INSIDER THREAT DETECTION MODEL OF POWER SYSTEM BASED ON LSTM-ATTENTION

Xiaojun ZUO[1], Fengru YAN[2], Botao HOU[3]*, Ze CHEN[4], Yuling GUO[5]

An insider threat detection model ITDBLA based on LSTM- Attention is proposed in this paper. First, the user behavior features, role behavior features, user behavior sequences, and psychological data are extracted from multi-source heterogeneous log files to describe users' daily activities. Secondly, we use the long-short term memory network and the attention mechanism to learn the user's normal behavior patterns and predict the next state, thereby calculating the deviation between the true data and the predicted data. Finally, the multi-layer perceptron is used to make comprehensive decisions based on these deviations and identify abnormal behaviors.

Keywords: LSTM; Attention; the power system; user and entity behavior analysis; insider threat detection

1. Introduction

As a critical national infrastructure, the power system guarantees people's daily lives and work and is related to the country's economic lifeline and energy security. The construction of electric power information security is mainly subject to internal and external threats, among which virus attacks, natural disasters, and hacker intrusions are external threats. However, with the improvement of the information security mechanism of the power system, the threshold of external threats has been increasing, which has curbed the external threats to a certain extent. The insider threat is more subtle in its manifestation, more harmful, and more difficult to detect. Domestic and foreign scholars have long and in-depth research on external threats, such as encryption isolation gateway [1], data encryption and authentication [2], and intrusion detection systems [3], which can effectively prevent external threats. However, applying external threat detection methods to insider threat detection didn't achieve the desired results. Therefore,

1 Eng., State Grid Hebei Electric Power Research Institute, Hebei Province, China, e-mail: ids2020114027@126.com
2 Master, School of Information Science and Engineering, Hebei University of Science and Technology, China, e-mail: yfr970510@163.com
3 Eng., State Grid Hebei Electric Power Research Institute, Hebei Province, China, e-mail: 2397994927@qq.com
4 Eng., State Grid Hebei Electric Power Research Institute, Hebei Province, China, e-mail: 2189137081@qq.com
5 Eng., State Grid Hebei Electric Power Research Institute, Hebei Province, China, e-mail: 402608074@qq.com

establishing an effective insider threat detection model has become a research priority for power system security experts.

Important information such as business information and user data is reserved in the internal network of the power system, and those information is transmitted through private lines. Since data transmission based on private lines is theoretically secure, most of the data in the power system are transmitted in plaintext, and the behavior monitoring of internal personnel is lacking. Whereas, with the penetration technology constantly changing, most of the attacks lurk in the intranet in the early stage and then gradually gain high privileges and cause damage to the system through system vulnerabilities or management flaws. The insiders are located inside the organization and have specific legal operation authority. Therefore, the abnormal behavior of insiders is less likely to be detected. The traditional insider threat detection technologies mainly include detection based on user commands [4], big data and machine learning [5], etc. Traditional threat detection mainly focuses on security incidents, such as Trojan horses and viruses. However, as attack methods become more and more complex, traditional threat detection technologies can no longer effectively detect threats in the network.

User and Entity Behavior Analytics (UEBA) technology [6] can be productively utilized for insider threat detection. It faces the behavior of users and entities and uses advanced data analysis methods to analyze user and entity behaviors. It also characterizes the baseline of user and entity behavior and then discovers the anomalies of users and entities. Gartner predicts [7] that user and entity behavior analysis technology will be used in 80% of network security detection methods by 2022. Designing an effective insider threat detection model that improves its accuracy, reduces false alarm rates, improves interpretability is one of the critical tasks of current power sector security research.

Most existing insider threat detection systems for user and entity behavior analysis [8-10] focus on learning user behavior patterns by comparing the behaviors of different users or the behaviors of the same user in different periods. These methods are aimed at the threat detection of a single domain and only establish a behavior baseline for the user, without considering the behavior baseline of the user's corresponding role and subjective factors like user personality. Although anomaly detection usually uses machine learning algorithms, such as SVM [11] and Isolation Forest [12], it is still difficult to select a suitable model with low false alarm rate but high accuracy rate. In addition, the existing algorithm is not specifically designed to detect insider threat. Consequently, the AUC score of the algorithm is not very high. The insider threat detection model using user and entity behavior analysis technology is improved in response to the above. The insider threat detection system based on LSTM-Attention (ITDBLA) model that consists of a combination of data pre-processing

and a user behavior analysis module based on LSTM-Attention is proposed. The model can perform insider threat detection based on the multi-source heterogeneous log data of the power system, and combine the psychological data, job role and other attributes of power system users to model the user's behavior habits from multiple perspectives. This model thereupon determines whether abnormal behaviors occur by deviations between real and predicted behaviors.

This article is arranged as follows. The related work is introduced in Section 1. Section 2 describes the insider threat detection model's details and workflow. In Section 3, the validity of the ITDBLA model is verified through experiments. The full text and proposes future work are summarized in Section 4.

2. Related Work

Experts and scholars in related fields have proposed different technical methods and solutions for insider threats. The rule-based method is one of the most utilized and developed methods. The general process is to establish a standard behavior profile by mining the association rules of the behavior, then to perform anomaly detection by analyzing the incoming instances and existing regulations. A method for detecting insider threats of intelligent grid data monitoring equipment based on behavior rules is proposed in literature [13]. By extracting three behavior regulations to describe the behavior specifications of each device, it can detect whether the behavior of the monitoring equipment deviates from the behavior specifications. A standard about stochastic Petri nets and behavior rules to control the insider threats of smart grid communication systems is proposed in literature [14]. The rule-based method has the advantages of simple process and rapid response. But there are several disadvantages of this method. Firstly, establishing a rule database requires a large amount of professional knowledge. Secondly, its effect depends on the update of the behavior database, and it cannot identify threats of unknown patterns.

