Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6623554, 10 pages
https://doi.org/10.1155/2021/6623554

Research Article

Network Intrusion Detection Based on an Improved Long-Short-Term Memory Model in Combination with Multiple Spatiotemporal Structures

Xiaolong Huang

School of Information Engineering, Baise University, Baise 533000, China

Correspondence should be addressed to Xiaolong Huang; hsl@bsuc.edu.cn

Received 8 December 2020; Revised 29 December 2020; Accepted 13 April 2021; Published 24 April 2021

Academic Editor: Yuanpeng Zhang

Aimed at the existing problems in network intrusion detection, this paper proposes an improved LSTM combined with spatiotemporal structure for intrusion detection. The unsupervised spatiotemporal encoder is used to intelligently extract the spatial characteristics of network traffic data samples. It can not only retain the overall/nonlocal characteristics of the data samples but also extract the most essential deep features of the data samples. Finally, the extracted features are used as input of the LSTM model to realize classification and identification for intrusion samples. Experimental verification shows that the accuracy and false alarm rate of the intrusion detection model based on the neural network are significantly better than those of other traditional models.

1. Introduction

With the continuous development of the Internet, the network has been integrated into all aspects of people's daily life, so network security has become a problem that network users have to face [1]. The importance of network security has risen to the height of a national strategy. If the network data is attacked, it will have an irreversible impact on the efficiency and security of the enterprise [2]. IDS is defined as an intrusion detection system [3] that can extract and analyze the characteristics of the input data and intercept the detected abnormal data, which greatly improves the security of the system. IDS refers to the establishment of a rule base based on existing knowledge or the training of abnormal behavior characteristics to detect malicious attacks such as computer worms and Trojan, so as to maintain information security. An accurate and stable intrusion detection system is very important to a network security system [4]. Therefore, an intrusion detection system has important research significance.

Intrusion detection is an indispensable defense line in the security system. It collects information from several key nodes in the computer network system, checks whether there are any violations of security policies and signs of attack in the network, identifies threats in the network, and generates alarms, so as to provide real-time protection for internal attacks, external attacks, and misoperations [5]. In fact, intrusion detection is to classify the network traffic packets into two or more categories, dividing each network connection into normal behavior or abnormal attack or further distinguishing which kind of attack it belongs to through a multi-object classification model.

However, with the continuous development of science and technology and the gradual opening of the public network, the operation efficiency and real-time situational awareness efficiency of many facilities and equipment are also improved simultaneously. The scale of the Internet is becoming larger and larger, and the structure is more and more complex [6]. Due to the long-term experience of the attacker, the attack mode is different from the conventional

attack mode in the past, which is more intelligent and complex. At present, many scholars have studied intrusion detection models.

In recent years, many machine learning methods have been widely used to identify various types of attacks in the network and the other applications [7–12]. However, most traditional machine learning algorithms belong to shallow learning, and shallow learning algorithms rely on the construction of model and selection of data features [7]. They cannot efficiently classify large-scale data in the real network environment. Shallow learning cannot meet the requirements of intelligent analysis of massive data and prediction of high-dimensional learning [8]. Intrusion detection based on shallow models, such as traditional data mining and machine learning methods, is difficult to effectively detect various new types of attacks. Deep learning has the potential to extract better representations from massive data to design better detection models [9].

