an unsupervised learning framework for graph-structured data based on the variational auto-encoder (VAE) and has been widely used in graph representation learning. Initially, Kipf and Welling [23] offered the VGAE method to learn interpretable latent representations of undirected graphs. By introducing encoder and decoder structures, VGAE is able to perceive low-dimensional representations from graph data and capture information about the relationship and structure between nodes in the graph. Xie et al. [24] used VGAE to combine topic learning and graph embedding for text categorization. Meanwhile, Zhang et al. [25] used the VGAE method to extract deep features of coupled networks for anomaly identification in nuclear power plant systems. Therefore, in this study, the topological features of the power system are extracted based on VGAE to indicate the connection relationship between nodes and the topological structure in the power system.

Long Short-Term Memory (LSTM) is a particular type of recurrent neural network (RNN) capable of understanding long-period dependence information, and can solve the gradient disappearance or gradient explosion problems in RNN. GRU is an advancement of LSTM that simplifies the structure, reduces the number of training parameters, and decreases the time complexity while maintaining the effectiveness of LSTM [26]. In addition, GRU has demonstrated significant outcomes in recognizing the cyber-attacks in internet of things (IoT) and smart grid [27,28]. Specially, TCN is a unique architecture combining CNN and RNN, widely used in pattern recognition and anomaly detection [29,30]. Using structures such as dilated causal convolution and residual connection, TCN can exceed the traditional CNN model and process multiple time series in parallel while consuming a miniature amount of running memory [31]. In our study, TCN and GRU are merged to take advantage of TCN in processing time-series data, and the gating mechanism of GRU is used to enhance the learning ability of long-term dependencies to improve the performance of FDIA localization and classification detection in power systems.

3. Preliminaries

3.1. Power system state estimation

Power system state estimation is the basis for energy management systems and wide area measurement systems to execute analysis and control functions such as optimal tide calculation, load forecasting, transient stability, etc. Its main functions include enhancing the accuracy of measurement data, deriving precise electrical parameters of the power system, and promoting the reliability of supervisory control and data acquisition (SCADA) systems [32]. In this study, the state estimation of direct current (DC) grids is investigated due to their advantages over alternating current (AC) grids, such as higher reliability, simpler control, and more efficient interfaces with renewable energy sources and energy storage devices. In the DC model, the relationship between measurements and state variables can be expressed as

$$z = Hx + e \tag{1}$$

where $z \in \mathbb{R}^m$ is the measurement vector, including the bus's injected power measurement and the transmission line's power flow measurement, $x \in \mathbb{R}^n$ is the system state vector, representing the voltage phase angle. In addition, m is the number of meter measurements, and n is the number of system state variables. e denotes the measurement error, $H \in \mathbb{R}^{m\times n}$ implies the Jacobian matrix.

Weighted Least-Squares (WLS) is used to obtain the estimates of the state variables as follows [33]:

$$\hat{x} = \left(H^T W H\right)^{-1} H^T W z \tag{2}$$

where W is a diagonal matrix composed of the variances of the Gaussian distribution of each meter error.

3.2. BDD and FDIA

In order to eliminate the influence of bad data on state estimation results, the traditional BDD method is to compare the ℓ_2-norm of the residuals with a threshold τ. If the ℓ_2-norm of the residual exceeds τ, the relevant measurement will be considered bad data [34]. The threshold range of τ is set by the 3σ rule. The BDD detector recognizes the presence of attack:

$$R = z - Hx_2^2 \geq \tau \tag{3}$$

FDIA explores the flaws of this detection method by injecting the attack vector $a = [a_1, a_2, \ldots, a_m]^T$ into the measured value, which is tampered with to be $z_{bad} = z + a$. At this point, the estimated value of the state variable is $x_{bad} = \hat{x} + c$, and $c = [c_1, c_2, \ldots, c_n]^T$ expresses the error variable introduced by the attacker in the state variables. The residual expression at this time can be indicated as follows:

$$r = z_{bad} - Hx_{bad} = z + a - H(\hat{x} + c) = z - H\hat{x} + a - Hc \tag{4}$$

When $a = Hc$, the following equation holds

$$r = z_{bad} - Hx_{bad} = z - H\hat{x} \tag{5}$$

At this time, the BDD detector cannot recognize the existence of the attack, and the attacker can arbitrarily interpolate the measurement values and state variables, yielding a threat to the safe operation of the power system.

The first attack involved in this study is the FDIA with complete information that constructs the attack vector once the attacker has already acquired the topological data of the power grid and electrical parameters.

Nevertheless, this information is not easily accessed. Hence, the second attack considered is the FDIA with incomplete information that forms the attack vector using the PCA approximation method. The construction of the attack vector depends only on the observed measurement information matrix.

The third type of attack is a topology attack that causes a modification in the topology information observed by the power system. To change the topology information, the attacker intercepts the data (s, z) from RTU by implementing a man-in-the-middle attack, $s \in \{0, 1\}^d$ is the power network topology data representing the state of the circuit breaker (0 for on, 1 for off). Moreover, the attacker modifies the intercepted data as $(\bar{s}, \bar{z})$ and sends it to the control center.

$$\bar{s} = s + b\,(mod 2) \tag{6}$$

$$\bar{z} = z + a_t \tag{7}$$

where a_t is the measurement attack vector, and b is the line state attack vector.

The topological attack can be produced by modifying the measurement values without the change of current state variables, and the corresponding measurement attack vector can be described as follows:

$$a_t = \bar{H}x - Hx \tag{8}$$

where H is the Jacobi matrix after the modification of topological information.

To construct the above attack vector, the attacker merely requires the local topology information of the specific attack line and the related measurements. Moreover, an attacker can alter the status of a line observed by the control center to disconnected by adjusting the relevant measurements of the line. Before the attack, it is assumed that z_{ij} is the branch flow from node i to node j, and the node injection powers at both ends of the line are z_i and z_j, respectively. For revising the line state to disconnected under the attack, the attacker needs to set the measurements z_{ij} and z_{ji} to 0. However, if these two measurements are directly switched to 0, they may be detected by the BDD detector. Therefore, the attacker reconditely modifies the measured values on both sides of the bus to avoid this situation. For line (i, j), the injected power measurement z_i at bus i is subtracted by z_{ij}, and the injected power measurement z_j at bus j is deducted by z_{ji}.

4. Detection model

This study proposes an attack detection model based on multimodal deep learning consisting of VGAE and TCN-GRU. The overall process architecture is shown in Fig. 1. The detection model includes modules for data collection and preprocessing, multimodal data fusion, and FDIA attack detection. The data acquisition and preprocessing module transmits various data types from the power grid through RTUs, PMUs, and other intelligent electronic devices to the SCADA system. The collected measurement data is then preliminarily identified for anomalies by a BDD detector. In the multimodal data fusion module, the VGAE algorithm represents the topological structure of the power system as a graph, utilizing graph neural networks to process node and edge information for gaining topological features. The measurement data preprocessed by the BDD is combined with these topological features side by side to form multimodal data. The FDIA attack detection module utilizes the TCN-GRU model to analyze multimodal data and pre-set labels. The TCN component of the model is designed to handle local patterns in long sequences, while the GRU component is built to retain long-term dependency information. The model can effectively extract temporal features from the multimodal data by combining these capabilities. The extracted features are then processed through a sigmoid layer to achieve attack localization and classification. Notably, this study refers to FDIA with complete information and FDIA with incomplete information as conventional FDIA since both cases involve falsification of measurement data. This study labels the power system's measurement data and topological features to signify whether they are compromised and whether the samples are subjected to conventional or topology attacks. Each nodal and branch measuring instrument can obtain a measurement corresponding to a location on the instrument. Thus, each measurement is associated with a location label. There are 34 location labels for the IEEE 14 bus system and 304 for the IEEE 118 bus system. A location label of 1 implies that the meter associated with that label has been attacked, while a label of 0 indicates no attack. Topological features are represented by a one-dimensional topological label for both systems. A topological label of 1 infers that the sample has been subjected to a topological attack, while a label of 0 denotes a conventional FDIA.

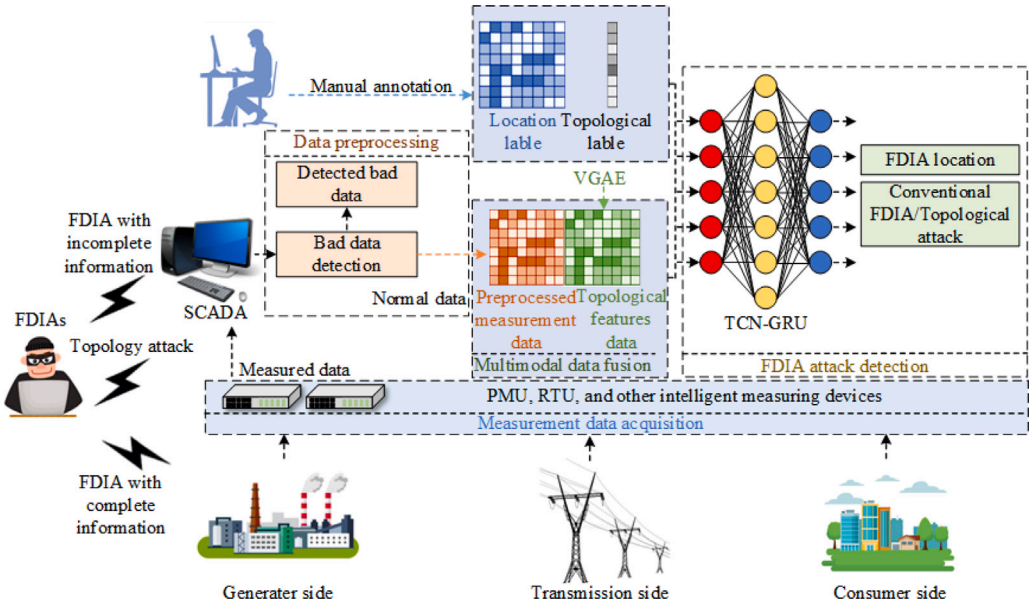


Fig. 1. Attack detection model based on multimodal deep learning.