Anomaly-based threat detection can effectively detect unknown threats. User and entity behavior analysis is behavior oriented. It has been used in enterprise internal behavior analysis [15], host intrusion detection [16], user profile research [17], complex behavior modeling [18], recommendation system [19], etc., and has achieved remarkable results. The development and improvement of UEBA provides an effective method for insider threat detection. Li et al. [20] study user behavior from the perspective of using the application window, collect and analyze user behavior data on the application window, and extract behavior features for detecting abnormal users and recognizing user's behavior changes. The detection method based on misuse cannot effectively detect unknown attacks, but its performance in detecting known attacks is much higher than the detection method based on anomaly. Mohammed et al. [21] propose a

hybrid detection method of anomaly and misuse detection. The random forest algorithm is used in the first layer, and the second layer uses the K nearest neighbor algorithm. Shashankam et al. [22] propose a solution for user and entity behavior analysis of the Niara's security analysis platform, which uses a singular value decomposition algorithm to detect anomaly from the traffic data collected by Niara's internal network. These user and entity behavior analysis technologies based on machine learning are not suitable for processing long sequences and have low efficiency in the application of insider threat detection, which leads to the bottleneck in the improvement of the threat detection rate.

With the development of neural networks, insider threat detection begins to utilize deep learning-based user and entity behavior analysis techniques widely. The key of insider threat detection is to establish a normal user behavior model and identify abnormal behaviors through behavior deviations. User behavior can be regarded as time-series data, and RNNs can handle serial data well. In the literature [23,24], recurrent neural networks (RNNs) are used to model user behavior as time series and predict the probability of anomalies. However, traditional RNNs suffer from the problem of "long-distance dependence," as the distance increases, RNNs cannot connect relevant information. Long Short-Term Memory Networks and attentional mechanisms can learn long-term sequential patterns better than RNNs, which can discover the implicit behavioral features in internal user behavior and significantly improve the detection rate. Literature [25] put forward a combination model of convolutional neural network based on the attention mechanism and the long short-term memory (AMCNN-LSTM) to detect accurately abnormalities. This model adopts a convolutional neural network based on the attention mechanism to capture significant fine-grained features to solve gradient dispersion and memory loss problems. Literature [26] uses various LSTMs on streaming data for scoring and then performs comprehensive scoring to achieve anomaly detection. Literature [27] uses the LSTM-CNN framework to discover the user's abnormal behaviors. First, it uses a long short-term memory network to learn user behaviors, extracts abstract temporal features, and then uses a convolutional neural network to detect insider threats. However, most of these user behavior modeling methods ignore the particularity of insider threats and do not take advantage of factors such as psychological data and the similarity of the behavior of the same role.

In response to the above-mentioned related work, on basis of user and entity behavior analysis technology, this study proposes an insider threat detection model based on the LSTM-Attention deep learning algorithm to advance the capability of power system insider threat detection. It not only fully utilizes the advantages of LSTM and attention mechanism in processing long sequences, but also considers the particularity of insiders. First, the model preprocesses multi-source heterogeneous power system log data and combines role behavior, personal

behavior, and psychological data for feature extraction. Then, the LSTM-Attention algorithm is used to learn the user behavior's normal pattern. When the behavior is abnormal, it outputs a higher mean square error. Finally, a Multilayer Perceptron (MLP) is used to make comprehensive decision-making to achieve the detection of insider threats in the power system.

3. Insider Threat Detection Model Based on LSTM-Attention

The power system insider threat detection model ITDBLA is mainly divided into two modules, the data preprocessing module and the user and entity behavior analysis module based on LSTM-Attention. The latter includes four models: behavior feature model, role feature model, behavior sequence model, and comprehensive decision model. The role feature is the mean value of the behavior feature under the same role, so the same model is used for the role feature and the behavior feature. The workflow of the ITDBLA model is seen in Fig. 1.

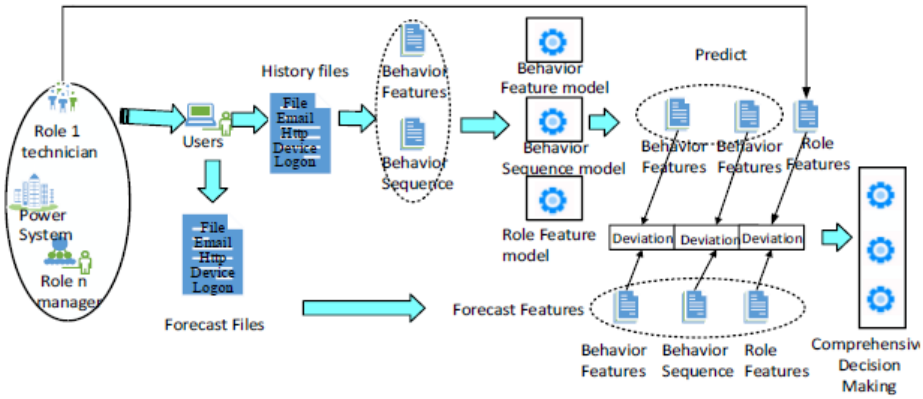

Fig. 1 The workflow of the ITDBLA model

In the data preprocessing module, first of all, we clean, integrate, and code the log data, and extract features from it. The design of the model in this research thinks over that the work of users in the same role is similar. According to their roles, the employees are grouped, such as technical personnel, human resources, DC transportation inspection personnel, etc. Then, the role features are extracted. Finally, the behavior sequence, the behavior feature, the role feature, and psychological data required by the model are generated respectively. In the user and entity behavior analysis module based on LSTM-Attention, the model considers three types of features, including behavior features, behavior sequences and role features. Three models are designed based on these three types of features, and historical data is used to learn the normal behavior patterns of users and then to make predictions about the next state. When insiders behave