Deep learning methods can improve the overall performance of intrusion detection systems, which is a popular research direction for many researchers, and relevant research results are also emerging in endlessly [13]. He et al. proposed an improved deep learning method for flow-based anomaly detection, and experiment results have showed that deep learning can be applied to software-defined network anomaly detection [14]. Umer et al. [15] proposed a deep belief network-based intrusion detection model to classify network connections and verified the effectiveness of the proposed method on the NSL-KDD dataset. However, these literatures mainly focus on using deep learning methods as part of a pretraining model and then using traditional methods such as decision trees and SVM for classification. Afterwards, some researchers have successively use deep learning methods to build intrusion detection models. Tan et al. [16] proposed an intrusion detection model constructed by a three-layer RNN and proved through experiments that the model can improve the performance of intrusion detection. However, the various layers in the network are partially connected. That is to say, the simplified RNN model does not reflect the ability to learn the deep feature of the data, and the classification performance of the model has not been analyzed in the binary classification. Subsequently, Liang et al. [17] proposed an intrusion detection model constructed by a fully connected RNN, where the model was studied for binary-class and five-class classification on the NSL-KDD dataset. The experimental results showed that the intrusion detection model based on RNN is better than traditional machine learning, which can get a higher detection rate. However, this model does not remove redundant features when training the RNN classification network and is doped with a lot of data noise, so the final classification performance of the model is not ideal. Deep learning methods have natural advantages over shallow learning methods in the face of large-scale data, where better detection results can usually be obtained in intrusion detection. In addition, as the dimensionality of data features increases, the hidden layer structure of deep neural networks will become more complex, and the difficulty of model training will continue to increase. Moreover, the features extracted from the network connection data are often redundant, which will also make the detection rate of the model reduced [18].

Literature [19] proposed an intrusion detection model based on a deep autoencoding network. Although it improves the classification accuracy and detection rate, it ignores the sequence characteristics of intrusion data. Literature [20] proposed the combination of a asymmetric convolutional autoencoder and supports a vector machine to be applied to intrusion detection systems, which significantly reduced the training time, but the detection time was longer and the robustness was not good; literature [21] proposed the combination of the encoder which extracts data features and the extreme learning machine which classifies quickly and effectively, but the disadvantage is that it is easy to fall into overfitting; the PCA-LSTM algorithm proposed in literature [22] can effectively remove the noise information in the sample data. However, when the amount of data is large, the accuracy of data extraction is low; literature [23] proposed a novel hybrid algorithm based on the PCA-ANN model, and the Artificial Neural Network (ANN) has great advantages in reducing training time, unsupervised learning, and highly nonlinear approximation capabilities, but it is easy to fall into a local minimum. In addition, the feed-forward neural network in ANN does not have the function of remembering and using long-term information dependence. In summary, in view of the current network attacks that are intelligent, complicated, and concealed, the amount of data is large, and complexity and feature dimensions are high; this paper proposes a deep learning-based LSTM neural network intrusion detection model. It can solve the problem of gradient disappearance and gradient explosion of the traditional recurrent neural network (RNN). Traditional shallow learning methods cannot detect current intelligent attack methods, such as hiding in the object-host for 1 to 2 years. LSTM has the advantage of using feature data related to long-term dependence and long-term span and has stronger training and detection capabilities than recurrent neural networks. With the rapid development of Internet technology in the current era of big data, there are not only more attacks but also greatly increased network traffic compared with the traditional Internet. Therefore, intrusion data presents the characteristics of large samples and high dimensions. When the PCA feature extraction model is applied to a large number of data samples, the problem of incomplete feature expression will lead to an increase in the false-positive rate of detection [24]. Deep learning has more advantages in processing large samples and high-dimensional data. SDAE adds noise to data information on the basis of the autoencoder model, which enhances the robustness of the input layer of the autoencoding network. Compared with DAE, it adds lost packet technology. Robustness of feature cascade between autoencoding networks has also improved the robustness of intrusion detection. Therefore, some scholars adopted the stacked denoise autoencoder to perform the spatial dimensionality reduction reconstruction for high-dimensional data.

Aimed at the existing problems in network intrusion detection, this paper proposes an improved LSTM combined with spatiotemporal structure for intrusion detection. The

unsupervised spatiotemporal encoder is used to intelligently extract the spatial characteristics of network traffic data samples. It can not only retain the overall characteristics of the data samples but also extract the deep features of the data samples. Deep learning extracts the spatial feature of network traffic data and extracts more complex features iteratively on the basis of preserving the overall characteristics of the data. Finally, the extracted features are used as input of the LSTM model to realize classification and identification for intrusion samples. The model is trained and tested by using the NSL-KDD dataset which is more in line with the data characteristics of the new era than the KDDCUP dataset. Experimental verification shows that the accuracy and false alarm rate of the intrusion detection model based on the neural network are significantly better than those of other traditional models.