4.1. Multimodal feature fusion

Because general autoencoders cannot handle the topological data in non-Euclidean space, VGAE is used to extract the topological features of the power system and map the topological features to 1-dimensional vector space for parallel combination with the measurements after BDD preprocessing (Fig. 2). The power system network topology can be regarded as an undirected graph, and the buses and branches of the power grid can be illustrated by the nodes and edges of the undirected graph. Assuming that there are a total of n buses and b branches in the power system, the power system can be represented as $\mathcal{G} = (A, Y)$, where $A \in \mathbb{R}^{n \times n}$ denotes the adjacency matrix. If bus i and bus j are directly connected by a line, then A_{ij}=1, otherwise A_{ij}=0. $Y = \left[y_1, y_2, \ldots, y_n\right]$ denotes the node characteristics, including bus labeling, bus type, bus injected power, and bus phase angle. In this study, a two-layer GCN is firstly applied as the encoder to learn the node representation distribution of the undirected graph, and the mean and variance of the node representation distribution can be calculated by

$$GCN(X, Y) = \bar{A} ReLU\left(\bar{A} Y W_0\right) W_1 \tag{9}$$

$$\bar{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \tag{10}$$

$$\mu = GCN_{\mu}(Y, A) \tag{11}$$

$$log\sigma = GCN_{\sigma}(Y, A) \tag{12}$$

where W_0, W_1 are the weight parameters to be learned, $\bar{A}$ is the symmetric normalized adjacency matrix, and D is the degree matrix of A.

$$q\left(z_i|Y, A\right) = N\left(z_i|\mu_i, diag\left(\sigma^2\right)\right) \tag{13}$$

$$q(Z|Y, A) = \sum_{i=1}^{N} q\left(z_i|Y, A\right) \tag{14}$$

where v_i is the element of feature embedding V, and V is the extracted topological feature.

Subsequently, the decoder reconstructs the graph by computing the probability of the existence of a branch between two nodes, Sigmoid is chosen as the activation function, and the process can be expressed as follows:

$$P\left(A_{ij} = 1|v_i, v_j\right) = sigmoid\left(v_i^T v_j\right) \tag{15}$$

$$P(A|V) = \sum_{i=1}^{N} \sum_{i=1}^{N} P\left(A_{ij}|v_i, v_j\right) \tag{16}$$

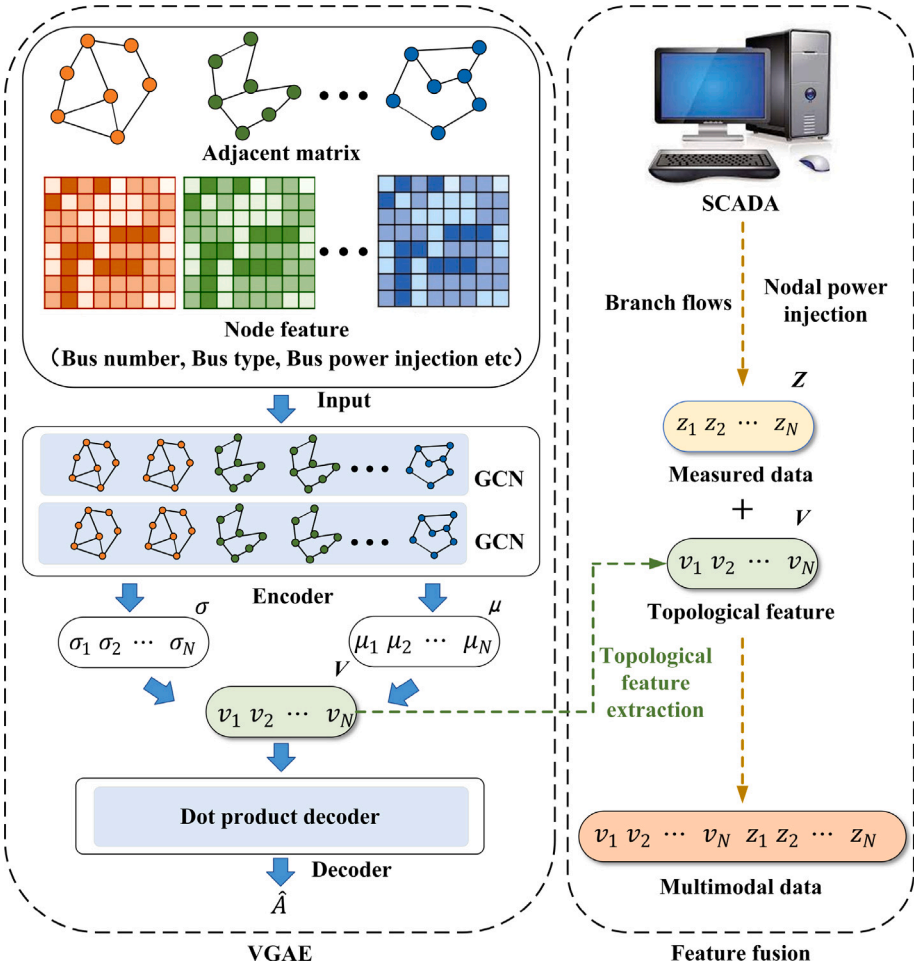


Fig. 2. Multimodal feature fusion process.

Finally, to ensure that the feature embedding V can effectively retain critical information about power system operation, VGAE is trained using cross entropy and Kullback–Leibler (KL) divergence as the loss function $\mathcal{L}$. Cross entropy assesses the similarity between the topology information of the original power system and the decoded information. KL divergence is employed to evaluate the similarity between the two distribution values.

$$L = E_{q(V|Y,A)}\left[log P\left(A|V\right)\right] - KL[q(V|Y,A)||P\left(V\right)] \tag{17}$$

The multimodal data is assembled by the measurement data Z acquired after BDD preprocessing and feature embedding V obtained after training via Eq. (17).

4.2. Attack detection model

The TCN-GRU-based multi-label classifier is used for attack detection and localization. The TCN-GRU model structure is consisted of input layer, temporal convolutional network layer, threshold recurrent unit layer, and an output layer (Fig. 3).

(1) Input layer: The input data is the multimodal data after combining the meter measurement data (node injected power and branch active power) and the topological features collected by VGAE. The input labels are the localization labels and topology labels.

(2) Temporal convolutional network layer: 4 layers of residual units are set up. Each layer of the residual unit contains 2 convolutional units and 1 nonlinear mapping. The ReLU function is used as an activation function to normalize the weights of the convolution kernel in convolutional units. The convolutional kernel size is 3, and the dropout coefficient is designated to 0.05 to prevent training overfitting and accelerating the speed of convergence. Moreover, the expansion factor is assigned to [1,2,4,8], and the number of filters is 3.

(3) Threshold recurrent unit layer: Two GRU layers are selected. The first layer contains 128 GRU units, and the second layer includes 64 GRU units. The output of the temporal convolutional network is used as the input of the threshold recurrent unit.

