Hindawi
Mobile Information Systems
Volume 2022, Article ID 6576023, 11 pages
https://doi.org/10.1155/2022/6576023

Research Article

Computer Network Intrusion Anomaly Detection with Recurrent Neural Network

Zeyuan Fu

School of Cyber Science and Engineering, Wuhan University, Wuhan, China

Correspondence should be addressed to Zeyuan Fu; 2019302180015@whu.edu.cn

Received 30 December 2021; Accepted 31 January 2022; Published 7 March 2022

Academic Editor: Hasan Ali Khattak

Network intrusion anomaly detection technique has been widely employed in computer network environments as a highly effective security prevention method. As network technology and network applications have advanced at a rapid pace, so too has network data traffic, resulting in an increase in virus and attack kinds. In the face of large-scale traffic and characteristic information, traditional intrusion detection will have problems such as low detection accuracy, high false negatives, and reliance on dimensionality reduction algorithms. Therefore, it is particularly important to establish a fast and efficient network intrusion anomaly detection method to deal with the current complex network environment. This work designs a computer network intrusion detection model with a recurrent neural network in order to explore a new intrusion detection method. The main purpose of this article include the following: (1) design a network security emergency response system architecture with the recurrent neural network model. This system consists of a management center module, a knowledge database module, a data acquisition module, a risk detection tool module, a risk analysis and processing module, a data protection module, and a remote connection auxiliary module. The modules cooperate with each other to complete system functions. (2) Aiming at the risk analysis and processing module, a network intrusion detection model combining bidirectional long short-term memory (BiLSTM) and deep neural network (DNN) is designed. In view of the lack of consideration of the before-and-after relevance of intrusion data features and the multifeature problem in existing models, the use of BiLSTM to extract the relevance between features and the use of DNN to extract deeper features are proposed. Aiming at the problem that the model lacks consideration of the importance of features, it is proposed to embed an attention mechanism into the network to increase consideration for the importance of features. (3) Massive experiments have verified the reliability and effectiveness of the method proposed in this work.

1. Introduction

In today's society, the Internet has affected the entire social process. The application for the Internet is becoming diversified, and the scale of netizens continues to show a trend of sustained and rapid development [1, 2]. At this stage, China's cyber security situation is becoming more severe and complex, and cyber security incidents are becoming more complex. And with the rapid development of the mobile Internet, cloud computing has been further implemented, and NFV (Network Function Virtualization) technology has continued to mature and be commercially available [3]. Under a huge IT architecture, security risks will further expand. For example, the probability of occurrence of security incidents, such as intrusion attacks, web page tampering, DDoS, and the scope of protection, will increase accordingly, and new security risks will also appear [4]. This also puts forward new requirements for safety protection technology, safety risk management, and emergency response methods [5].

Network and information security always affect people's lives and national security. With the increase of attackers, intrusion technology has become more and more mature, with complex and changeable attack tools and techniques [6]. Only using firewall technology can no longer meet the needs of important departments for security technology. Therefore, the network defense system needs to adopt some deeper and more diverse methods [7, 8]. The increasingly complex network environment requires that defense equipment must be continuously upgraded and leaked.

These problems make the workload of network administrators more and more, and inconspicuous negligence may cause major security risks [9]. As an effective supplement to the firewall, the intrusion detection system can quickly and effectively detect intrusions, ensure the normal daily operations of the network, and ensure its security. Therefore, intrusion detection system has become a research hotspot and has received more and more attention from people. No matter what environment it is in, it plays a very important role [10].

Security vulnerabilities are will be the main cause of network vulnerabilities. In addition, there are political, economic, and cultural interest trends that have led to more and more serious network security problems [11, 12]. Among traditional network security technologies, most of them are passive defense systems. The main purpose is to protect the data from being infringed during the transmission process. When there are external malicious and unauthorized requests, the defense system will be activated to protect the transmitted data [13]. To make up for the lack of defense problems, intrusion detection technology is slowly entering people's lives. Therefore, a popular network system must have not only firewall defense capabilities but also an intrusion detection system that can control the network in real time for security, attack, and antiattack. Security protection from both the internal and external aspects of the network can better protect the network system and further facilitate people's daily life. In recent years, intrusion detection systems have received close attention from researchers in the security industry and industry professionals [14, 15].

Nowadays, in the process of emergency response to security incidents, all localities have gradually established a one-click emergency response platform to realize emergency response to security incidents and improve overall security emergency response capabilities. The existing technology still has certain shortcomings; that is, the network security incidents are mainly discussed in the analysis link of the network security incident, and there is not much description of the emergency response link. In actual operations, the emergency handling of network security incidents is often handled manually, which is greatly affected by personal factors, such as the unstable efficiency of manual handling and the fluctuating efficiency. The manual treatment process is not uniform, and it takes a long time, the treatment result is inaccurate, and it is not convenient to manage the treatment result.

This work comprehensively considers these problems and proposes a computer network intrusion anomaly detection strategy with the recurrent neural network. The contributions of this article are as follows: (1) design a network security emergency response system architecture with the recurrent neural network model. This system includes seven different modules, each of which cooperates with each other to monitor network intrusions in real time and take corresponding protective measures when an intrusion occurs. (2) Aiming at the risk analysis and processing module in the system, a computer network intrusion anomaly detection method combining BiLSTM and DNN is designed, and the attention mechanism is embedded to complete efficient intrusion detection.

2. Related Work

Literature [16] first proposed the introduction of machine learning in the process of intrusion detection. Its purpose was to monitor the abnormal situation of the network in real time so as to judge the attack behavior on the network and further identify the judged attack behavior. Literature [17] proposed two algorithms, KNN and SVM, which were the most frequently used data mining methods in intrusion detection. Literature [18] used support vector machines for intrusion detection systems. Literature [19] proposed an intrusion detection method based on a one-dimensional CNN, which could extract the features for original data. Literature [20] analyzed the feasibility of RNN for intrusion detection and detected the behavior of network traffic by modeling network traffic as a sequence of states. Literature [21] verified the performance of LSTM in intrusion traffic classification, and the results showed that LSTM could learn attacks hidden in training data. Literature [22] proposed a feature algorithm, that is, a conditional random feature selection algorithm and a feature selection algorithm with linear correlation coefficients, which used the existing CNN to select the most contributing feature. Literature [23] proposed a hierarchical deep learning system based on big data. This system used behavioral characteristics and content characteristics to understand network traffic features and information. Each model strived to learn a unique data distribution to further improve intrusion detection based on deep learning.