2. Long-Short-Term Memory Network

The core of deep learning is to train the weight parameter matrix and bias parameter deeply through multiple neural networks, so as to minimize the error function between the calculated value and the real value in the deep learning model. The derivative of the error function with respect to the weight and bias parameters is obtained, and the weight and bias parameters under the minimum value are determined by using the change trend of the error function. The cost of calculating the loss function of each parameter is very high. The error backpropagation has been regarded as a gradient descent method, which can usually be used to solve the gradient problem. The global gradient solution is transformed into the local gradient solution, which simplifies the calculation process [24].

The current output of a sequence of recurrent neural networks (RNN) is affected by both the current input and the previous output. It has the function of "memory" to the information in the front of the network. When recursing in time, it can be regarded as a limited multilayer deep learning network [25]. For RNN, the basic function of each hidden state layer is to memorize data and add new information to each layer through each iteration so that the information is passed down.

Although RNN can effectively deal with nonlinear time series data, there are still two problems: (1) due to gradient vanishing and gradient explosion, RNN cannot process long time series data and (2) the training of the RNN model needs to determine the intercept length, and its optimal parameters are difficult to obtain by experience. LSTM can effectively solve these problems. LSTM and RNN have the same input and output, but the difference is the internal structure of the hidden layer. LSTM is an improved recurrent neural network and is mainly to overcome the defect of RNN that is difficult to deal with long-distance dependence in practical application, which is the most popular RNN at present. Their structural differences are shown in Figure 1. LSTM has made milestone achievements in many fields such as speech recognition, image description, and natural language processing.

In addition to the external RNN recurrent, the LSTM recurrent network also introduces an internal self-recurrent. The weight in the self-recurrent depends on the context, so that the neural network can selectively forget the old state and ensure the continuous flow of the gradient. Therefore, LSTM is not a simple nonlinear model with element by element operation after transforming input and recurrent units but a complex nonlinear system of a gating unit system including more parameters and control information flow. Its gating unit system consists of three parts: forgetting gate, input gate, and output gate. Its structure is shown in Figure 2.

The weight of the internal self-recurrent is mainly controlled by the forgetting gate. Its activation function adopts the tanh function, so that the value of the weight parameter ranges from 0 to 1, and its output shape is the same with C_{t-1}. They are multiplied point by point to determine the old state of forgetting; the input gate is also called the memory gate, which is used to determine the new state of memory. The output gate is the result of combining the forgetting gate and the memory gate to calculate the value of the next hidden state. The state updating method of the internal recurrent in LSTM is shown and written as follows.

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

$$g_t = \sigma\left(b_g + \sum U_g x_t + \sum W_g h_{t-1}\right), \tag{2}$$

$$q_t = \sigma\left(b_o + \sum U_o x_t + \sum W_o h_{t-1}\right), \tag{3}$$

$$C_t = f_t C_{t-1} + \tanh\left(b + \sum U_i x_t + \sum W_i h_{t-1}\right) g_t, \tag{4}$$

$$h_t = \tanh\left(C_t\right) q_t, \tag{5}$$

where x_t is the current input, h_t and $h_{t\text{-}1}$ are the hidden state of the t-th time and the $(t-1)$-th time, f_t is the forgetting gate, g_t is the input gate, q_t is the output gate, U, W, and B are the weight parameters in the self-recurrent, $\sigma(\cdot)$ and $\tanh(.)$ are sigmoid and hyperbolic tangent activation functions, respectively. The LSTM model training process adopts the time backpropagation algorithm, which is roughly divided into the following four steps [26].