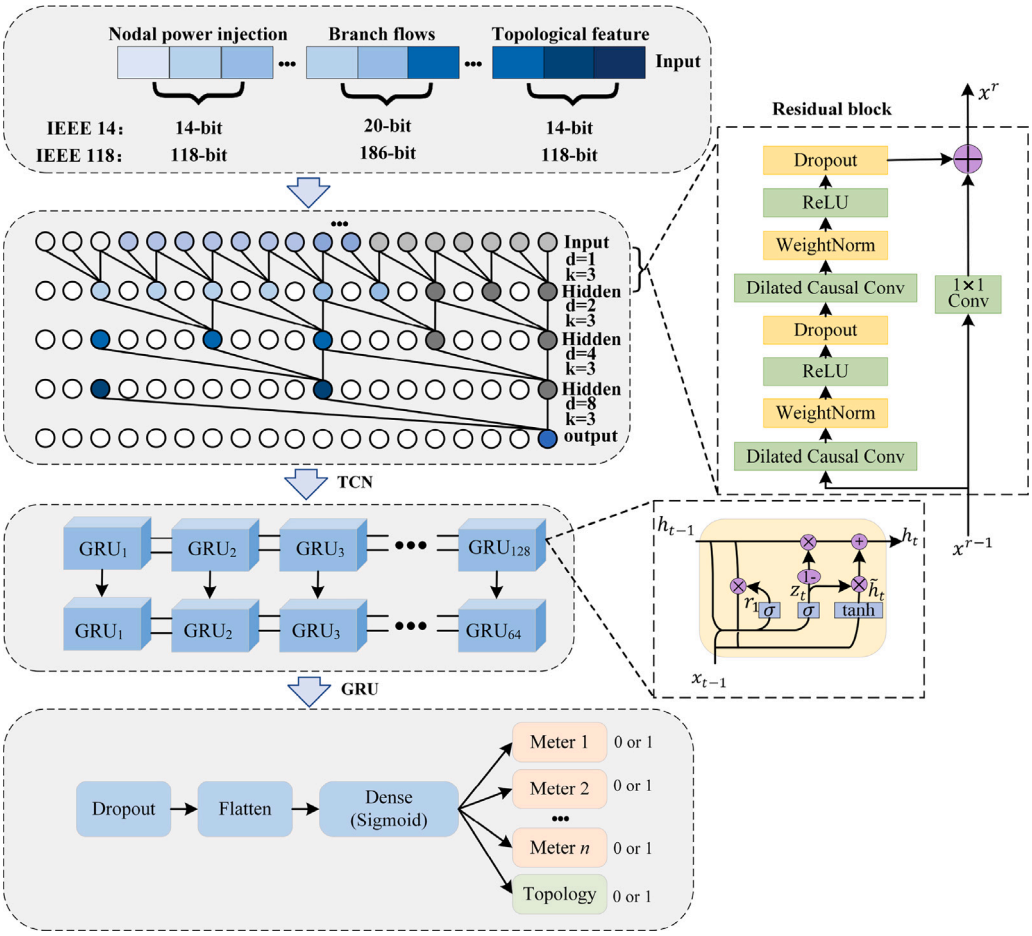


Fig. 3. Attack detection model.

(4) Output layer: This network layer uses the Sigmoid as the activation function, and the outputs are the predicted values of the localization and topological labels.

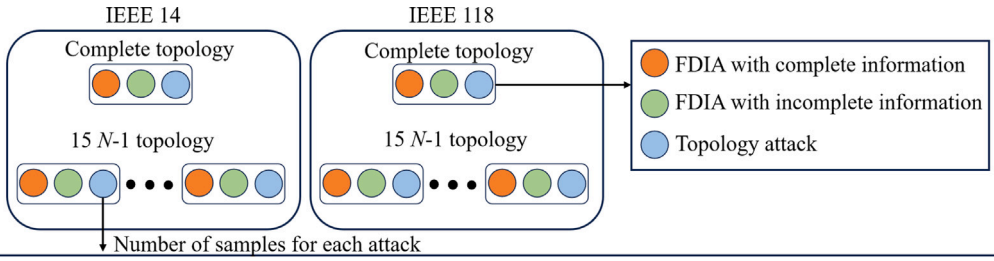
To improve the performance of the TCN-GRU algorithm, this study systematically optimized its hyperparameters. Specifically, the critical parameters include the number of hidden layers [35], learning rate [36], and batch size [37], which have a decisive impact on model performance. This study adjusts these model parameters by applying a grid search method. During optimization, 70% of the samples in each batch are used as the training set, and the remaining 30% are used as the validation set. Additionally, to tackle the problem of imbalanced labels in the dataset, this study utilizes a balanced cross-entropy loss function. This loss function incorporates weighting coefficients to the standard cross-entropy method, which assigns weights to attack data with uneven category distribution. Attack samples that occur less frequently are given higher weights, while those that occur more frequently are allocated lower weights. The technique helps mitigate the imbalances between attack samples at different positions, improving the model's detection capabilities. The optimization steps for the number of hidden layers, learning rate, and batch size are as follows:

(1) The batch size is set at 64, and the learning rate is fixed at 0.003. Different numbers of hidden layers (2, 3, 4, 5, 6) are tested by comparing recall, precision, F1 score, RACC, accuracy of classification detection, and total number of parameters for determining the optimal number of hidden layers.

(2) With the number of hidden layers designated at the optimal level, different combinations of learning rates (0.1, 0.03, 0.01, 0.003, 0.001) and batch sizes (32, 64, 128, 256, 512) are experimented with, comprising 25 variations. The optimal parameter combination is identified by comparing different combinations' F1 scores and RACC.

5. Results and analysis

Numerical simulations are performed on a computer with 32 G of RAM, an Intel Core i7 11800H CPU, and an NVIDIA GeForce RTX3080Ti GPU. The dataset is generated by the simulation tool PYPOWER, and the detection method is implemented using PyTorch. In the meantime, the proposed model is tested on IEEE 14 and IEEE 118 bus systems.



Attack intensity	Noise intensity	Sample size	Attack intensity	Noise intensity	Sample size
1	0.2	3000	2	0.1	3000
2	0.2	3000	2	0.2	3000
3	0.2	3000	2	0.3	3000
4	0.2	3000	2	0.4	3000
5	0.2	3000	2	0.5	3000

Fig. 4. Dataset setting diagram.

5.1. Data generation

(1) Normal data generation: The FDIA detection mechanism is evaluated in IEEE14 and 118 bus systems. Normal data is generated by using the original load distribution provided by the cases in the PYPOWER toolkit and setting the new load distribution to a random value between 80% and 120% of the initial value of the original load distribution, resulting in a total of 960,000 sets of normal data.

(2) Attack data generation in different topology structures: The FDIA samples with complete information, the FDIA samples with incomplete information, and the topology attack samples are yielded under complete topology structure, respectively. Three types of FDIA samples are generated for 15 randomly selected N-1 topologies in the IEEE 14 and IEEE 118 bus systems to verify the performance of detection models under different topologies. Each topology structure generates 30,000 sets of attack data.

(3) Attack data generation for different attack intensities: To assess the robustness of detection algorithms under different attack strengths, the ℓ_2-norm of the attack vectors is used as a metric for evaluating the attack strength when implementing FDIA with complete information and FDIA with incomplete information. The number of target state variables follows a discrete uniform distribution (2, 5) in the 14-bus system and (2, 10) in the 118-bus system. Meanwhile, the ℓ_2-norms of the attack vectors are set to be 1, 2, 3, 4, and 5, respectively. To carry out the topology attack, 1, 2, 3, 4, and 5 lines are randomly deleted from the 14-bus system, and 10, 20, 30, 40, and 50 lines are randomly removed from the 118-bus system. Every attack intensity yields 3000 sets of attack data.

(4) Attack data generation for various noise intensities: To estimate the robustness of detection methods under different noise intensities in the data acquisition environment, the standard deviation of Gaussian noise is used as an index to evaluate the noise intensity, and the standard deviation is set to 0.1, 0.2, 0.3, 0.4, and 0.5, respectively. Accordingly, each noise intensity generates 3000 sets of attack data.

The specific dataset configuration is shown in Fig. 4.

5.2. Evaluation index

For FDIA localization detection, recall is used to estimate the proportion of all actual attacked meter locations that are correctly located. The precision is applied to measure the ratio of meter locations that are actually under attack out of the sites detected by the model as attacked. The F1 score is the geometric mean of precision and recall. Moreover, row accuracy (RACC) is another crucial evaluation metric in localization detection, indicating the proportion of all correctly located samples to the number of all samples. Concretely, recall, precision, and F1 score are defined as follows:

$$recall = \frac{TP}{TP + FN} \tag{18}$$

$$precision = \frac{TP}{TP + FP} \tag{19}$$

$$F1\ value = 2 \times \frac{precision \times recall}{precision + recall} \tag{20}$$

Table 1
Detection comparisons between TCN-GRU and conventional algorithms.

Dataset	Detection algorithms	Location detection indicators							
		IEEE 14 bus system				IEEE 118 bus system			
		Recall	Precision	F1	RACC	Recall	Precision	F1	RACC
Measurement datasets	TCN-GRU	0.874	0.911	0.892	0.819	0.835	0.852	0.843	0.782
	LSTM	0.783	0.818	0.800	0.729	0.721	0.757	0.738	0.692
	CNN	0.837	0.877	0.856	0.751	0.786	0.826	0.805	0.739
	CNN-LSTM	0.833	0.890	0.861	0.776	0.801	0.832	0.816	0.745
Multi-modal datasets	VGAE-TCN-GRU	0.901	0.943	0.922	0.876	0.892	0.931	0.911	0.854
	VGAE-LSTM	0.814	0.840	0.827	0.808	0.807	0.825	0.815	0.784
	VGAE-CNN	0.857	0.893	0.875	0.823	0.839	0.886	0.861	0.808
	VGAE-CNN-LSTM	0.865	0.904	0.884	0.837	0.857	0.896	0.876	0.821

where *TP*, *FN*, and *FP* are the number of meter locations correctly detected as being under attack, the number of meter locations accurately assessed as non-attacked, and the number of non-attacked meter locations classified as attacked meter locations, respectively. RACC is calculated by

$$RACC = \frac{N_{true}}{N_{sum}} \tag{21}$$

where N_{true} denotes the number of samples in which all the attacked meters are properly located for a sample, and N_{sum} represents the total number of samples.

Accuracy is used as the evaluation metric of attack classification and is described as follows:

$$accuracy = \frac{TP_c + TN_c}{TP_c + TN_c + FP_c + FN_c} \tag{22}$$

where TP_c implies the number of samples in the topological label that are correctly detected as being under topology attack; TN_c expresses the number of samples suffering from a topology attack detected as not subject to a topology attack; FP_c denotes the number of samples not subject to a topology attack detected as being under a topology attack; FN_c means the number of samples rightly detected as not being subjected to a topology attack.