Literature [24] designed a network intrusion detection strategy with autoencoder network (AE) and LSTM. By superimposing multiple self-encoding networks and mapping information to low-dimensional space, a self-encoding network model was constructed. Then, the optimized LSTM model was used to extract features, train data, and predict the type of intrusion detection. Literature [25] conducted related research on deep learning models, focusing on using unsupervised deep learning methods and semisupervised learning methods to detect abnormal network traffic from stream-based data. More specifically, the method of the autoencoder and variational autoencoder was used to learn traffic characteristics to identify unknown attacks so as to improve intrusion detection. Literature [12, 26, 27] also applied convolutional neural networks in intrusion detection, using CNN to classify network traffic characteristics, and compared with traditional algorithms, this algorithm had better results. This did not pay attention to the issue of imbalance in the network traffic dataset, which made the classification result of this algorithm for a small number of attack types relatively poor. To solve the issue of imbalance in network traffic characteristics, literature [28] proposed an autoencoder (AE) and generative adversarial network (GAN) model based on the unsupervised learning model in the deep learning process. It aimed to solve the problem of data imbalance and high-performance intrusion detection. To improve the performance, a GAN under the condition of

autoencoder (AE-CGAN) is proposed. The model was based on the GAN model to sample rare data, and after the features of the data were processed to a lower level using the autoencoder model, the performance degradation caused by the imbalance of data was solved. Literature [29] set the weight coefficient of the cost function according to the number of each class for the above-mentioned imbalance problem. It effectively solved the problem of data imbalance and proposed a new convolutional neural network model. Literature [30] proposed a new unsupervised dimensionality reduction method to detect attacks. It used t-SNE combined with a hierarchical neural network to map network data. Literature [31] proposed a hybrid method of multiobjective genetic algorithm and neural network to establish a set of integrated solutions for effective network intrusion detection. This method had achieved good results.

3. Method

3.1. Network Security Emergency Response System. In actual operations, the emergency response links of network security incidents are often handled manually, which is greatly affected by personal factors. The manual treatment process is not uniform, and it takes a long time, the treatment result is inaccurate, and it is not convenient to manage the treatment result. In response to this problem, this work first designed a network security emergency response system with a recurrent neural network model to overcome the above technical problems in the existing related technologies.

The system consists of a management center module, a knowledge database module, a data acquisition module, a risk detection tool module, a risk analysis and processing module, a data protection module a,nd a remote connection auxiliary module. The system architecture is illustrated in Figure 1.

The management center module is used to perform system management and complete the connection and call of each module. The knowledge database module is utilized to store risk data and expert knowledge data and update it regularly. The risk data stored in the knowledge database module includes web page tampering, domain name hijacking, intrusion attacks, viruses, Trojan horses, and malicious codes. Expert knowledge is a solution corresponding to the risk data. The data acquisition module is used to monitor and collect network data in real time. The collected network data includes website source code, operating system logs, website web page access logs, and middleware log information. The risk detection tool module is used for security detection of network data and detection and classification of abnormal data. It includes system vulnerability verification tools, website vulnerability verification tools, database vulnerability verification tools, virus detection tools, Trojan horse detection tools, and malicious code detection tools.

The risk analysis and processing module is used to analyze, match, and process abnormal data. It includes a recurrent neural network model building module, a recurrent neural network model predictive analysis module, a risk information extraction module, a risk processing module, and a data transmission module. The recurrent neural network model building module is utilized to realize the construction of the recurrent neural network model. Recurrent neural network model evaluation module is used to analyze and predict abnormal data. The risk information extraction module is used to process abnormal data and obtain clue tree and attacker information. The risk processing module is used to match risks and promptly repair them. The data transmission module is used to transmit the expert knowledge data, the clue tree, and the attacker information to the remote connection auxiliary module.

The risk analysis and processing module is used to analyze, match, and process abnormal data. It includes the following steps: the recurrent neural network model building module uses different types of risk data stored in the knowledge database module to construct and train the recurrent neural network model, respectively. The recurrent neural network model predictive analysis module receives the abnormal data detected by the risk detection tool module, extracts query parameters, and inputs the corresponding recurrent neural network model for analysis and prediction according to its type. The risk processing module matches the prediction results, and if there is corresponding expert knowledge in the knowledge database module, automatic system repair and manual rectification are performed. If it does not exist, upload the prediction result to the data delivery module. The risk information extraction module extracts the clue tree and attacker information in the abnormal data. The data transmission module sends the prediction result, the clue tree, and the attacker information to the remote connection auxiliary module.

In the designed network security emergency response system, the recurrent neural network in the risk analysis and processing module is the core of the system. In the following chapters, we will introduce in detail the recurrent neural network designed for computer network intrusion anomaly detection.

3.2. Intrusion Detection with Deep Neural Network. A deep neural network (DNN) was proposed by Professor Hinton in 2006. It is a dense network composed of interconnected neurons and has multiple hidden layers. It is also a feedforward neural network. Therefore, it can also be divided into input layer, output layer, and hidden layer. The first two layers are a single-layer structure, and the number of nodes is determined by the input matrix and the output matrix, respectively. The layers between the two are collectively called the hidden layer. The overall length of the network determines the depth of the model. From a purely structural analysis point of view, it is the same as a multilayer perceptron. The structure of DNN is illustrated in Figure 2.