abnormally, their real behavior will deviate from the predicted behavior, and the model will calculate the deviation between the behaviors. The specific workflow of this module is as follows. Firstly, user behavior and role behavior are modeled using LSTM. Secondly, the sequence of user behavior is analyzed and modeled by the attention mechanism. Finally, the comprehensive decision model MLP uses three deviations to detect anomalies in user behavior. When the model exhibits a high mean square error, it demonstrates that the user has aberrant behavior during this period.

3.1 Data Preprocess

The purpose of data preprocess is to extract relevant information from multi-source heterogeneous log files and convert them into a standardized representation. When abnormal behavior occurs, deep learning algorithms can detect deviations from them. The main process of the pre-processing phase is as follows. Firstly, data cleaning is performed for each log file, including the modification of error information, the deletion of redundant fields and the filling of missing values. Secondly, data integration of multiple log files. Each log file in the dataset is the operation of all users, which needs to be integrated after information extraction according to the users will be performed and sorted in chronological order. Since the user's behavior in a day can better reflect the user's behavior habits, the user's behavior is described in units of days in this paper. Finally, the extracted data are encoded. The numeric variables can be directly employed as the inputs for deep learning, while the categorical variables cannot be directly employed as input for deep learning algorithms, and the categorical data need to be encoded {'logon':1,'Connect':2,'Disconnect':3,'http':4,'email':5,' logoff':6}. Appropriate features play a crucial role in capturing the deviations between the real behavior and the user behavior predicted by the model. These deviations can indicate abnormal behavior and describe the degree of suspicious threat to the user. Recognizing abnormal users requires a comprehensive analysis of various user behaviors. Therefore, this paper extracts behavior features, behavior sequences, psychological data, and role features from different log files such as logon.csv, email.csv, and http.csv. The behavior features are described as shown in Table 1.

Table 1

Description of behavior features

Data	Describe
logon.csv	Working Day log on/log off, Weekend log on, Online time
	Num. weekday/weekend email send, Num. emails sending
email.csv	Num. internal/external email receive
	Num. attachments, Size of emails
file.csv	Num. exe/jpg/txt/pdf/zip file copy
http.csv	Num. websites, Num. career/news/tech sites
device.csv	Num. device, Num. weekday/weekend device

The sequence, role and psychological features are described as shown in Table 2, which portrays the features required by the model.

Table 2

Description of sequence, role and psychological features

Feature	Data	Describe
Sequence Feature	Behavior sequence within the time period	logon:1 Connect:2 Disconnect:3 http:4 email:5 logoff:6
Role Feature	Common behavior features of all users in the role	Average value of behavior features in the same role
Psychological Data	Five-Factor Model	O (Openness) E (Extraversion) A (Agreeableness) N (Neuroticism) C (Conscientiousness)

The ITDBLA model performs feature extraction and model training for each user, generating a total of 1000 user behavior files and 45 role behavior files. The behavior feature is to transform each instance into a fixed-length vector. Behavior sequence is a feature frequently used in user behavior anomaly detection, representing users' behavioral habits. Role features and psychological data are related to user attributes, which are missing dimensions in other user behavior anomaly detection methods.

- Behavior Feature. Behavior feature is counting features used to represent users' daily activities in each period. For example, the user's first login time and login duration can be obtained from the login.csv, and the http.csv can obtain the number of network activities. Experts define the features according to the specificities of each activity, and these features indicate the user's daily behavior.
- Sequence Feature. The sequence feature is the user's behavior sequence over some time. First, the user's activities are extracted from different log files, and then these multi-source heterogeneous log lines are merged and sorted by time. For example, a user first logs on to a computer then browses two webs, sends an email, uses a removable drive, and finally logs off. Its behavior sequence is {login, web, web, email, drive connection, drive disconnection, logout}. Activities performed by different users over a while may have different execution sequences even if the activities are the same. Therefore, the behavior sequence also manifests the user's behavior habits.
- Role Feature. The role feature is based on the count feature of the user's job role. It is the common feature of the user based on the role. Employees in the same job role have analogous jobs, and their behavior features are also very analogous. In this article, the average value of employees' behavior in the same role is defined as the role feature.
- Psychological Data. To some extent, psychological data can identify abnormal behaviors. Psychological data reflects the employee's

personality. For example, an employee with a personality of N is more prone to impulsive threats than a C personality.

3.2 User and Entity Behavior Analysis Module

The ITDBLA model uses three deep learning algorithms to detect suspicious behavior in multi-source heterogeneous log lines. Generally speaking, when the model detects apparent changes in behavior, such as changes in access to sensitive files and the frequency of u-disk usage, it indicates that the user may have performed threatening operations of data leakage. In order to learn a normal behavior model of users and detect abnormal behaviors, the ITDBLA model utilizes historical data to predict the next state, and detects abnormal behaviors by calculating the error between the prediction and the actual behaviors. The ITDBLA model uses a new combination of attention mechanisms, long and short-term memory networks, and multilayer perceptron for insider threat detection. These algorithms have achieved remarkable results in natural language processing or computer vision.