5.3. Validity of topological features

To validate the effectiveness of the combination of measurement data and topology features extracted by VGAE for detecting three types of FDIA, the model is configured with three hidden layers, a batch size of 64, and a learning rate of 0.03. The dataset sets the attack intensity at 2 and the noise intensity at 0.2, establishing both a measurement dataset and a multimodal dataset for simulation experiments. The measurement dataset consisting solely of measurement data (only bus active power injection, and line active power flow) is used to evaluate the FDIA detection capabilities of the model when relying only on sensor measurements. The multimodal dataset combining measurement data with topological features (including bus active power injection, line active power flow, and topological features extracted by VGAE) is applied to assess the enhancement effect of topological information on FDIA detection. We used TCN-GRU, LSTM [38], CNN [21], and CNN-LSTM [39] models for FDIA localization detection. The related results are shown in Table 1.

It is easy to observe that in IEEE14 and IEEE118 bus systems, compared with only using measured data, the F1 scores of TCN-GRU, LSTM, CNN, and CNN-LSTM algorithms using multi-modal fusion data increased by 0.049, 0.052, 0.038, 0.042 on average, and the average RACC raises by 0.064, 0.086, 0.071, 0.069, respectively. This result can be caused by the measurement data only providing information about the operational status of the power grid. In contrast, the topological features derived by VGAE can offer the information about the structure and topological relationships of the power system. By merging these two features, we can depict the state of the power system more comprehensively to enhance the accuracy of localization detection.

5.4. Model comparisons

To determine the best number of hidden layers for the VGAE-TCN-GRU model and evaluate its detection performance, we employ comparative algorithms, including VGAE-LSTM, VGAE-CNN, and VGAE-CNN-LSTM. These models are set up with 2, 3, 4, 5, and 6 hidden layers, a batch size of 64, and a learning rate of 0.03. Comparative trials are performed on a multimodal dataset with a fixed attack intensity of 2 and noise intensity of 0.2. The dataset consists of bus active power injection, line active power flow, and topology features extracted by VGAE. The results of the tests are presented in Table 2.

It can be obviously noticed that VGAE-TCN-GRU outperforms the other models in terms of F1 score and RACC for localization detection in the IEEE14 and IEEE 118 bus system, achieving more than 90% classification accuracy. The F1 score and RACC gradually boost in the circumstance that the number of hidden layers increases from 2 to 4. When the number of hidden layers grows to 5–6, there is a trend of performance decline. This phenomenon can be regarded as a degradation problem, i.e., with the increasement of depth of the network, the performance arrives at a saturation point and deteriorates rapidly [40]. When the number of hidden

Table 2
Comparison results of VGAE-TCN-GRU with various algorithms.

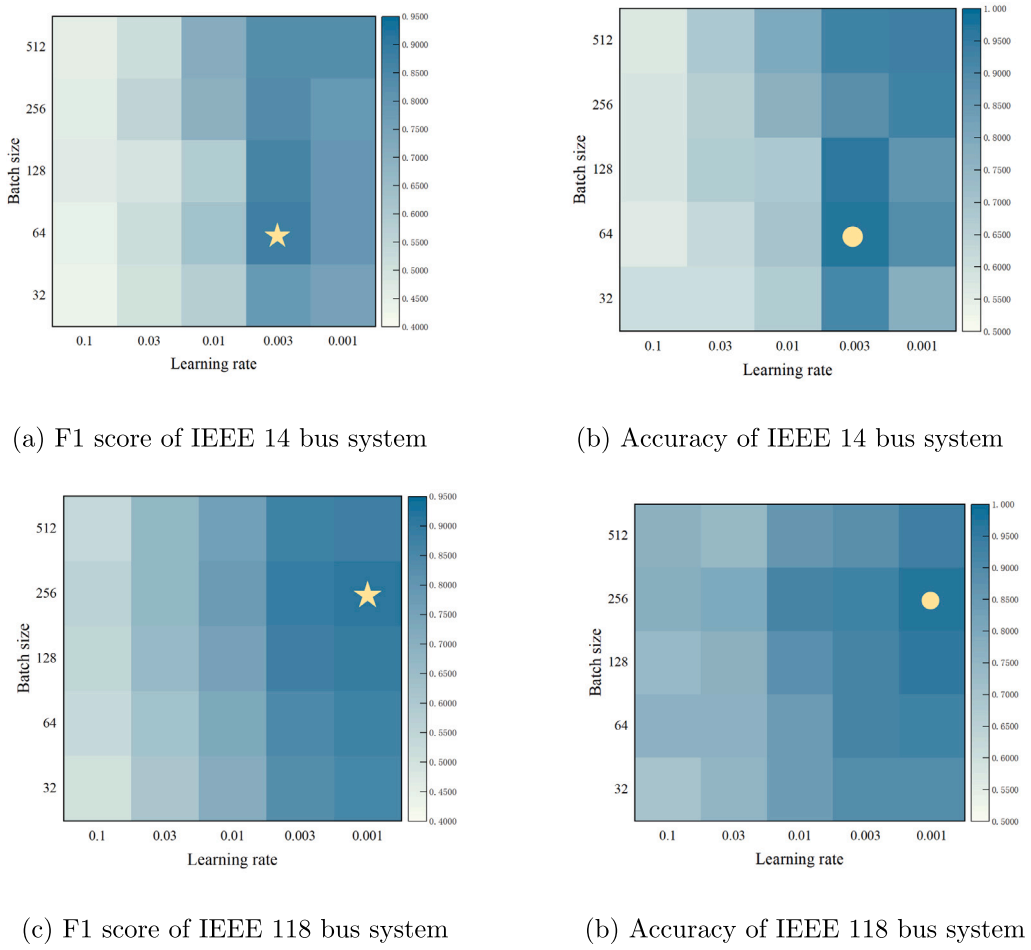
System	Detection algorithm	Number of layers	Location detection indicator				Classification detection indicator	Number of parameters	Time of localization detection /ms
			Recall	Precision	F1	RACC	Accuracy		
IEEE 14 bus system	VGAE-TCN-GRU	2	0.878	0.949	0.912	0.853	0.968	143,299	10
		3	0.901	0.943	0.922	0.876	0.975	225,603	15
		4	0.905	0.942	0.923	0.898	0.982	250,435	18
		5	0.889	0.941	0.914	0.862	0.971	349,123	34
		6	0.879	0.919	0.899	0.851	0.965	447,811	47
	VGAE-LSTM	2	0.801	0.824	0.812	0.786	0.844	1,134,371	35
		3	0.814	0.840	0.827	0.808	0.865	1,298,723	53
		4	0.823	0.835	0.829	0.812	0.907	1,430,307	62
		5	0.818	0.831	0.824	0.805	0.856	1,561,891	96
		6	0.812	0.827	0.819	0.799	0.848	1,693,475	128
	VGAE-CNN	2	0.856	0.867	0.862	0.776	0.832	232,803	4
		3	0.862	0.871	0.866	0.816	0.912	298,467	5
		4	0.857	0.893	0.875	0.825	0.926	339,555	6
		5	0.846	0.878	0.861	0.821	0.918	380,643	7
		6	0.83	0.888	0.858	0.819	0.913	421,731	8
	VGAE-CNN-LSTM	2	0.834	0.865	0.849	0.815	0.91	956,451	19
		3	0.846	0.881	0.863	0.826	0.928	1,120,547	21
		4	0.865	0.904	0.884	0.837	0.941	1,284,643	23
		5	0.854	0.892	0.872	0.834	0.935	1,448,739	31
		6	0.842	0.883	0.862	0.828	0.933	1,580,323	33
IEEE 118 bus system	VGAE-TCN-GRU	2	0.849	0.894	0.871	0.817	0.911	1,788,945	180
		3	0.892	0.931	0.911	0.854	0.971	1,813,713	320
		4	0.901	0.937	0.919	0.863	0.977	1,838,545	381
		5	0.900	0.931	0.915	0.852	0.966	1,863,313	523
		6	0.898	0.919	0.908	0.835	0.958	1,888,081	661
	VGAE-LSTM	2	0.765	0.822	0.792	0.768	0.817	7,279,537	898
		3	0.807	0.825	0.815	0.784	0.841	7,443,889	1120
		4	0.808	0.826	0.817	0.798	0.852	7,575,473	1356
		5	0.7995	0.819	0.809	0.776	0.835	7,707,057	1652
		6	0.796	0.790	0.7932	0.756	0.806	7,838,641	1928
	VGAE-CNN	2	0.806	0.837	0.8217	0.774	0.829	3,543,601	49
		3	0.826	0.866	0.8459	0.789	0.848	3,576,497	66
		4	0.839	0.886	0.861	0.808	0.867	3,609,393	75
		5	0.835	0.874	0.854	0.804	0.851	3,642,289	89
		6	0.828	0.856	0.842	0.786	0.846	3,676,185	93
	VGAE-CNN-LSTM	2	0.814	0.864	0.838	0.798	0.854	3,658,673	408
		3	0.835	0.883	0.858	0.813	0.906	3,839,153	428
		4	0.857	0.896	0.876	0.821	0.921	4,176,433	559
		5	0.850	0.908	0.878	0.833	0.933	4,365,105	591
		6	0.844	0.882	0.862	0.827	0.928	4,553,777	642

layers is increased from 3 to 4, there is an increasement in the metrics, but the change is small. Moreover, considering the issue that the number of model parameters and runtime boost with the number of hidden layers, we design the model architecture as 3 layers including 2 GRU layers and 1 TCN layer to obtain a balance between detection effect and computational complexity.