The key to deep neural network learning lies in parameter learning and adjustment, which involves the calculation process of forward propagation and backward propagation. In the process of forward propagation, the output of k -th layer is as follows:

$$y_k = f\left(W_k y_{k-1} + b_k\right), \tag{1}$$

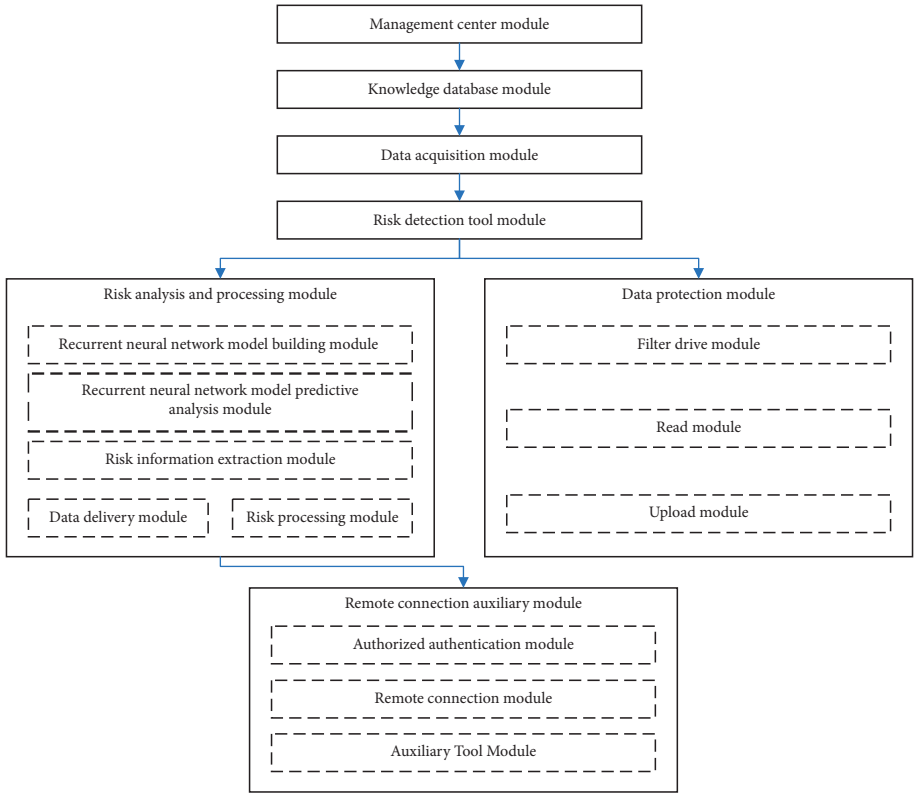

Figure 1: Network security emergency response system with the recurrent neural network.

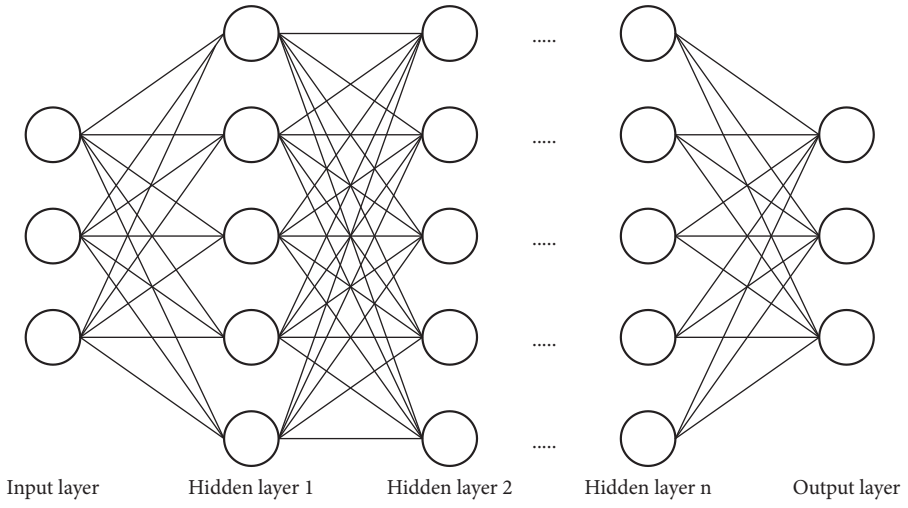

Figure 2: The structure of DNN.

where y_k represents the output of the k -th layer, $f(\cdot)$ represents the activation function, W_k represents the connection weight, and b_k represents the bias of the k -th layer. It can be seen that with certain input parameters, the final output result is related to weight and bias. Therefore, the backward propagation is the solution and update for weights and biases.

Although the DNN network has a strong nonlinear expression ability, it also faces some problems. The activation function uses the sigmoid function, the convergence of the model is slow, the problem of gradient disappearance is prone to occur. The deepening of the model depth increases the number of network parameters and matrix calculations, and it is easy to overfit. Therefore, in actual use, DNN often uses dropout regularization to alleviate overfitting. Dropout refers to that when the network is training, some of the nodes are randomly selected for inactivation operation and then restored after the parameters are updated.

The basic framework of DNN-based intrusion detection is shown in Figure 3. The basic idea is to use the good

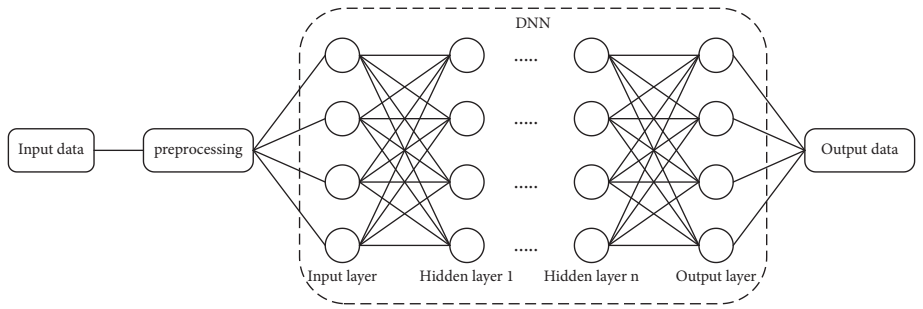

Figure 3: DNN-based intrusion detection framework.

nonlinear expression ability of DNN to build a detection model by learning the characteristics of intrusion data and extracting deep-level features. The nonlinear expression ability of DNN is related to its hidden layer. The number of nodes can refer to the empirical formula $s = (m + n)/2$. m and n are the number of input and output nodes, adjusted by actual conditions.