3.2.1 Behavior and role feature Model-LSTM

Long Short-Term Memory (LSTM) network is a supervised method proposed by the literature [28], which is an improvement of recurrent neural network. Different from the single tanh recurrent structure, LSTM is a unique network structure with the three "gates" structure. LSTM can effectively solve the gradient explosion or disappearance problem of simple recurrent neural networks. Since the situations of working days and non-working days are fundamentally different, only working days are kept in this paper. You can also train a second model to simulate user behavior on non-working days if necessary. Due to experimental limitations, this article does not simulate a model of non-working day behavior. Specifically, the task of the LSTM is to predict the features of the fifth day by the behavior features of the first four days. Fig. 2 shows the cyclic unit structure of the LSTM network. The calculation process is as follows. First, three gates are calculated using the external state h_{t-1} of the last moment and the input x_t of the current moment, as well as the candidate states. Then, the forgetting gate f_t and the input gate i_t are combined to update the memory cell c_t. Finally, combined with the output gate o_t to transfer the information of the internal state to the external state h_t. The three gates are shown in equations (1) to (3).

$$\mathrm{i}_t = \sigma\left(W_i x_i + U_i h_{t-1} + b_i\right) \tag{1}$$

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{2}$$

$$\mathrm{o}_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{3}$$

The output interval of the logistic function is (0,1), x_t is the input at the

current moment, and h_{t-1} is the external state at the last moment.

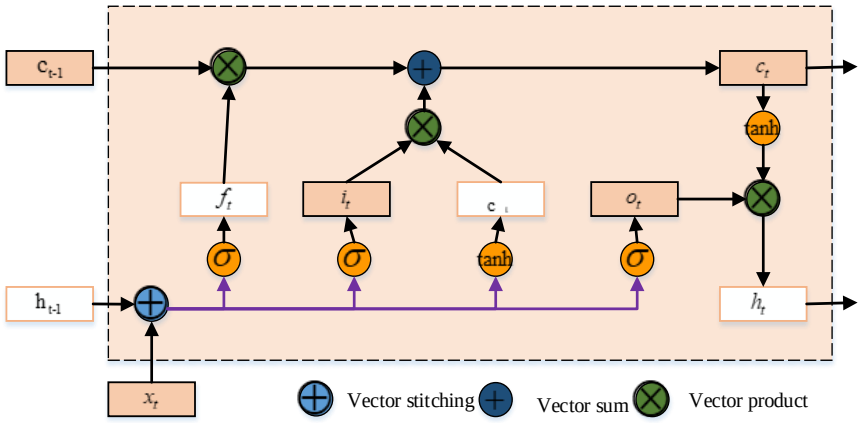

Fig. 2 The cyclic unit structure of the LSTM network

Through the LSTM cycle unit, the entire network can establish a longer-distance timing dependence relationship, as shown in formulas (4) ~ (6).

$$\begin{bmatrix} \tilde{c}_t \\ o_t \\ i_t \\ f_t \end{bmatrix} = \begin{bmatrix} \tanh \\ \sigma \\ \sigma \\ \sigma \end{bmatrix} \left(W \begin{bmatrix} x_t \\ h_{t-1} \end{bmatrix} + b \right) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{5}$$

$$h_t = o_t \odot tan\, h(c_t) \tag{6}$$

Where $W \in R^{4d \times (d+e)}$ and $b \in R^{4d}$ are the network parameters, $x_t \in R^{\ell}$ is the input at the current moment.

3.2.2 Behavior sequence Model-Attention

The attention mechanism was first applied to the image field. The Google Mind team proposed using the RNN model's attention mechanism for image classification and realized good results in 2014. The core of the attention mechanism is to assign weights, which is an idea and does not depend on any framework. The fully connected network incorporating the attention mechanism is trained to learn the user's normal behavior sequence and predict the following state's behavior sequence based on the historical data. The fully connected network model with the attention mechanism is shown in Fig. 3. First, the similarity and correlation between Query and Key are calculated by dot product, and then the softmax function is introduced to transform the first stage score numerically and normalize it so that the sum of weights of all elements is 1. Finally, the weight coefficients corresponding to α and v are calculated and then

weighted to obtain the Attention value. In the experiment of this article, the behavior sequence of N days is used to project the behavior sequence of the next state. The length of each user's behavior sequence is not same. For instance, the sequence {login, webpage, webpage, drive connection, drive disconnection, email, logout} will be coded and entered into the model.