5.5. Parameter optimization of VGAE-TCN-GRU

To choose the optimal batch size and learning rate combination for VGAE-TCN-GRU, the number of hidden layers is set to the optimal level, and various combinations of learning rates (0.1, 0.03, 0.01, 0.003, 0.001) and batch sizes (32, 64, 128, 256, 512) are tested. Comparative experiments are conducted on a multimodal dataset with a fixed attack strength of 2 and noise intensity of 0.2, comprising bus active power injection, line active power flow, and topological features acquired by VGAE.

Fig. 5 shows the F1 scores for localization and accuracy for classification of VGAE-TCN-GRU for different combinations of learning rate and batch size. In the IEEE 14 bus system, the maximum F1 score and accuracy are achieved when the learning rate is set to 0.003 and the batch size is 64, reaching 0.922 and 0.975, respectively. In the IEEE 118-bus system, the maximum F1 score and accuracy are obtained with a learning rate of 0.001 and a batch size of 256, reaching 0.911 and 0.971, respectively. When the learning rate is 0.1, 0.03, 0.01 in IEEE 14 bus system and 0.1, 0.03, 0.01, 0.003 in IEEE 118 bus system, the model parameters neglect the optimal solution of the loss function due to the high learning rate, resulting in VGAE-TCN-GRU failing to converge to the optimal state. In addition, for the IEEE 14 bus system, the lower learning rate (0.001) may cause the training process to fall into a local optimum. On the other hand, some of the combinations display a decrease in the F1 score and RACC with the remarkable increase of batch size, leading to overfitting of the model. Based on the results of different combinations of learning rates and batch sizes, we found that VGAE-TCN-GRU showed the most satisfactory performance in IEEE 14 and IEEE 118 bus systems with learning rates of 0.003 and 0.001 and batch sizes of 64 and 256, respectively.



(a) F1 score of IEEE 14 bus system

(b) Accuracy of IEEE 14 bus system

(c) F1 score of IEEE 118 bus system

(b) Accuracy of IEEE 118 bus system

Fig. 5. Parameter optimization for learning rate and batch size.

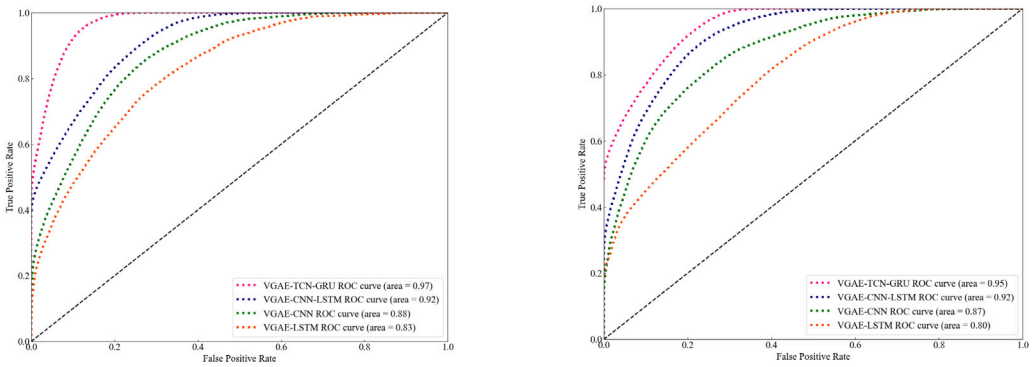
5.6. Model performance

To evaluate the detection performance of VGAE-TCN-GRU, the number of hidden layers, batch size, and learning rate are set to the optimal values identified in the previous experiments. Comparison tests are implemented on a multimodal dataset with a fixed attack strength of 2 and noise intensity of 0.2, including bus active power injection, line active power flow, and topological features achieved by VGAE. The performance of the localization detection model is estimated using the Area Under the Curve (AUC) of the ROC curve, which is the area between the false positive rate and the true positive rate.

Fig. 6 shows the receiver operating characteristic curve (ROC). We use the area under the curve (AUC), i.e., the area between the false positive rate (FPR) and the true positive rate TPR, to evaluate the performance of the detection model. It can be seen that the AUC of the proposed method is higher than that of VGAE-CNN, VGAE-LSTM, and VGAE-CNN-LSTM in both IEEE 14 and IEEE 118 bus systems (Approximating to 1). This outcome indicates the ideal detection capability of our method. Because VGAE can effectively extract grid topological features, GRU has the memory capability to capture the temporal patterns and trends in the grid data, and TCN can capture the features on different time scales through convolution operation, combining the three algorithms can better deal with the temporal features in the grid data.

5.7. Robustness of detection model

To assess the robustness of the proposed method under various attack and noise intensities, this study sets the number of hidden layers, batch size, and learning rate to optimal values, with the noise standard deviation set as 0.2. The robustness under different attack intensities is verified by setting the attack strength for both complete and incomplete information FDIA to 1, 2, 3, 4, and 5, respectively. The robustness under different noise intensities is assessed with the attack strength designated at 2 and noise intensities set to 0.1, 0.2, 0.3, 0.4, and 0.5, respectively.



(a) ROCs of IEEE 14 bus system (b) ROCs of IEEE 118 bus system

Fig. 6. Performance comparison of VGAE-TCN-GRU with different algorithms for IEEE 14 and IEEE 118 bus systems.

Fig. 7 shows the detection performance of each model under the same noise intensity with different attack intensities. The results show that VGAE-TCN-GRU is superior to VGAE-CNN, VGAE-LSTM, and VGAE-CNN-LSTM in terms of location detection and classification detection under 5 attack intensities. As the attack intensity increases, the F1 score, RACC, and accuracy of all detection models also grow. This is because the abnormal behavior of the power system has more obvious spatio-temporal characteristics and is easier to capture by the model. Specifically, in the IEEE 14 and 118 bus systems, when the attack intensity is 1, the average values of recall, precision, F1 score, RACC, and accuracy of VGAE-TCN-GRU reach 0.868, 0.92, 0.891, 0.845, and 0.946, respectively. When the attack intensity is 5, the average recall, precision, F1 score, RACC, and accuracy of VGAE-TCN-GRU achieve 0.93,0.95, 0.939, 0.914, and 0.986, respectively, indicating that the proposed detection model is robust under different attack intensities.

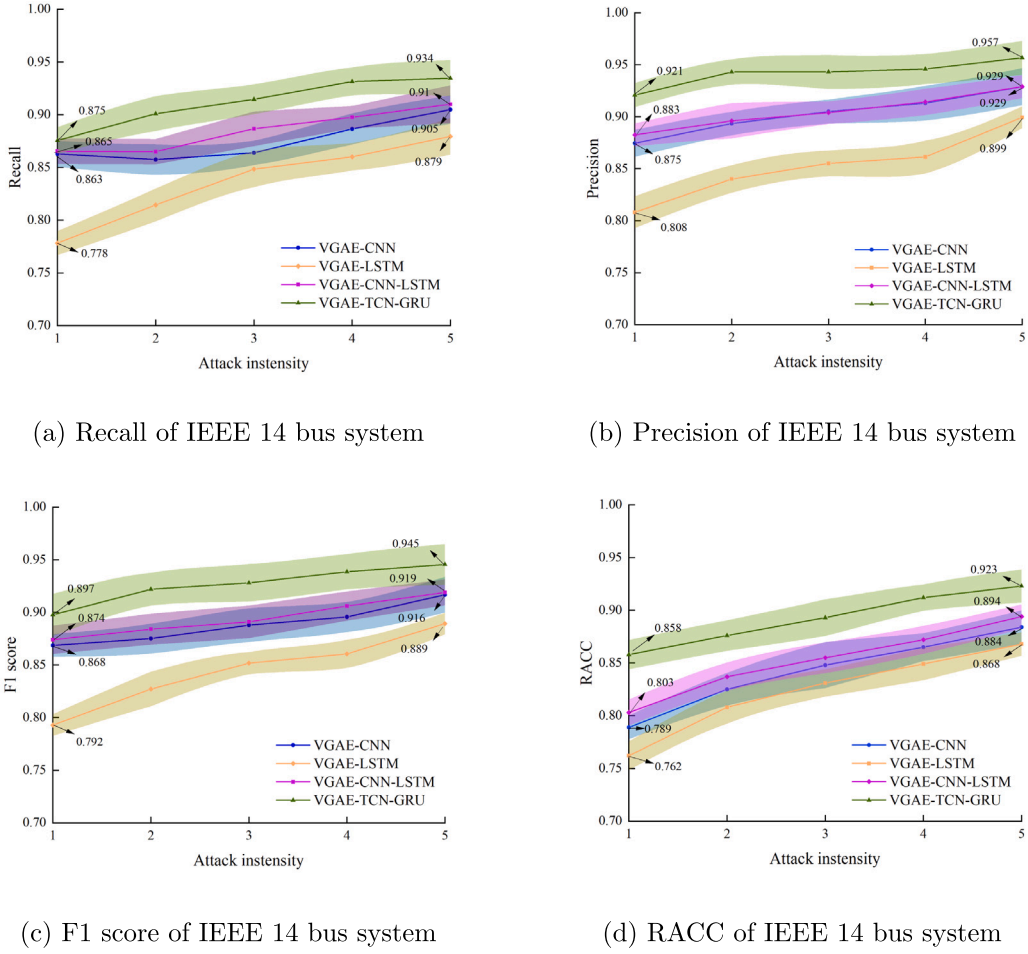
Fig. 8 shows the detection results of different models under the identical attack intensity with varied noise intensities. For a constant attack strength, as the noise strength increases, the F1 score and RACC of every model decrease to some extent. This is because the larger the noise standard deviation is, the more interference will have a negative impact on the model to obtain effective features. The growth in noise strength enhances the difficulty of distinguishing between real and anomalous data while the intensity of the attack remains unchanged. For the IEEE 14 and 118 bus systems, in the case that the noise intensity equals 0.1, the average values of recall, precision, F1 score, RACC, and accuracy of VGAE-TCN-GRU for FDIA detection are 0.92, 0.941, 0.929, 0.885, and 0.979. In the condition that the noise intensity is 0.5, the average values of recall, precision, F1 score, RACC, and accuracy are 0.871, 0.881, 0.891, 0.845, and 0.946. Although the performance drops at higher noise intensities, it stays above the level of 0.8, suggesting that the proposed detection model has more reasonable robustness under different noise intensities and can effectively cope with the effect of altered noise.