The basic workflow of this method is as follows: (1) preprocessing the original dataset into standard format data; (2) using the training data to train the DNN model; (3) using the trained model to predict and classify unknown data.

3.3. Intrusion Detection with Recurrent Neural Network. In traditional neural networks, neurons in the same layer do not transmit information to each other, which cannot meet the needs of many contextual tasks. Due to the particularity of its network, RNN has unique advantages for processing time series tasks and is usually regarded as a neural network that works according to time sequence. It simulates the order in which people read the article and achieves a certain memory ability by retaining certain information on the processed content so that the subsequent content can be better processed. The structure diagram of the recurrent neural network is illustrated in Figure 4.

In Figure 4, x represents the input, h represents the hidden state, y represents the prediction result, and W, U, and V represent the weight matrix. Since RNN is a time series model, it is necessary to analyze the behavior and state of the network in terms of time. At time t, the neuron state h_t is determined by the network's current input x_t and the network state h_{t-1} at the previous time. At this time, the neuron state h_t can be calculated as follows:

$$h_t = f\left(Ux_{t-1} + Wh_{t-1} + b_h\right), \tag{2}$$

where f is activation function and b_h is a bias term.

The neuron state h_t at time t is used as the output, at the same time as the input of the network state at the next time $t + 1$. However, h_t cannot be directly output as a result. It needs to be multiplied by a coefficient V, then the offset ob is added, and normalization is required. The mathematical calculation formula is as follows:

$$y_t = act\left(Vh_t + b_y\right), \tag{3}$$

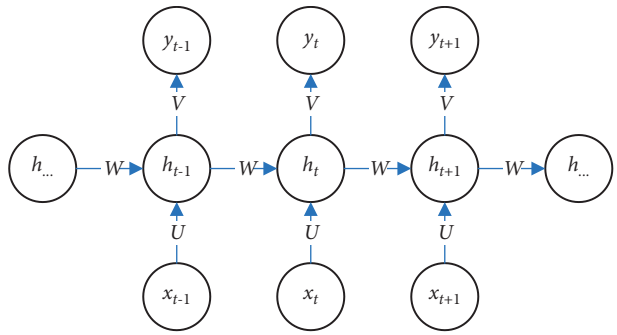

Figure 4: RNN structure diagram.

where act is activation function and b_y is a bias term.

The parameters of the model are shared by the RNN at different times, which reduces parameters that need to be learned but also leads to very unstable model parameters when updating. At the same time, RNN theoretically has the ability to deal with long-term dependence issues, but there are problems with gradient explosion or gradient disappearance, and due to gradient disappearance, only short-term memory is available.

The researchers studied the difficult-to-solve gradient explosion or gradient disappearance in traditional recurrent neural networks and long-term dependence problems. They have proposed a variety of solutions, among which the gated-based recurrent neural network is the most representative, and the LSTM network is one of the commonly used ones. LSTM is a kind of RNN with a special structure. It determines the retention or forgetting of input information by adding gating units to the model, thus effectively solving the problem of long sequence dependence. In layman's terms, LSTM has better performance than ordinary RNN when dealing with longer sequence tasks.

LSTM model is composed of a cell unit and three gating units. The cell unit is used to store necessary content information. The Forgotten Gate screens the content passed down from the previous moment, retains important information, and forgets useless information. The input gate selectively retains the network input at the current moment, which can effectively eliminate useless information. The output gate selectively outputs the current state of the cell. The calculation formula is as follows:

$$
\begin{aligned}
f_t &= \sigma\left(W_f\left[h_{t-1}; x_t\right] + b_f\right) \\
i_t &= \sigma\left(W_i\left[h_{t-1}; x_t\right] + b_i\right) \\
o_t &= \sigma\left(W_o\left[h_{t-1}; x_t\right] + b_o\right) \\
h_i &= o_t * \tanh\left(c_t\right) \\
c_t &= f_t * c_{t-1} + i_t * \tanh\left(W_s\left[h_{t-1}; x_t\right] + b_s\right),
\end{aligned} \tag{4}
$$

where W_f, W_i, W_o, and W_s, respectively, represent the weight matrix of each gating unit and cell unit. b_f, b_i, b_o, b_s represent the bias terms of each gating unit and cell unit, respectively. σ represents the sigmoid function.

The output of each gating unit at time t is determined by the input x_t of the entire model at that time and the implicit output h_{t-1} at the previous time, but their respective weights and biases are different. The weights and biases are learned during training.

The basic framework of intrusion detection with the recurrent neural network is shown in Figure 5. The basic idea is to utilize the ability of the recurrent neural network to process time series, analyze the intrusion data by extracting the correlation information between features, and build a detection model.

The basic workflow is as follows: (1) using preprocessing to convert the original dataset into standard format data; (2) building a recurrent neural network model and training it; (3) using the trained model to predict and classify unknown data.

3.4. Intrusion Detection with BiLSTM-DNN. Traditional intrusion detection methods are often purely analyzed from a certain aspect, it is difficult to carry out a more comprehensive description of the intrusion behavior. Intrusion detection usually needs to extract in-depth features from the numerous features of massive intrusion data to realize the detection of intrusion behavior, which is an extremely complicated process. Although deep neural networks have a good ability to express complex problems, there are still two problems in the application of intrusion detection.

The model itself has problems of overfitting and the disappearance of gradients. In the training process, ordinary neural network models are prone to overfitting problems; that is, although the model has a very good performance on the training set, it performs very poorly on the test set. Overfitting usually means that the model is learning too thoroughly. There are many reasons for overfitting, such as excessive noise data interference and insufficient training data. The disappearance of the gradient means that the weight of the node in the neural network is no longer updated, which will make the model difficult to train. The reason for the disappearance of the gradient is the inappropriate activation function used in the deep network. Since the neural network uses the BP algorithm to update the weights, if the selected activation function derivative is less than 1, then the more the previous layer is, the slower the weight update will be. In the end, the weights of the foremost layer are hardly updated.