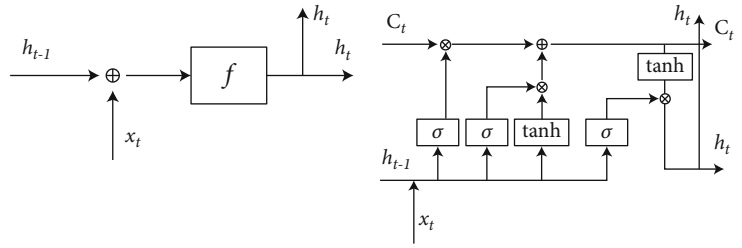
(1) The forward propagation algorithm in the LSTM model is obtained by using the internal self-recurrent state update method in the LSTM model, and the output value of the single module is calculated

(2) From two directions of the time and network layer, the error term of each LSTM is calculated reversely

(3) The gradient of each weight is calculated for the error term

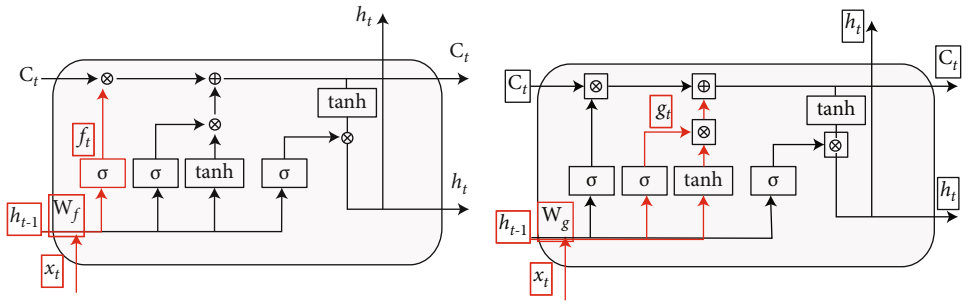
(4) The gradient-based optimization algorithm is applied to update the weight.

The commonly used parameter updating optimization methods include stochastic gradient descent, momentum, AdaGrad, RMSProp, and Adam [27]. In this paper, Adam, namely, adaptive momentum estimation algorithm, is selected as the optimization algorithm. The algorithm combines the advantages of the SGD algorithm and momentum

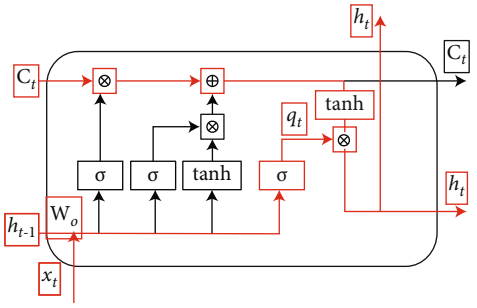


(a) (b)

Figure 1: Hidden layer internal structure diagram: (a) RNN; (b) LSTM.



(a) Forgetting gate (b) Input gate



(c) Output gate

Figure 2: Schematic diagram of LSTM gate structure.

algorithm to search for the parameter space efficiently, and it can be used to perform the "bias correction" of the super parameters, which is less dependent on the initial value and takes less computer resources in the calculation process.

3. Multiple Spatiotemporal Models for LSTM

In order to make the characteristics of each node in the network structure represent different variables, the spatiotemporal characteristics extracted by each node can be used to independently predict different variables [28]. Therefore, after the network is trained, the characteristics of each node can correspond to different variables. This learning method can construct multiple independent channels, and each channel represents a variable and is used to learn each variable. According to the above ideas, it is possible to realize the mining of the spatiotemporal characteristics in the Internet. When the characteristics of each node not only correspond to different variables but also contain spatiotemporal information, the convolution operation is performed on the basis of the spatiotemporal relationship in the Internet, and then, the extracted features can have both the temporal and spatial characteristics of the process. In the process of training, the network can coordinate the learning of temporal and spatial relationships to obtain the intrusion characteristics of spatiotemporal information in the fusion process [29]. Therefore, the soft measurement application can be realized on the basis of collaborative learning of process spatiotemporal characteristics. According to the above ideas, this paper adopts a multichannel network structure to realize spatiotemporal collaborative learning for the intrusion detection system.