The high robustness of the FDIA detection is due to the ability of VGAE to capture complex spatial dependencies between nodes through learning in latent space. This helps in identifying potential anomalies or inconsistencies in measurement data. The TCN can grab long-term dependencies in measurement time series, whereas the GRU enhances the response to short-term critical events within these series. These three algorithms complement each other's weaknesses to enhance the robustness of FDIA detection.

5.8. Detection performance of different attack models

In order to evaluate the detection ability of the proposed algorithm across different attack models, the best configurations of the number of hidden layers, batch size, and learning rate are selected. With the attack intensity fixed at 2 and the noise standard deviation appointed at 0.2, the tests are conducted using FDIA data under complete information, FDIA data under incomplete information, and topological attack data.

Fig. 9 reveals the F1 score and RACC for localization detection of FDIA with complete information, FDIA with incomplete information, and topology attack under an attack strength of 2 and a noise strength of 0.2. It can be seen that FDIA with incomplete information has the highest F1 score (0.952) and RACC (0.876) in the IEEE 14 bus system, respectively. This is followed by the topology attack, with F1 score and RACC of 0.921 and 0.885, respectively. FDIA with complete information has the lowest F1 score and RACC of 0.896 and 0.834, respectively. A similar trend is shown for the IEEE 118 bus system. This is because for FDIA with complete information, the attacker can obtain all the information about the power system to construct the attack vector accurately. Conversely, when FDIA with incomplete information uses the PCA approximation method to forge false data, the attacker can just gather partial details and cannot simulate the whole system behavior, resulting in a more significant deviation in the constructed attack vectors compared to that of FDIA with complete information. Accordingly, the temporal characteristics of anomalous data are more prominent and more easily detected. Moreover, the topological features of topology attacks may be affected by changes in the topology of the power system, leading to a slightly lower performance in localization detection. In general, the F1 score and



(a) Recall of IEEE 14 bus system

(b) Precision of IEEE 14 bus system

(c) F1 score of IEEE 14 bus system

(d) RACC of IEEE 14 bus system

Fig. 7. Identification comparison of VGAE-TCN-GRU with various models under varied attack strengths.

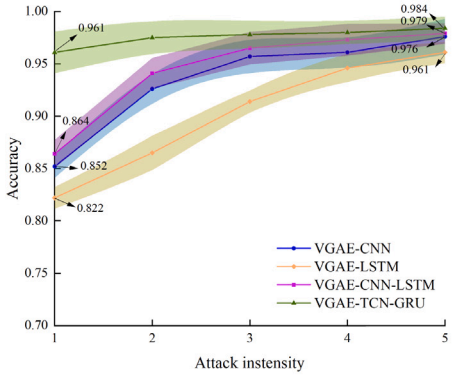
Table 3
Comparison of advantages and disadvantages between the proposed algorithm and related algorithms.

Algorithm	Forte	Deficiency
CNN	Spatial feature extraction	Weak time series data processing capability
LSTM	Time series feature extraction	Inadequate space data processing capability
CNN-LSTM	Spatial and temporal feature extraction	Feeble non-Euclidean data processing and Incomplete temporal feature extraction
VGAE-GRU-TCN	Remarkable detection capability, high robustness	No applicability for AC FDIA detection

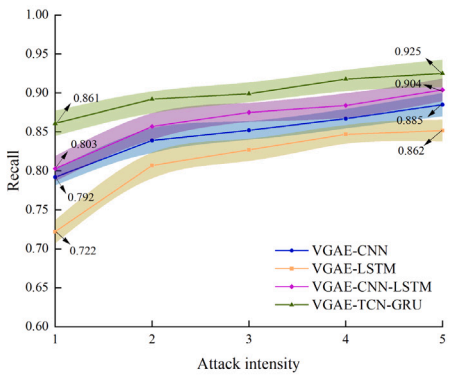
RACC of the proposed method are above 0.8 for localization detection of the three types of attacks in the IEEE 14 and IEEE 118 systems.

Based on the above analysis, the advantages and disadvantages of the proposed method and comparative methods are summarized in Table 3.

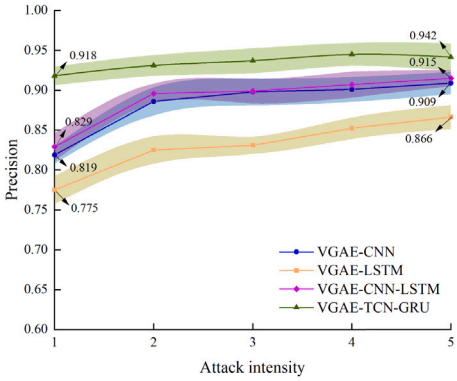
The method proposed in this study can accurately identify and locate FDIA in DC systems. In the future, incorporating methods such as extended Kalman filtering into this approach could enable FDIA detection in AC systems. In addition, the proposed method, when combined with cloud computing technologies, can be utilized to build an online FDIA detection system to properly deal with the challenges of large-scale data storage in real smart grids and the difficulty of switching FDIA detection between DC and AC systems.