There are a lack of consideration of the relevance of feature attributes and lack of consideration of the importance of features. Although both traditional RNN and gating-based LSTM can handle time series problems, they are both one-way and consider the impact of historical information on the current moment. In many tasks, the output at the current moment is not only related to the previous state but may also be related to the future state. Through analysis and research on the background, current situation, and methods of intrusion detection, it is found that the intrusion detection models proposed by current researchers rarely take the importance of features into consideration. However, different attack characteristics have different degrees of importance to the model when identifying the type of attack.

Aiming at several problems in intrusion detection, this work designed an intrusion detection method combining BiLSTM and DNN and named it BiLSTM-DNN. Moreover, in this method, Batch Normalization (BN) and attention mechanism are introduced. The structure of BiLSTM-DNN is illustrated in Figure 6.

The model uses BiLSTM to extract the correlation between features, uses DNN to extract deeper features, and uses the ReLU function and batch normalization to alleviate the problems of overfitting and gradient disappearance. It introduces an attention mechanism and increases the consideration of the importance of features. Therefore, the model can be divided into three parts: BiLSTM, DNN, and attention mechanism.

The proposal of BiLSTM reasonably and effectively solves the problem of the context of information. It also considers the impact of historical information and future information on the current moment. The basic idea is to construct a recurrent neural network that works in a forward direction in time sequence and a recurrent neural network that works in reverse in time sequence at the same time and combines the two together. The basic structure of the BiLSTM network can be divided into four layers, namely, input layer, forward transmission layer, reverse transmission layer, and output layer. The input layer is responsible for serially encoding the input data so that the input data meet the input requirements of the network. The forward transmission layer is responsible for extracting the forward features of the input sequence from the front to the back. The reverse transmission layer is responsible for extracting the reverse features of the input sequence from back to front. Moreover, the output layer is responsible for integrating the data output by the forward and reverse transmission layers.

For the problem of easy overfitting and gradient disappearance caused by DNN, this article has made the following two considerations in this part: (1) using ReLU as the activation function. Through the understanding of the activation function of the neural network, although the sigmoid function can represent the activation state of the neuron very well because it is an exponential function, there is the problem of a large amount of calculation. Moreover, in its saturation region, the derivative is almost equal to 0, which can easily cause the gradient to disappear. The ReLU function is a linear function, and in comparison, the amount

Figure 5: RNN-based intrusion detection framework.

Figure 6: The structure of BiLSTM-DNN.

of calculation is smaller. And the output of some nodes will be 0, which will help increase the fault tolerance of the network, reduce the influence of the interaction between parameters, and have a certain mitigation effect on overfitting. (2) Use batch standardization. Batch standardization is a method proposed in 2015 and has been widely used in network training. In a deep neural network, the input of the latter layer is the output of the previous layer, and the layers are closely connected. Then the input data distribution of the later layer will change continuously due to the update of the parameters of the previous layer. As the number of layers increases, the change becomes more obvious. BN first solves the mean and variance of each small batch of data and then uniformly standardizes the input. Normalization is usually carried out before the input of each layer to reduce the adverse effects of data distribution, avoid too concentrated data in certain dimensions, and alleviate the problems of gradient disappearance and overfitting.

The attention mechanism was proposed very early, but it has not received extensive attention and in-depth research. Relevant research has introduced it into the image classification task and achieved good results, which made the attention mechanism reenter the researcher's field of vision. It has attracted the attention of scholars in various fields, has become one of the current research hotspots, and has been widely used. The attention mechanism simulates the attention model of the human brain; that is, at a certain point in time, the person's attention is always focused on a certain focal part of the object being viewed while ignoring other parts. Essentially, it is a resource allocation model. Its basic working principle is to give more attention to the key parts we need to pay attention to. However, less attention is given to other parts to achieve the purpose of reasonable allocation of attention resources to reduce or eliminate the adverse effects of noncritical factors. In recent years, the attention mechanism has been widely used in various fields, among which the three major areas of natural language processing, image recognition, and speech recognition are particularly prominent. It can be used alone or as a layer in other models. Common attention mechanisms can be divided into two types: additive attention [32] and dot-product attention [33].

In the attention mechanism, Q represents a certain element of the target ,$\langle K, V\rangle$ represents the element in the original information and a certain value corresponding thereto, and AV represents the attention value obtained by element Q in the current state. As can be seen from the schematic diagram of the attention mechanism, the calculation of the attention value is the process of first solving the correlation between Q and K, and then calculating the weighted sum of V were conducted as follows:

$$Att(Q, S) = \sum_i Sim(Q, K_i) \times V, \tag{5}$$

where S is the source and Sim is the similarity function.

The main difference between the additive attention and the dot-product attention mechanism lies in the calculation of correlation, which is reflected in the calculation method of the correlation between Q and K. However, due to the difference in correlation calculation methods, the processing efficiency of the two is also different under the same conditions, and the latter is higher than the former. Based on this, this article uses scaled dot-product attention proposed in literature [34]. The attention calculation formula of scaled dot-product attention is as follows:

$$Att(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{6}$$

where d_k is the dimension of K, and it eliminates the problem of too small back-propagation gradient due to too large input, which is difficult to learn.

4. Experiments and Discussion

4.1. Dataset and Evaluation Metric. In the experiment, two self-made datasets are used for testing, namely, IADA and IADB, respectively. These two datasets collected the network connection data of two kinds of computer networks for ten consecutive weeks. The training dataset collected six weeks of network traffic data, and the test dataset collected four-weeks of network traffic data. The data distribution of each dataset is illustrated in Table 1. The evaluation metrics used in this work are precision, recall, and F1 score.

4.2. Comparison with Other Methods. To verify the validity and reliability of the computer network intrusion detection algorithm designed in this article, this article compares BiLSTM-DNN with other intrusion detection algorithms. The methods compared include GA-LR [35], PCA-LSTM [36], NDAE [37], and AMSOM [38]. The experimental result is illustrated in Table 2.

Obviously, the method designed in this work can obtain the best performance. On IADA, 95.6% precision, 92.7% recall, and 94.5% F1 score can be obtained. On IADB, 97.2% precision, 93.9% recall, and 95.8% F1 score can be obtained. Compared with the best-listed method, BiLSTM-DNN can obtain 2.1% precision gain, 2.9% recall gain, and 2.8% F1 score gain on IADA, and this method can obtain 1.5%

Table 1: The division of the dataset.