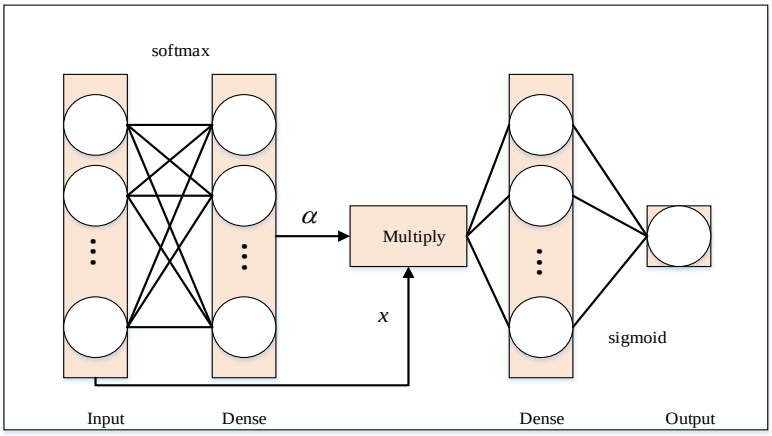

Fig. 3 Structure of fully connected neural network with attention mechanism

3.2.3 Comprehensive decision Model-MLP

The last part of the ITDBLA model is comprehensive decision-making. This article uses Multilayer Perceptron to perform comprehensive anomaly detection on the above-mentioned deep learning model results. MLP is a neural network composed of the input, hidden, and output layers. The input elements and weights product are fed to the summing node with neuron bias. The main advantage of MLP is that it can solve complex problems quickly and is often used to solve nonlinear problems. The weights and biases of each connection are determined by gradient descent. All parameters are first randomly initialized and then trained iteratively, continuously computing the gradient and updating the parameters until the error is sufficiently small. The MLP needs to be trained using historical data to learn the association between these features in the ITDBLA model. MLP judges whether deviant behavior is based on the deviation of behavior sequence, behavior feature, and role feature, alike to performing classification tasks.

4. Experiment and Result Analysis

This section describes the dataset and experimental environment used, and the hyperparameters are adjusted. To verify the model's validity, series of experiments are conducted on the CERT r4.2 dataset. Firstly, the LSTM and the fully connected network incorporating the attention mechanism are used to learn users' behavior patterns, and the weighted deviation degree (WDD) is used to

weigh the deviation between the actual and predicted data. Secondly, the MLP makes a comprehensive decision based on the WDD of the three features to achieve anomaly detection. Finally, to ensure the consistency of subsequent results comparison, the area under the ROC curve (Area Under Curve, AUC) is employed for uniform evaluation. To prove the threat detection capability of the ITDBLA model, this paper compares the insider threat detection models proposed by other scholars and discusses the experimental results.

4.1 Dataset Description

This study conducts experiments on the CERT insider threat dataset [29], which contains classification information such as user attribute. It is proposed by Carnegie Mellon University and widely used in researching, developing, and testing insider threat detection methods. It is collected from an actual enterprise and generated using thematic, behavioral, and psychometric models to simulate three types of attacks: system disruption, information theft, and insider fraud perpetrated by malicious insiders. The dataset has 1,000 users, including 70 insider attackers. The CERT dataset includes five different CSV files: device.csv, file.csv, login.csv, email.csv, and http.csv. The description of the files in the CERT dataset is shown in Table 3. In addition, it provides users' job information, psychological data, and monthly employee attendance, allowing this paper to analyze them using their psychological data and job roles.

Table 3

File description in CERT dataset

File Name	Description
LDAP	LDAP file describing the ontology (role, department, supervisor, etc.)
device.csv	Connect and disconnect removable devices (such as USB hard drives)
http.csv	User web browsing/uploading/downloading
file.csv	File access activities, including file open/copy/write/delete
logon.csv	User activity based on logging in and logging out of computing devices
email.csv	User mail log including mail receiving, sending and size, etc.
psychometric.csv	Provide personality and job satisfaction variables for each of the 1,000 users

4.2 Experimental Environment and Parameter Settings

This article uses Keras to build the LSTM-Attention network model. The computer configuration processor used for model training and testing is i5-6200U, CPU @ 2.30GHz 2.40 GHz, RAM 8GB. In the experiment, the user file under "ProductionLineWorker" is used for testing. This paper randomly shuffles the order of the training set for improving the model's robustness. In this experiment, the LSTM model consists of two LSTM layers, with 100 and 120 units, respectively. After each LSTM layer, there is a "tanh" activation layer, a 37-unit Dense layer, and a "relu" activation layer. LSTM network hyperparameter settings are described in Table 4. The Attention model is the attention mechanism built on a fully connected network. The attention layer comprises a Dense layer, a

Multiply operation, and a "softmax" activation layer. The hyperparameter settings are shown in Table 5.

Table 4

LSTM hyperparameter settings

Behavior feature model-LSTM	Hyperparameter Settings
Input	Dim = 148
Reshape	Dim = (4,37)
LSTM	Units = 100
Activation	Function = tanh
LSTM	Units = 120
Activation	Function = tanh
Dense	Dim = 37
Activation	Function = relu

Table 5

Attention hyperparameter settings

Sequence feature model-Attention	Hyperparameter Settings
Input	Dim = 132
Attention_vec: Dense	Dim = 132
Activation	Function = softmax
Multiply	Dim =[(None,132), (None,132)]
Dense	Dim = 33
Activation	Function = softmax

4.3 Result Analysis

In training the model, this experiment uses the features of the first four days to predict the features of the fifth day. It uses WDD to calculate the deviation between the real and predicted values, which is linearly weighted according to a weighted square error. The calculation formula of WDD is shown in (7). For the model to learn the normal behavior patterns of users, the data for training LSTM and attention mechanism are all benign samples.

$$WDD = \frac{1}{|V|}\sum_{\forall y \in V} \mathrm{w}\left(y - \hat{y}\right)^2 \tag{7}$$

Where V is the set of all real features, y is a feature in V, $\hat{y}$ is the predicted feature corresponding to y, and w is the weight value determined according to the feature y.