In the training process of the deep network, the LSTM and the spatiotemporal model can cooperate with each other in the learning of time series characteristics and spatial characteristics [29] and have the advantages of both. First, the process variables are input into p separate channels, and each channel represents a variable, and the LSTM layer is used to extract the timing features for different variables; then, the timing features extracted by each channel are used as nodes

in the spatiotemporal structure. It is worth noting that although the feature of each node does not directly represent the variable but is extracted from the process variable, it can be targeted to the prediction of each variable after the reverse training of the network. Combined with the adjacency matrix learned from sparse coding, the use of cross-channel spatiotemporal convolution operations can fuse the timing characteristics of different variables in different channels of spatiotemporal structure and associate channels with strong spatial relationships. Therefore, the extracted spatiotemporal features can reflect the inherent characteristics of the process. Since the operation of the spatiotemporal convolutional layer merges the timing characteristics in the related channels, the fused spatiotemporal characteristics can still retain the timing characteristics of the original characteristics; finally, after each channel is processed by two FC layers, the prediction results of each variable are obtained. Through the design of the multichannel network structure, the characteristics of the LSTM and spatiotemporal module can learn the spatiotemporal characteristics which reflect the process characteristics in the process of network training. Each variable can not only learn independently in its own channel but also be effectively correlated to make the learning process of different variables independent and shared.

Therefore, a network intrusion detection model based on multiple spatiotemporal models in LSTM is proposed in this paper [30]. In fact, the essence of the intrusion detection is a classifier model, in which the network flow mixed with abnormal data can be detected. In this paper, a spatiotemporal module is added to the LSTM structure of the classifier to extract features from the data. The process can retain the overall features of the data, and the more complex features are iteratively extracted from the low-level features, thus reducing the feature dimension of the original data. After the spatiotemporal model completes pretraining, the state vectors of each forgetting gate are combined to construct a multilayer neural network. In order to detect abnormal intrusion data, the LSTM with discriminative ability is used as the output layer for network abnormal intrusion detection, identification features, and classification, so as to output the classification detection results of traffic flow data. As shown in Figure 3, the internal loop structure of the LSTM neural network is expanded in time. The input of each time step is the output feature vector sequence of the spatiotemporal model, where x_t, $x_{t\text{-}1}$, and x_{t+1} represent the current state, the previous state, and the next state, respectively. LSTM gradually transfers the memory state backwards in time order by controlling its internal gate structure. The interface between spatiotemporal and LSTM neural networks requires that the input of the LSTM neural network is a sequence of feature vectors. In other words, it is composed of a collection of feature vectors at consecutive T time steps [29]. Therefore, the input sequence of LSTM must be constructed before training. The construction method is as follows: let x_t be the feature vector of the t-th time step; then, the first input sequence is $\{x_1, x_2, \cdots, x_T\}$, and the second is $\{x_2, x_3, \cdots, x_{T+1}\}$. By analogy, all interface input sequences of the spatiotemporal model and LSTM are obtained.

In order to better adapt the anomaly intrusion detection, we built a unidirectional 3-layer LSTM stacked neural network. As shown in Figure 3, the model is a stack of three LSTM layers. $\{x_0, x_1, \cdots, x_N\}$ is the input data after preprocessing. The final output layer is a classification network using the softmax function. The softmax function is essentially a form of probability distribution of neuron output, and the number of nodes in the last output layer is equal to the number of the classification task. The final output data uses one-dimensional arrays $[0, 1]$ and $[-1, 0]$ to represent intrusion traffic and normal traffic. In actual prediction, we use the subscript of the maximum value in the one-dimensional array to represent the prediction result; for example, $[0.0134871, 0.9875801]$ represents intrusion traffic and $[0.9965685, 0.0031123]$ represents normal traffic. The increase in network capacity can easily lead to overfitting of the model. In order to prevent overfitting and improve the generalization performance of the model, we use recurrent dropout regularization to reduce overfitting in the training process. In other words, the input unit of a certain layer is randomly set to 0 with a certain probability, and the purpose is to break the accidental correlation in the training data of this layer [30]. On the input data of the model, the dataset is transformed into the form of input data required by the model through feature extraction and preprocessing of the imported dataset. Each vector contains data with N feature values, where the batch size is 60 and the time step is 100 in the model.