Dataset	Training	Test	Total
IADA	640385	410539	1050924
IADB	680108	438951	1119059

Table 2: Comparison with other methods.

Methos	IADA			IADB		
	Precision	Recall	F1	Precision	Recall	F1
GA-LR	84.9	81.5	83.2	85.7	82.8	84.1
PCA-LSTM	88.2	83.6	86.4	89.4	85.7	88.3
NDAE	91.7	87.8	90.1	93.4	89.9	91.4
AMSOM	93.5	89.8	91.7	95.7	91.2	93.6
BiLSTM-DNN	95.6	92.7	94.5	97.2	93.9	95.8

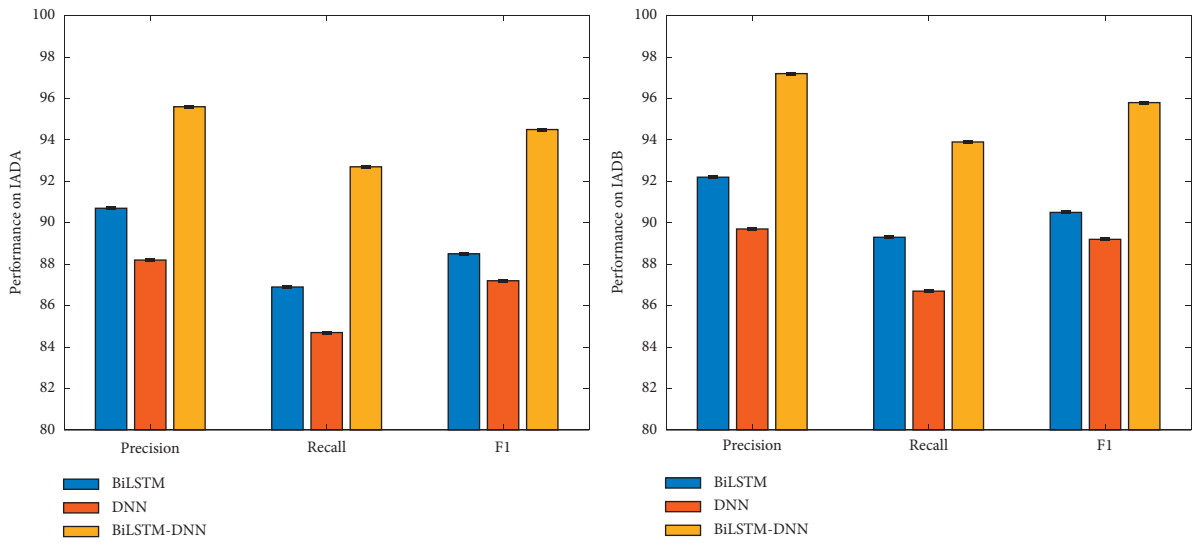

Figure 7: Evaluation on the combination of BiLSTM with DNN.

precision gain, 2.7% recall gain, and 2.2% F1 score gain on IADB. These data can prove the reliability and effectiveness of the BiLSTM-DNN method for computer network intrusion detection.

4.3. Evaluation on Combination of BiLSTM with DNN. As mentioned in the method section, this work combines BiLSTM with DNN. To verify effectiveness for this strategy, this work conducts a comparative experiment to compare the intrusion detection performance in the case of a single BiLSTM and a single DNN. The experimental results are illustrated in Figure 7.

Obviously, the intrusion detection performance when only using BiLSTM is better than when only using DNN. However, neither of these can achieve optimal performance. Only by combining the two to construct the BiLSTM-DNN model designed in this article can the greatest performance improvement be obtained.

4.4. Evaluation on BiLSTM. As mentioned in the method section, this work designs BiLSTM to replace LSTM. To verify effectiveness for this strategy, this work conducts a comparative experiment to compare the intrusion detection performance in the case of BiLSTM and LSTM. The experimental results are illustrated in Figure 8.

Obviously, the method using BiLSTM can obtain the best performance. Compared with the method using LSTM, BiLSTM-DNN can obtain 1.7% precision gain, 1.6% recall gain, and 1.7% F1 score gain on IADA, and this method can obtain 1.7% precision gain, 2.6% recall gain, and 2.1% F1 score gain on IADB. These data can prove the reliability and effectiveness of the BiLSTM method.

4.5. Evaluation on BN. As mentioned in the method section, this work utilizes a BN structure to process the feature. To verify effectiveness for this strategy, this work conducts a comparative experiment to compare the intrusion detection

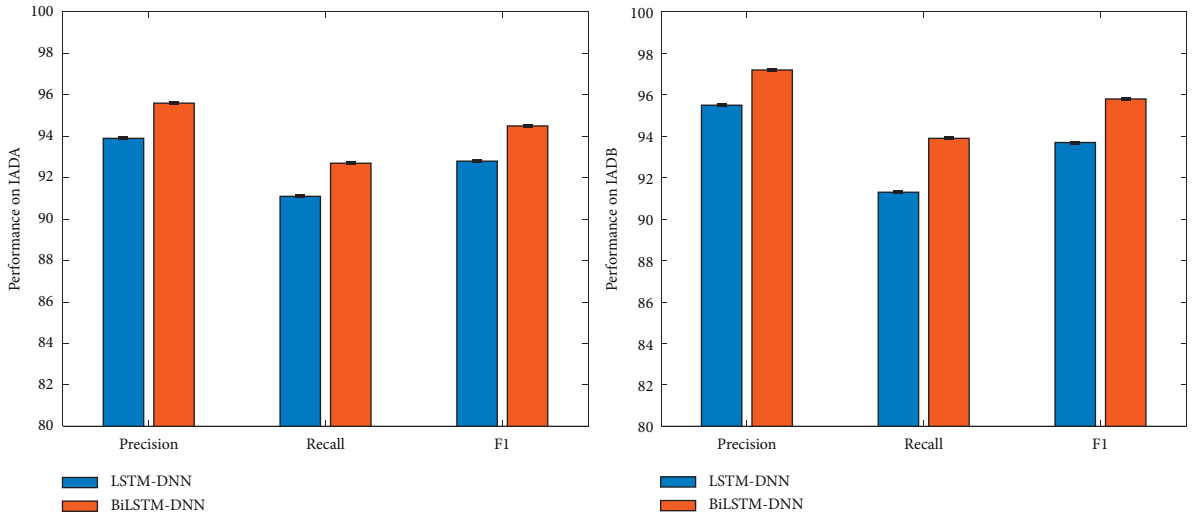

Figure 8: Evaluation on BiLSTM.