There is a severe imbalance between positive and negative categories in the CERT insider threat detection dataset. The test data's distribution of positive and negative samples may also change over time. However, ROC and AUC can eliminate the impact of imbalance between sample categories on the indicator results. The ROC curve is also called the receiver operating characteristic curve. Its ordinate is the True Positive Rate (TPR), and the abscissa is the False Positive Rate (FPR). Therefore, considering the accuracy and recall rate, this article uses

AUC as the evaluation metrics. The larger the AUC, the better the model effect. The following will analyze the experimental results of LSTM, attention mechanism, and MLP on CERT r4.2.

- Result analysis of user's behavior features

The features of user behavior are diverse and interrelated. This article uses every five days to train the LSTM model. The behavior features of the first four days is employed to predict the behavior features of the fifth day, and the deviation between the real fifth day data and the data predicted by the LSTM model is calculated and optimized in the training phase. It is important to note that all users under the identical role use the identical features and LSTM model, but each user keeps its own parameters. As shown in Fig. 4, the deviation of the first 200 days is basically between 0 and 2, and the deviation after 200 days has increased significantly, which is obviously different from the previous deviation. From the figure we can see that there are many abnormal deviations after 200 days, which correspond to the actual situation, which illustrates that the LSTM network has a powerful ability to learn user behavior patterns.

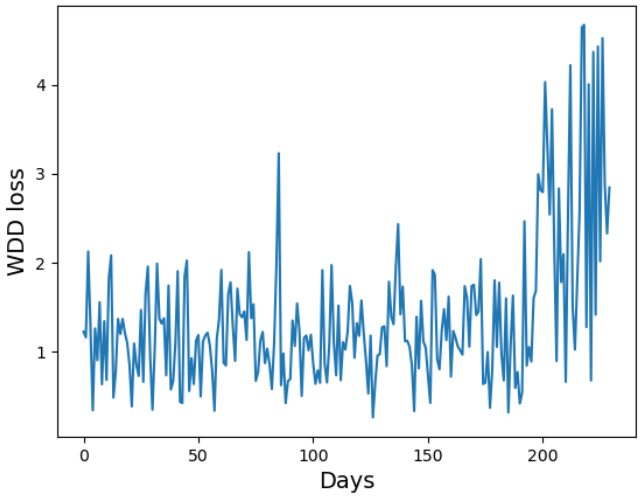

Fig. 4 WDD of behavior features

- Result analysis of user behavior sequence

As mentioned in section 2.3.2, the fully connected network with Attention is used to train and learn the normal mode of user behavior sequence. Due to time and data limitations, the Attention network is trained using a behavior sequence with a time unit of 5 days. In the model, the first four days are employed as known data to predict the behavior sequence of the fifth day and then compared with the authentic fifth day data, and the deviation calculation and optimization of the Attention network model are performed. As shown in Fig. 5, it can be found that the test data of the previous 40 days and the training data of the previous 160 days have a similar distribution, and the loss range is 0-4, indicating that the

Attention network model has learned the behavior sequence of users well. The figure shows that the user has some abnormal behaviors around 200 days, which makes the abnormal deviation between the prediction and the actual sequence to become larger. In the light of the above-obtained results and analysis, it is evident that the attention mechanism can be used to learn the behavior sequence's normal mode, and it is meaningful in practical applications. However, it is necessary to pay attention to the situation of overfitting.

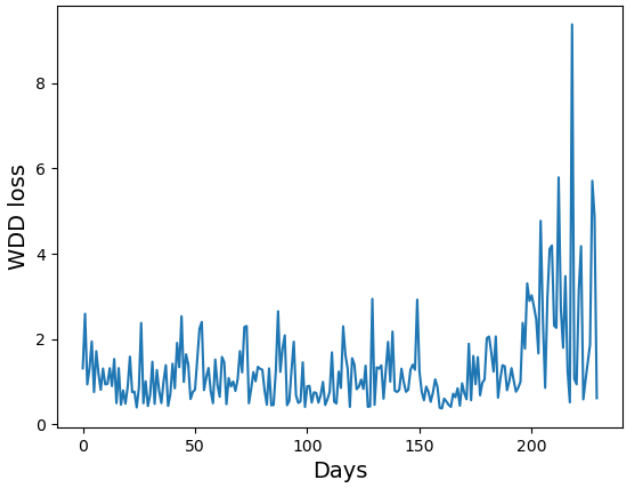

Fig. 5 WDD of behavior sequence

- Analysis of the results of MLP's comprehensive decision-making

The ITDBLA model obtains the deviation between actual and predicted from three perspectives: behavior sequences, behavior features, and role features, represented in Fig. 6. As shown in Fig. 6, it can be found that the normal point and the abnormal point are separable.

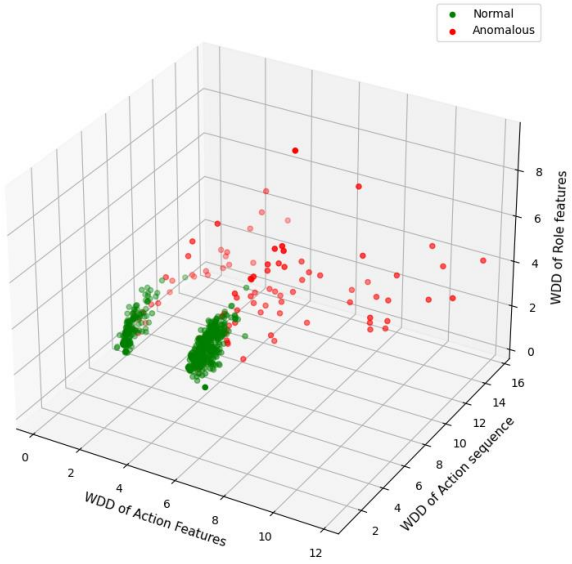

Fig. 6 The deviation distribution of normal and abnormal data

Although there are certain false positives and false negatives, these three features mirror the user's abnormal behavior to a large extent. In order to issue an alarm more accurately when abnormal behavior occurs, the MLP is used to learn the relationship of the three deviations and determine whether a user has abnormal behavior on a particular day in this experiment.