The LSTM layer and spatiotemporal layer in the above multichannel network structure are defined by equations (1)–(5), and the FC layer is defined by

$$H^{(l)} = \sigma\left(H^{(l-1)}\mathrm{W}^{(l)} + b^{(l)}\right), \tag{6}$$

where $H^{(l)}$ is the implicit feature of layer l, $W^{(l)}$ and $b^{(l)}$ are weight and bias parameters, respectively, and $\sigma(\cdot)$ is the ReLu activation function. The loss function of our proposed model is defined by formula (7). The model parameters in each layer are optimized by gradient descent, where Y and $\widehat{Y}$ represent the real and predicted values of variables, respectively. In order to avoid overfitting in training, the l^2 norm regularization term for model parameters is introduced, and the weight coefficient of the regularization term is γ.

$$\mathrm{loss}\left(Y, \widehat{Y}\right) = \sum_{i=1}^{m}\sum_{j=1}^{p}\left(Y_{ij} - Y\wedge_{ij}\right)^2 + \gamma\|W\|^2. \tag{7}$$

In practical application, first of all, we need to use the variable matrix in the training set and use the sparse coding to learn the spatiotemporal structure among variables and keep the spatiotemporal structure unchanged in the process of subsequent model training and testing. Our proposed model can be obtained by constructing the multichannel network structure. Through the multichannel network structure designed by our proposed model, LSTM and spatiotemporal modules are successively used to mine and learn the temporal and spatial characteristics of the intrusion detection process,

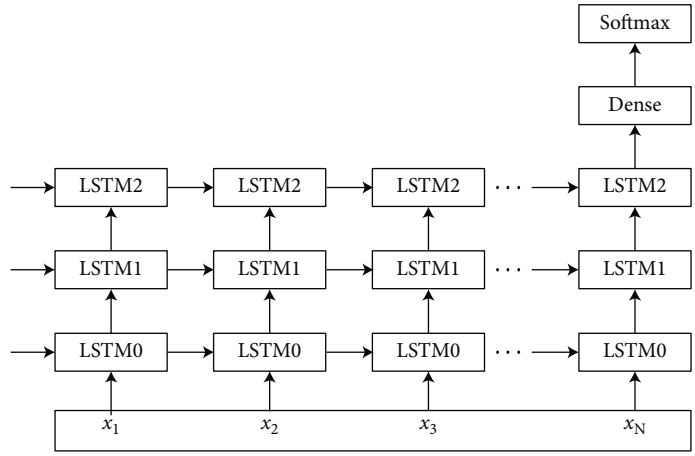


Figure 3: Improved LSTM combined with spatiotemporal structure.

and the features that can represent the internal temporal and spatial properties of the process are obtained for the intrusion detection system. Compared with the traditional model which only considers the temporal sequence characteristics of the process, our proposed model makes full use of the advantages of LSTM and spatiotemporal module and effectively realizes the intrusion detection on the basis of spatiotemporal collaboration.

4. Experimental Results and Analysis

4.1. Experimental Environment. In order to verify the abnormal network attack detection model based on the multispatiotemporal and long-short-term memory model in this paper, a simulation experiment environment is built [31]. The experiment uses the Keras2.2.4 deep learning framework based on tensorflow-GPU1.13 for simulation, the operating system is Windows 10, the Intel i5-6300HQ 4-core processor is CPU, the memory size is 8 G, and the NVIDIA GTX960 graphics card is used to accelerate the running speed of the model.