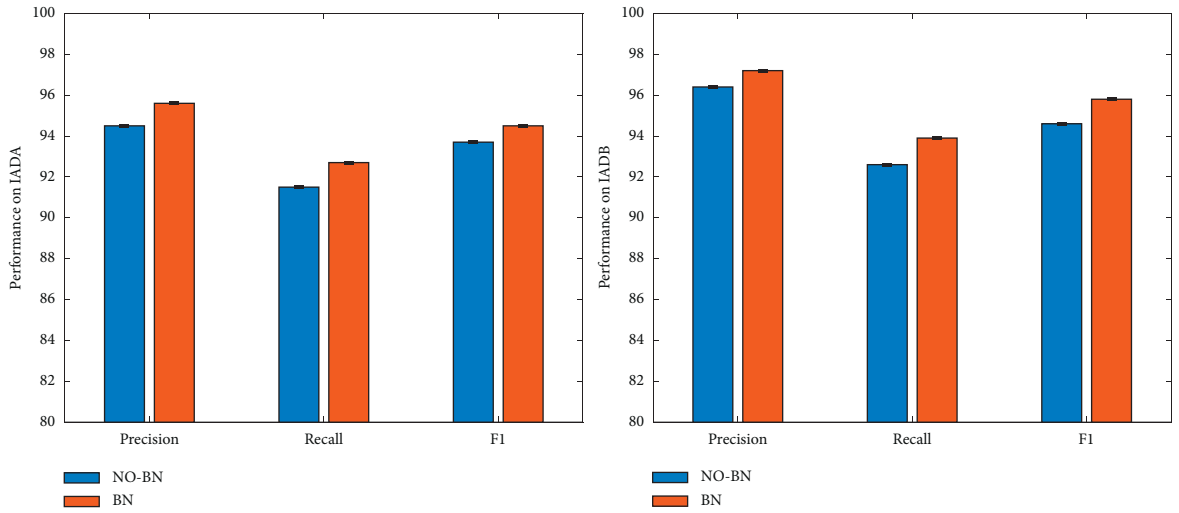

Figure 9: Evaluation on BN.

performance in the case of using BN and not using BN. The experimental results are illustrated in Figure 9.

Obviously, the method using BN architecture can obtain the best performance. Compared with the method not using BN architecture, BiLSTM-DNN using BN can obtain 1.1% precision gain, 1.2% recall gain, and 0.8% F1 score gain on IADA, and this method can obtain 0.8% precision gain, 1.3% recall gain, and 1.2% F1 score gain on IADB. These data can prove the reliability and effectiveness of the BN architecture.

4.6. Evaluation on Attention. As mentioned in the method section, this work utilizes an attention mechanism to learn more discriminatory features. To verify effectiveness for this strategy, this work conducts a comparative experiment to compare the intrusion detection performance in the case of using attention and not using attention. The experimental results are illustrated in Table 3.

Table 3: Evaluation on attention.

Methos	IADA			IADB		
	Precision	Recall	F1	Precision	Recall	F1
NO-ATT	94.4	90.8	93.1	95.7	92.5	93.3
ATT	95.6	92.7	94.5	97.2	93.9	95.8

Obviously, the method using an attention mechanism can obtain the best performance. Compared with the method not using attention mechanism, BiLSTM-DNN using attention can obtain 1.2% precision gain, 1.9% recall gain, and 1.4% F1 score gain on IADA, and this method can obtain 1.5% precision gain, 1.4% recall gain, and 2.5% F1

score gain on IADB. These data can prove the reliability and effectiveness of the attention mechanism.

5. Conclusion

Intrusion anomaly detection is an important part of network security protection. With the development of network technology, there are more and more methods of network intrusion. Traditional intrusion detection technology has been unable to meet the existing network security needs. The extensive application of recurrent neural networks in various fields has brought new ideas to the research of intrusion detection methods. Research on the application of machine learning methods in intrusion detection is one of the current research hotspots in the field of network security. Based on the above research background, this article has carried out in-depth research on the application of recurrent neural networks in network intrusion anomaly detection. The main research work is as follows: (1) a network security emergency response system architecture based on the recurrent neural network is proposed. This system consists of a management center module, a knowledge database module, a data acquisition module, a risk detection tool module, a risk analysis and processing module, a data protection module, and a remote connection auxiliary module. The modules cooperate with each other to complete system functions. (2) Aiming at the risk analysis and processing module, a network intrusion detection model combining BiLSTM and DNN is proposed. In view of the lack of consideration of the before-and-after relevance of intrusion data features and the multifeature problem in existing models, the use of BiLSTM to extract the relevance between features and the use of DNN to extract deeper features are proposed. Aiming at the problem that the model lacks consideration of the importance of features, it is proposed to introduce an attention mechanism into the model to increase the consideration of the importance of features. (3) Massive experiments have verified the reliability and effectiveness of the method proposed in this work.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] M. Fuentes-García, J. Camacho, and G. Maciá-Fernández, "Present and future of network security monitoring," *IEEE Access*, vol. 9, pp. 112744–112760, 2021.

[2] R. Nazir, K. Kumar, S. David, and A. A. Laghari, "Survey on wireless network security," *Archives of Computational Methods in Engineering*, vol. 1, pp. 1–20, 2021.

[3] G. Liu, B. Peng, and X. Zhong, "A novel epidemic model for wireless rechargeable sensor network security," *Sensors*, vol. 21, no. 1, p. 123, 2021.