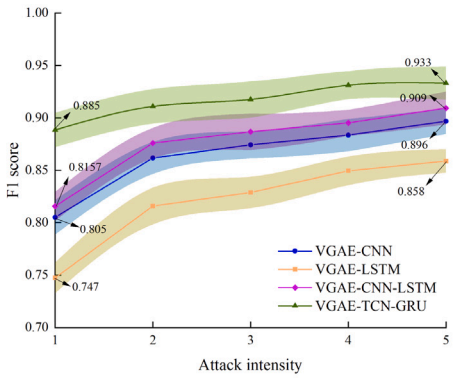
(e) Accuracy of IEEE 14 bus system



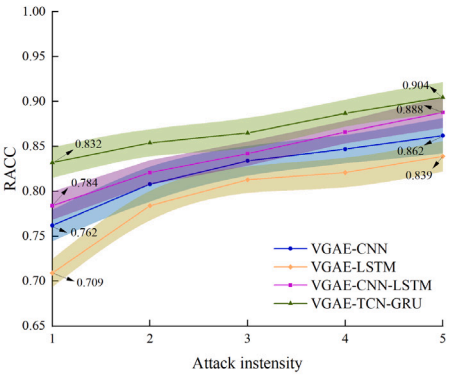
(f) Recall of IEEE 118 bus system



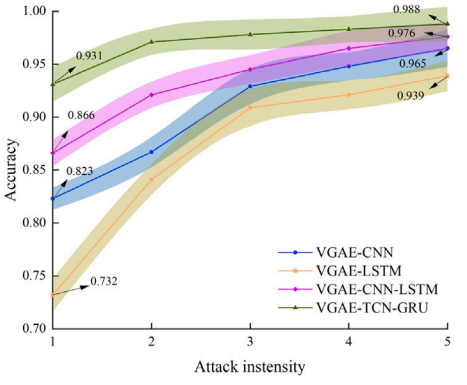
(g) Precision of IEEE 118 bus system



(h) F1 score of IEEE 118 bus system

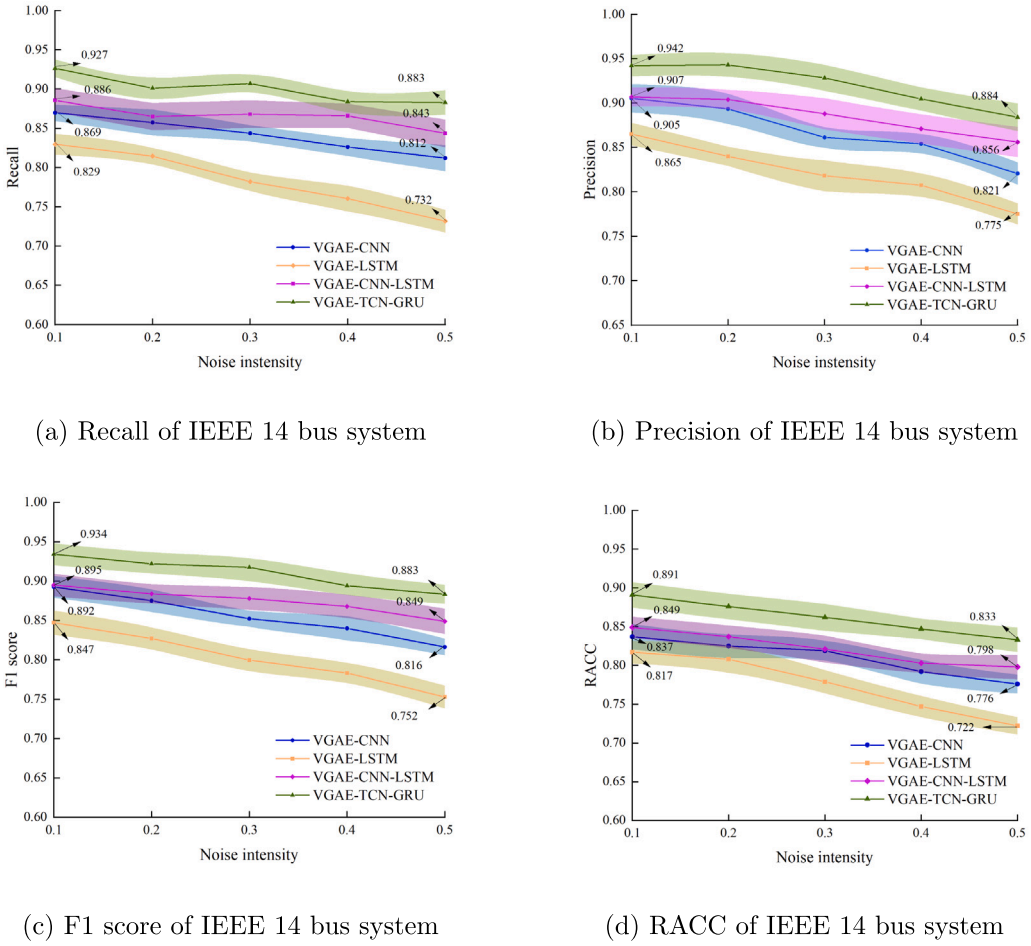


(i) RACC of IEEE 118 bus system



(j) Accuracy of IEEE 118 bus system

Fig. 7. (*continued*).



(a) Recall of IEEE 14 bus system

(b) Precision of IEEE 14 bus system

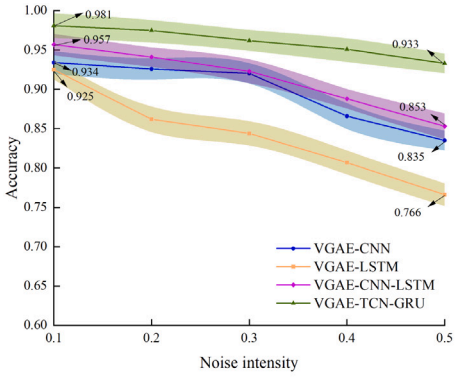
(c) F1 score of IEEE 14 bus system

(d) RACC of IEEE 14 bus system

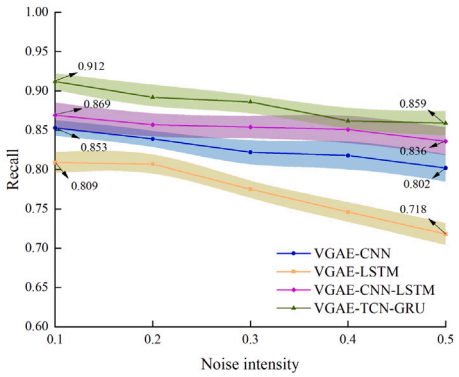
Fig. 8. Recognition comparison of VGAE-TCN-GRU with various models under varying noise intensities.

6. Conclusion

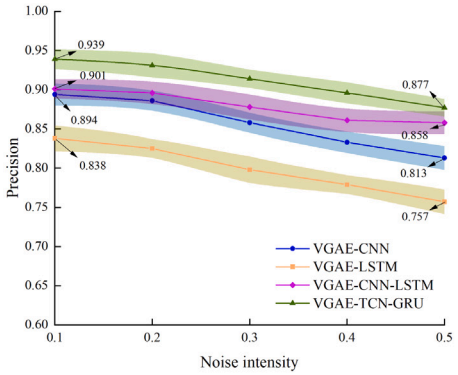
The VGAE-TCN-GRU proposed in this study can accurately localize and classify different types of FDIAs. Through numerical experiments in IEEE 14 and 118 bus systems, it can be seen that by use of the combination of topological features and measurement data, the F1 score and RACC are on average 0.049 and 0.064 higher than those of employing measured data alone as the dataset. In the IEEE 14 bus system, when the number of hidden layers is 3, the learning rate is 0.003, and the batch size is 64. For the IEEE 118 bus system, when the number of hidden layers is 3, the learning rate is 0.001, and the batch size is 256, the performance of the VGAE-TCN-GRU algorithm is the most satisfactory. Moreover, the AUC value of the proposed detection method is improved by 0.085, 0.145, and 0.04 on average compared with VGAE-CNN, VGAE-LSTM, and VGAE-CNN-LSTM, respectively. The F1 score, RACC, and accuracy remain above 0.867, 0.818, and 0.912 under different attack and noise intensities. In addition, the addressed method has the best detection effect on FDIA with incomplete information, followed by topology attack, and finally FDIA with complete information. In addition, the detection model of this paper is created as a multi-label classification model, which can accurately detect topological and non-topological attacks by treating each label as a binary classification issue. However, it cannot differentiate between different types of attacks. Future work includes using ensemble learning to stack multiple classification models to achieve FDIA localization detection and classify multiple types of attacks. On the other hand, DC systems are currently primarily used in scenarios such as hydroelectric and thermal power plants and various substations. However, with the vigorous development of new energy and the integration of source-grid-load-storage, the proportion of DC microgrids will gradually increase. Obviously, the covert nature of FDIA can cause severe harms to these systems. The method proposed in this study effectively solves the predicament of deteriorating detection performance induced by the diverse topologies of power systems and the various FDIA construction methods in DC microgrids, future work may consider using nonlinear state estimation techniques such as extended Kalman filter and unscented Kalman filter for FDIA detection under AC conditions.



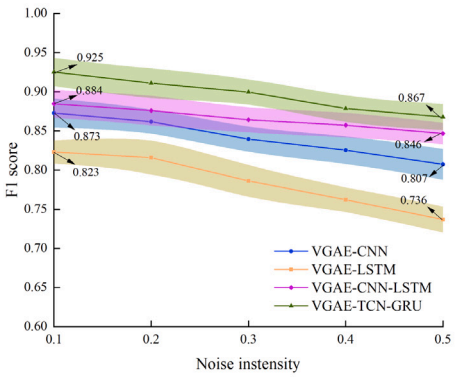
(e) Accuracy of IEEE 14 bus system



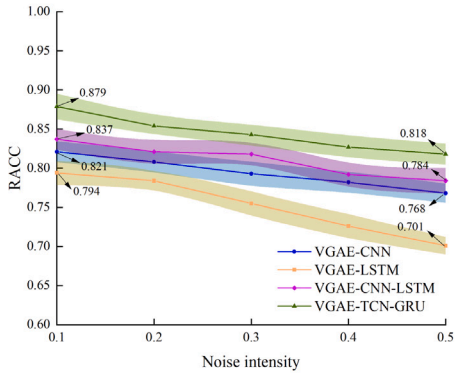
(f) Recall of IEEE 118 bus system



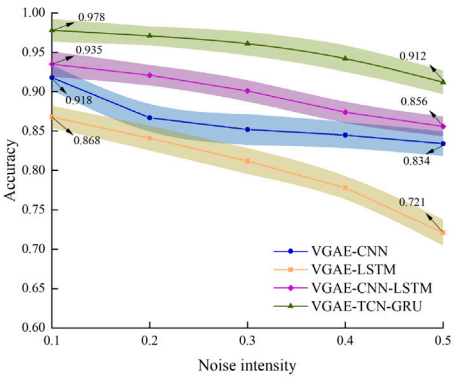
(g) Precision of IEEE 118 bus system



(h) F1 score of IEEE 118 bus system

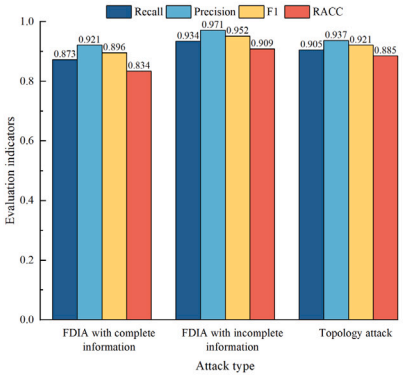


(i) RACC of IEEE 118 bus system

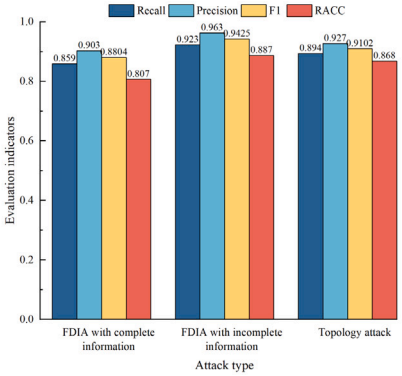


(j) Accuracy of IEEE 118 bus system

Fig. 8. (*continued*).



(a) Performance of IEEE 14 bus system



(b) Performance of IEEE 118 bus system

Fig. 9. Performance comparison of VGAE-TCN-GRU with different algorithms for IEEE 14 and IEEE 118 bus systems.