This paper compares the AUC scores with several other insider threat detection models that have worked on the CERT r4.2 dataset in recent years. The AUC scores of the ITDBLA model are all significantly better than the other models, as shown in Fig. 7. Table 6 and Fig. 8 show the comparison of experimental results between the ITDBLA and SUS in the literature [30], which can show that ITDBLA has achieved good results in accuracy and AUC scores.

Table 6

Comparison of experimental results of ITDBLA and SUS

Model	Accuracy	AUC	TPR	FPR
ITDBLA	0.98	0.964	0.98706	0.28125
SUS	0.94	0.934	0.96532	0.2222

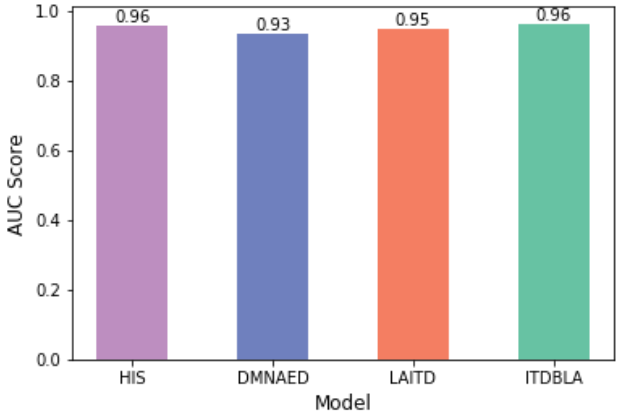

Fig. 7 AUC scores of ITDBLA and other models

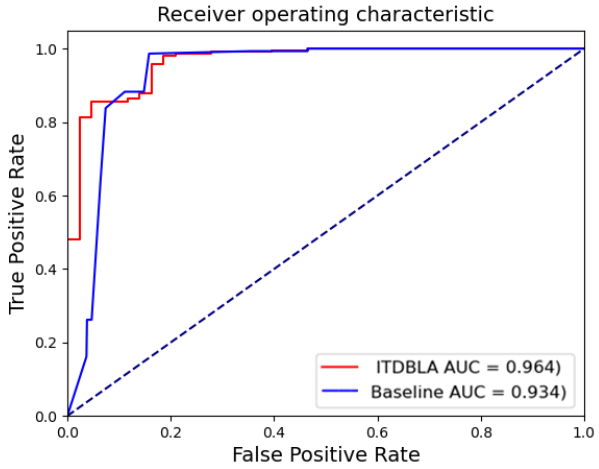

Fig. 8 ROC curve of ITDBLA and SUS model

The disadvantage is that the false alarm rate has increased, and the focus of future work is to reduce the false alarm rate and increase the interpretability. The ROC curve trajectory of SUS and ITDBLA is similar, but the area under the ROC curve of ITDBLA achieves 0.964, significantly better than SUS. It is proof of the effectiveness of the ITDBLA model, proving that it has strong anomaly detection capabilities.

5. Conclusions and Future Works

In this research, we propose an insider threat detection model ITDBLA for the power system based on LSTM-Attention user and entity behavior analysis techniques, which fully uses the advantages of LSTM and attention mechanism that can handle long sequences well. It models the behavior of insiders from multiple perspectives such as behavior features, behavior sequences, role features,

and psychological data in a comprehensive manner. The MLP utilizes the deviation of multiple features to perform comprehensive decision-making, thus realizing the abnormal behavior detection of insiders in the power system. The ITDBLA model proposed in this article and the model proposed in the latest literature are compared and analyzed. The abnormal detection ability of the ITDBLA model is better than that of a single machine learning model and most fusion models. With the implementation of the network security law and the development trend of information security enhancement, the behavioral security audit of power information systems is becoming more critical. In the next step, the problem of unbalanced sample categories and very few insiders in the insider threat dataset of the power system will be considered to classify the emerging unknown insider threats by using different small sample learning algorithms.

REFERENCES

[1] *L Fan, J Du, Y P Guo, et al.* A Security Defense Scheme for Encryption and Network Isolation Gateway in Power System. 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC). IEEE, 2019.

[2] *C Lai, P Cordeiro, A Hasandka, et al.* Cryptography Considerations for Distributed Energy Resource Systems. 2019 IEEE Power and Energy Conference at Illinois (PECI). IEEE, 2019.

[3] *V K Singh, H Ebrahem, M Govindarasu.* Security Evaluation of Two Intrusion Detection Systems in Smart Grid SCADA Environment. North American Power Symposium (NAPS) 2018. 2018.

[4] *C Wu, J L Shuai, T Long, et al.* Research on Detection Method of User Abnormal Operation Based on Linux Shell Commands. Netinfo Security, 2021, 21(5): 31-38.

[5] *X Y Ye, S S Hong, M M Han.* Feature Engineering Method Using Double-Layer Hidden Markov Model for Insider Threat Detection. International Journal of Fuzzy Logic and Intelligent Systems, 2020, 20(1): 17-25.