[4] S. Sengupta, A. Chowdhary, A. Sabur, and D. Huang, A. Alshamrani, S. Kambhampati, A survey of moving target defenses for network security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1909–1941, 2020.

[5] M. Zhao, G. Wei, C. Wei, and Y. Guo, "CPT-TODIM method for bipolar fuzzy multi-attribute group decision making and its application to network security service provider selection," *International Journal of Intelligent Systems*, vol. 36, no. 5, pp. 1943–1969, 2021.

[6] S. Gamage and J. Samarabandu, "Deep learning methods in network intrusion detection: A survey and an objective comparison," *Journal of Network and Computer Applications*, vol. 169, Article ID 102767, 2020.

[7] Y. Zhou, T. A. Mazzuchi, and S. Sarkani, "M-adaboost-a based ensemble system for network intrusion detection," *Expert Systems with Applications*, vol. 162, Article ID 113864, 2020.

[8] J. Pei, K. Zhong, J. Li, J. Xu, and X. Wang, "ECNN: Evaluating a cluster-neural network model for city innovation capability," *Neural Computing & Applications*, pp. 1–13, 2021, https://doi.org/10.1007/s00521-021-06471-z.

[9] T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, "BAT: Deep learning methods on network intrusion detection using NSL-KDD dataset," *IEEE Access*, vol. 8, pp. 29575–29585, 2020.

[10] M. Pawlicki, M. Choraś, and R. Kozik, "Defending network intrusion detection systems against adversarial evasion attacks," *Future Generation Computer Systems*, vol. 110, pp. 148–154, 2020.

[11] Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 1, Article ID e4150, 2021.

[12] C. Deng and H. Qiao, "Network security intrusion detection system based on incremental improved convolutional neural network model," in *International Conference on Communication and Electronics Systems*, pp. 1–5, Coimbatore, India, 21-22 Oct. 2016.

[13] Z. Wu, J. Wang, L. Hu, Z. Zhang, and H. Wu, "A network intrusion detection method based on semantic re-encoding and deep learning," *Journal of Network and Computer Applications*, vol. 164, Article ID 102688, 2020.

[14] R. Magán-Carrión, D. Urda, I. Díaz-Cano, and B. Dorronsoro, "Towards a reliable comparison and evaluation of network intrusion detection systems based on machine learning approaches," *Applied Sciences*, vol. 10, no. 5, p. 1775, 2020.

[15] W. Elmasry, A. Akbulut, and A. H. Zaim, "Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic," *Computer Networks*, vol. 168, p. 107042, 2020.

[16] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," *Security and Privacy (SP)*, pp. 305–316, 2010.

[17] C. F. Tsai, Y. F. Hsu, and C. Y. Lin, "Intrusion detection by machine learning: A review Science Direct," *Expert Systems with Applications*, vol. 36, no. 10, pp. 11994–12000, 2009.

[18] H. Lee, J. Y. Song, and D. Park, "Intrusion detection system based on multi-class SVM," *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, pp. 511–519, 2005.

[19] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in *IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 712–717, Beijing, China, 22-24 July 2017.

[20] C. L. Yin, Y. F. Zhu, J. L. Fei, and X. He, "A deep learning Approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, no. 99, pp. 21954–21961, 2017.
[21] R. C. Staudemeyer, "Applying long short-term memory recurrent neural networks to intrusion detection," *South African Computer Journal*, vol. 56, no. 1, pp. 136–154, 2015.
[22] B. Riyaz and S. Ganapathy, "A deep learning approach for effective intrusion detection in wireless networks using CNN," *Soft Computing*, vol. 24, no. 22, pp. 17265–17278, 2020.
[23] W. Zhong, N. Yu, and C. Ai, "Applying big data based deep learning system to intrusion detection," *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 181–195, 2020.
[24] Y. Zhang, N. Zhang, Y. Zhang, and M. Xiao, "A network intrusion detection method based on deep learning with higher accuracy," *Procedia Computer Science*, vol. 174, pp. 50–54, 2020.
[25] S. Zavrak and M. Iskefiyeli, "Anomaly-based intrusion detection from network flow features using variational auto encoder," *IEEE Access*, vol. 8, no. 99, pp. 108346–108358, 2020.
[26] W. H. Lin, H. C. Lin, P. Wang, B.-H. Wu, and J.-Y. Tsai, "Using convolutional neural networks to network intrusion detection for cyber threats," in *IEEE International Conference on Applied System Invention*, pp. 1107–1110, Chiba, Japan, 13-17 April 2018.
[27] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," *Electronics*, vol. 9, no. 6, pp. 916–920, 2020.
[28] J. H. Lee and K. H. Park, "AE-CGAN model based high performance network intrusion detection system," *Applied Sciences*, vol. 9, no. 20, pp. 4221–4229, 2019.
[29] K. Wu, Z. Chen, and W. Li, "A novel intrusion detection model for a massive network using convolutional neural networks," *IEEE Access*, vol. 6, pp. 50850–50859, 2018.
[30] H. Yao, C. Li, and P. Sun, "Using parametric t-distributed stochastic neighbor embedding combined with hierarchical neural network for network intrusion detection," *International Journal on Network Security*, vol. 22, no. 2, pp. 265–274, 2020.
[31] G. Kumar, "An improved ensemble approach for effective intrusion detection," *The Journal of Supercomputing*, vol. 76, no. 1, pp. 275–291, 2020.
[32] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.
[33] M. T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," 2015. arXiv preprint arXiv:1508.04025.
[34] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," *Advances in Neural Information Processing Systems*, pp. 5998–6008, 2017.
[35] C. Khammassi and S. Krichen, "A GA-LR wrapper approach for feature selection in network intrusion detection," *Computers & Security*, vol. 70, pp. 255–277, 2017.
[36] Z. Gao, Y. Su, and Y. Liu, "Study on intrusion detection based on PCA-LSTM," *Computer Science*, vol. 46, no. 2, pp. 473–476, 2019.
[37] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, 2018.
[38] D. Wu and Y. Liu, "Intrusion detection model based on self-organizing map of variable network structure[J]," *Computer Engineering and Applications*, vol. 56, no. 12, pp. 81–86, 2020.