CRediT authorship contribution statement

Jun Wang: Conceptualization, Methodology, Writing – original draft. **Haoran Chen:** Conceptualization, Methodology, Writing – review. **Yifei Si:** Data processing and analysis. **Yonghai Zhu:** Methodology, Supervision. **Tianci Zhu:** Algorithm, Writing – related work. **Shanshan Yin:** Conceptualization, Methodology. **Bo Liu:** Data processing and analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Science and Technology Development Project in Henan Province, China (232102241042).

References

[1] Fadlullah Zubair Md, Fouda Mostafa M, Kato Nei, Takeuchi Akira, Iwasaki Noboru, Nozaki Yousuke. Toward intelligent machine-to-machine communications in smart grid. IEEE Commun Mag 2011;49(4):60–5.
[2] Raja D Jim Solomon, Sriranjani R, Parvathy A, Hemavathi N. A review on distributed denial of service attack in smart grid. In: 2022 7th international conference on communication and electronics systems. ICCES, IEEE; 2022, p. 812–9.
[3] Zhang Yichi, Wang Lingfeng, Xiang Yingmeng, Ten Chee-Wooi. Power system reliability evaluation with SCADA cybersecurity considerations. IEEE Trans Smart Grid 2015;6(4):1707–21.
[4] Dragičević Tomislav, Lu Xiaonan, Vasquez Juan C, Guerrero Josep M. DC microgrids—Part I: A review of control strategies and stabilization techniques. IEEE Trans Power Electron 2015;31(7):4876–91.
[5] Liu Yao, Ning Peng, Reiter Michael K. False data injection attacks against state estimation in electric power grids. ACM Trans Inf Syst Secur 2011;14(1):1–33.
[6] Esmalifalak Mohammad, Nguyen Huy, Zheng Rong, Han Zhu. Stealth false data injection using independent component analysis in smart grid. In: 2011 IEEE international conference on smart grid communications (smartGridComm). IEEE; 2011, p. 244–8.
[7] Yu Zong-Han, Chin Wen-Long. Blind false data injection attack using PCA approximation method in smart grid. IEEE Trans Smart Grid 2015;6(3):1219–26.
[8] Bi Suzhi, Zhang Ying Jun. Using covert topological information for defense against malicious attacks on DC state estimation. IEEE J Sel Areas Commun 2014;32(7):1471–85.
[9] Kim Jinsub, Tong Lang. On topology attack of a smart grid: Undetectable attacks and countermeasures. IEEE J Sel Areas Commun 2013;31(7):1294–305.
[10] Musleh Ahmed S, Chen Guo, Dong Zhao Yang. A survey on the detection algorithms for false data injection attacks in smart grids. IEEE Trans Smart Grid 2019;11(3):2218–34.
[11] Manandhar Kebina, Cao Xiaojun, Hu Fei, Liu Yao. Detection of faults and attacks including false data injection attack in smart grid using Kalman filter. IEEE Trans Control Netw Syst 2014;1(4):370–9.
[12] Gu Yun, Liu Ting, Wang Dai, Guan Xiaohong, Xu Zhanbo. Bad data detection method for smart grids based on distributed state estimation. In: 2013 IEEE international conference on communications. ICC, IEEE; 2013, p. 4483–7.
[13] Sahoo Subham, Mishra Sukumar, Peng Jimmy Chih-Hsien, Dragičević Tomislav. A stealth cyber-attack detection strategy for DC microgrids. IEEE Trans Power Electron 2018;34(8):8162–74.

[14] Li Boda, Ding Tao, Huang Can, Zhao Junbo, Yang Yongheng, Chen Ying. Detecting false data injection attacks against power system state estimation with fast go-decomposition approach. IEEE Trans Ind Inf 2018;15(5):2892–904.

[15] He Youbiao, Mendis Gihan J, Wei Jin. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. IEEE Trans Smart Grid 2017;8(5):2505–16.

[16] Xiong Xiaoping, Hu Siding, Sun Di, Hao Shaolei, Li Hang, Lin Guangyang. Detection of false data injection attack in power information physical system based on SVM–GAB algorithm. Energy Rep 2022;8:1156–64.

[17] Zhang Ying, Wang Jianhui, Chen Bo. Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach. IEEE Trans Smart Grid 2020;12(1):623–34.

[18] James JQ, Hou Yunhe, Li Victor OK. Online false data injection attack detection with wavelet transform and deep neural networks. IEEE Trans Ind Inf 2018;14(7):3271–80.

[19] Zhu Jianxin, Meng Wenchao, Sun Mingyang, Yang Jun, Song Zhuo. FLLF: A fast-lightweight location detection framework for false data injection attacks in smart grids. IEEE Trans Smart Grid 2023.

[20] Zia Muhammad Fahad, Inayat Usman, Noor Wafa, Pangracious Vinod, Benbouzid Mohamed. Locational detection of false data injection attack in smart grid based on multilabel machine learning classification methods. In: 2023 IEEE IAS global conference on renewable energy and hydrogen technologies (globConHT). IEEE; 2023, p. 1–5.

[21] Wang Shuoyao, Bi Suzhi, Zhang Ying-Jun Angela. Locational detection of the false data injection attack in a smart grid: A multilabel classification approach. IEEE Internet Things J 2020;7(9):8218–27.

[22] Mukherjee Debottam, Chakraborty Samrat, Ghosh Sandip. Deep learning-based multilabel classification for locational detection of false data injection attack in smart grids. Electr Eng 2022;104(1):259–82.

[23] Kipf Thomas N, Welling Max. Variational graph auto-encoders. 2016, arXiv preprint arXiv:1611.07308.

[24] Xie Qianqian, Huang Jimin, Du Pan, Peng Min, Nie Jian-Yun. Inductive topic variational graph auto-encoder for text classification. In: Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: human language technologies. 2021, p. 4218–27.

[25] Zhang Le, Cheng Wei, Liu Xue, Chen Xuefeng, Chang Fengtian, Hong Junying, et al. System-level anomaly detection for nuclear power plants using variational graph auto-encoders. In: 2021 IEEE international conference on sensing, diagnostics, prognostics, and control. SDPC, IEEE; 2021, p. 180–5.

[26] Chung Junyoung, Gulcehre Caglar, Cho KyungHyun, Bengio Yoshua. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014, arXiv preprint arXiv:1412.3555.

[27] ur Rehman Saif, Khaliq Mubashir, Imtiaz Syed Ibrahim, Rasool Aamir, Shafiq Muhammad, Javed Abdul Rehman, et al. DIDDOS: An approach for detection and identification of distributed denial of service (ddos) cyberattacks using gated recurrent units (GRU). Future Gener Comput Syst 2021;118:453–66.

[28] Xu Congyuan, Shen Jizhong, Du Xin, Zhang Fan. An intrusion detection system using a deep neural network with gated recurrent units. IEEE Access 2018;6:48697–707.

[29] Dickey Joshua, Borghetti Brett, Junek William. Improving regional and teleseismic detection for single-trace waveforms using a deep temporal convolutional neural network trained with an array-beam catalog. Sensors 2019;19(3):597.

[30] Cheng Yongliang, Xu Yan, Zhong Hong, Liu Yi. HS-TCN: A semi-supervised hierarchical stacking temporal convolutional network for anomaly detection in IoT. In: 2019 IEEE 38th international performance computing and communications conference. IPCCC, IEEE; 2019, p. 1–7.

[31] Tong Cheng, Zhang Linghua, Li Hao, Ding Yin. Temporal inception convolutional network based on multi-head attention for ultra-short-term load forecasting. IET Gener Transm Distribution 2022;16(8):1680–96.

[32] Monticelli Alcir. Electric power system state estimation. Proc IEEE 2000;88(2):262–82.

[33] Meriem Majdoub, Bouchra Cheddadi, Abdelaziz Belfqih, Jamal Sabri Omar Boukherouaa, Nazha Cherkaoui, et al. Study of state estimation using weighted-least-squares method (WLS). In: 2016 international conference on electrical sciences and technologies in maghreb. CISTEM, IEEE; 2016, p. 1–5.

[34] Ariantara Hazel, Sarjiya Sarjiya, Hadi Sasongko Pramono. The solution for optimal power flow (OPF) method using differential evolution algorithm. IJITEE (Int J Inf Technol Electr Eng) 2017;1(1):19–24.

[35] Uzair Muhammad, Jamil Noreen. Effects of hidden layers on the efficiency of neural networks. In: 2020 IEEE 23rd international multitopic conference. INMIC, IEEE; 2020, p. 1–6.

[36] He Fengxiang, Liu Tongliang, Tao Dacheng. Control batch size and learning rate to generalize well: Theoretical and empirical evidence. Adv Neural Inf Process Syst 2019;32.

[37] Radiuk Pavlo M. Impact of training set batch size on the performance of convolutional neural networks for diverse datasets. Inf Technol Manag Sci 2017;20(1):20–4.

[38] Mukherjee Debottam, Chakraborty Samrat, Abdelaziz Almoataz Y, El-Shahat Adel. Deep learning-based identification of false data injection attacks on modern smart grids. Energy Rep 2022;8:919–30.

[39] Wu Yi, Wang Qiankuan, Guo Naiwang, Tian Yingjie, Li Fengyong, Su Xiangjing. Efficient multi-source self-attention data fusion for fdia detection in smart grid. Symmetry 2023;15(5):1019.

[40] He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, p. 770–8.