[6] *Z H Tian, C C Luo, H Lu, et al.* User and Entity Behavior Analysis under Urban Big Data. ACM Transactions on Data Science, 2020, 1(3): 1-19.

[7] *S Gorka, C Jonathan, M Neil, T Henrique.* 2019 Market Guide for User and Entity Behavior Analytics. Gartner, 2019.

[8] *M Du, F F Li, G N Zheng, et al.* DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning. ACM SIGSAC. The 2017 ACM SIGSAC Conference on Computer and Communications Security, Oct 30, 2017, New York, USA. New York: ACM SIGSAC, 2017: 1285-1298.

[9] *Y B Guo, C H Liu, J Kong, et al.* Study on User Behavior Profiling in Insider Threat Detection. Journal on Communications, 2018, 39(12): 141-150.

[10] *Pauls, S Mishra.* LAC: LSTM AUTOENCODER with Community for Insider Threat Detection. ACM. The 4th International Conference on Big Data Research (ICBDR'20), Nov 27, 2020, Tokyo, Japan, New York, NY: ACM, 2020: 71-77.

[11] *A Al-Dhamari, R Sudirman, N H Mahmood.* Transfer deep learning along with binary support vector machine for abnormal behavior detection. IEEE Access, 2020, 8: 61085-61095.

[12] *M U Togbe, M Barry, A Boly, et al.* Anomaly detection for data streams based on isolation forest using scikit-multiflow. International Conference on Computational Science and Its Applications. Springer, Cham, 2020: 15-30.

[13] *H Y Bao, R X Lu, et al.* BLITHE: Behavior Rule-Based Insider Threat Detection for Smart Grid. IEEE Internet of Things Journal, 2015, 3(2): 190-205.

[14] *B Li, R Lu, G Xiao, et al.* Towards Insider Threats Detection in Smart Grid Communication Systems. IET Communications, 2019, 13(12): 1728-1736.

[15] *P Legg, O Buckley, et al.* Automated Insider Threat Detection System Using User and Role-Based Profile Assessment. IEEE Systems Journal, 2017, 11(02): 503-512.
[16] *M Jagielski, A Oprea, B Biggio, et al.* Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning. IEEE. 2018 IEEE Symposium on Security and Privacy (SP), May 20, 2018, San Jose, CA: IEEE, 2018: 204-221.
[17] *F Liu, E T Lim, H X Li, et al.* Disentangling Utilitarian and Hedonic Consumption Behavior in Online Shopping: An Expectation Disconfirmation Perspective. Information & Management, 2020, 57(3): 100-123.
[18] *Y Wen, W P Wang, D Meng.* Mining User Cross-Domain Behavior Patterns for Insider Threat Detection. Chinese Journal of Computers, 2016, 39(08): 1555-1569.
[19] *H Yu, T T Zhang, J X Chen, et al.* Web Items Recommendation Based on Multi-View Clustering. IEEE.2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Jul 23, 2018, Tokyo, Japan: IEEE, 2018: 153-158.
[20] *Z Li, L P Song.* Research on Internal Threat Detection Based on User Window Behavior. Computer Engineering, 2020, 46(4): 135-142,150.
[21] *M N Al-Mhiqani, R Ahmad, Z Z Abidin, et al.* A new intelligent multilayer framework for insider threat detection. Computers & Electrical Engineering, 2021: 107597.
[22] *Shashankam, M Y Shen, J S Wang.* User and Entity Behavior Analytics for Enterprise Security. IEEE. 2016 IEEE International Conference on Big Data (Big Data), Dec 5-8, 2016, Washington D.C., USA: IEEE, 2016:1867-1874.
[23] *D Sun, M Liu, M Li, et al.* DeepMIT: A Novel Malicious Insider Threat Detection Framework based on Recurrent Neural Network. 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2021: 335-341.
[24] *M N Al-Mhiqani, R Ahmad, Z Z Abidin, et al.* New insider threat detection method based on recurrent neural networks. Indones. J. Electr. Eng. Comput. Sci, 2020, 17(3): 1474-1479.
[25] *Y Liu, S Garg, J Nie, et al.* Deep anomaly detection for time-series data in industrial iot: a communication-efficient on-device federated learning approach. IEEE Internet of Things Journal, 2020, 8(8): 6348-6358.
[26] *S B Teodora, Caglayanb, and A Haytham.* DeepAD: A Generic Framework Based on Deep Learning for Time Series Anomaly Detection. Cham: Springer, 2018: 577–588.
[27] *F F Yuan, Y N Cao, Y M Shang, et al. I*nsider Threat Detection with Deep Neural Network. Springer. International Conferenceon Computational Science (ICCS), June 11, 2018, Wuxi, China, Cham: Springer, 2018:43-54.
[28] *S Hochreiter and J Schmidhuber.* Long Short-Term Memory. Neural Computation, 1997, 9(8): 1735–1780.
[29] *J Glasser, B Lindauer.* Bridging the Gap: A Pragmatic Approach to Generating Insider Threat Data. IEEE. 2013 IEEE Security and Privacy Workshops.May23-24, 2013, San Francisco, CA, USA: IEEE, 2013: 98-104.
[30] *D C Le, A N Zincir-Heywood.* Evaluating Insider Threat Detection Workflow using Supervised and Unsupervised Learning. IEEE. 2018 IEEE Security and Privacy Workshops, May 24, 2018, San Francisco, CA. USA: IEEE, 2018:270-275.