4.2. Evaluation Index. In this experiment, four indicators are constructed to evaluate the advantages and disadvantages of the model, including the accuracy rate (ACC), false-positive rate (FPR), false-negative rate (FNR), and prediction rate (DR). Through the evaluation indexes, we can get the advantages and disadvantages of the model and then adjust the model parameters until the evaluation index is optimal [31], which means that the model is optimal at this time. As shown in Table 1, T^+ represents that the predicted result is intrusion attack, namely, the number of successful predicted samples; T^- represents the number of samples predicted as normal and the actual value as the number of samples with successful normal prediction. Therefore, we can use the ACC to denote the accuracy: $\mathrm{ACC} = (T^- + T^+)/(T^- + T^+ + F^- + F^+)$. F^- represents that the predicted result is normal, but the actual value is attacked; namely, the prediction fails. Therefore, the system fails to detect the number of samples which can be denoted as the false-negative rate: $\mathrm{FNR} = F^-/(F^+ + T^-)$; F^+ represents that the prediction is an intrusion attack and the actual is normal; namely, prediction fails. Therefore, the number of samples of false positives in the system is denoted as the false-positive rate: $\mathrm{FPR} = F^+/(F^+ + T^+)$. The detection rate (DR) can be written as $\mathrm{DR} = T^+/(F^- + T^+)$.

Table 1: Evaluation matrix.

		Prediction	
		Intrusion attack	Normal
Actual	Intrusion attack	T^+	F^-
	Normal	F^+	T^-

4.3. Experimental Dataset. Recently, most of the standard examples used in the field of intrusion detection are still the KDDCUP99 dataset, and the test results in KDDCUP99 are better under certain conditions [32]. However, the KDDCUP99 dataset simulated 20 years ago is no longer suitable for the modern intelligent and complex attack methods, such as penetration utilization, SQL injection, APT, and complex hidden attack forms. The UNSW-NB15 dataset was created by the Australian Network Security Center (ANSC) in 2015 [33], which is a new dataset in the field of intrusion detection, which reflects the modern network traffic pattern. This dataset contains a large number of low occupancy intrusion and deep structured network traffic information and has 9 different types of modern attacks and 49 features. It has 5 attack types more than NSL-KDD, including 2540044 samples and 9 types of attacks, which are fuzzers, DoS, analysis, reconnaissance, exploit, shellcode, worm, backdoor, and generic. The UNSW-NB15 dataset contains 5.5 million records. Since the normal data sample in the UNSW-NB15 dataset is more than 10 times that of attack-type data sample, using a small number of sample oversampling will lead to excessive attack class duplicate samples, resulting in an overfitting phenomenon. If a simple undersampling strategy is used, normal samples will lose key information. Therefore, in order to reduce the error of experimental results caused by unbalanced data samples, the SMOTE oversampling method is adopted in this paper, as

Table 2: The description of the used dataset.

Type	Fuzzers	DoS	Analysis	Reconnaissance	Exploit	Shellcode	Worm	Backdoor	Generic	Normal
Training	17245	11357	2000	4000	33652	1285	125	1587	625	58220
Testing	6005	4072	655	15892	12854	320	42	550	127	32125

shown in formula (8). The principle is to use the existing attack sample data to find the random samples in the same kind of samples, and the linear difference will generate a new sample x, and repeat this process until the samples are balanced.

$$\hat{x} = x + \delta(x_i - x), \tag{8}$$

where $\hat{x}$ is the new sample, x is the actual sample, and x_i is a randomly selected sample from a sample near x. The dataset is divided into two parts, including the training dataset and the test dataset, as shown in Table 2.

4.4. Qualitative and Quantitative Analysis

4.4.1. Effectiveness Analysis for the Spatiotemporal Module. In order to verify the effectiveness of the introduced spatiotemporal module, different modules are tested under the condition that other parameter settings remain unchanged, as shown in Figure 4. It is not difficult to see from Figure 4 that the detection rate of the spatiotemporal module selected in this paper is higher than that of the other three modules, especially for the accuracy rate of normal data compared with the efficiency of other modules; the maximum difference is 4.30%, which can reflect the significant advantage of adding a spatial structure feature between classes in the spatiotemporal module.

4.4.2. Effectiveness Analysis for LSTM Depth Level. This paper analyzes the intrusion detection performance of the proposed convolutional neural network model and other four convolutional neural network structures and takes the detection accuracy and false alarm rate as the evaluation criteria of this experiment. The results are shown in Table 3. Through the comparative analysis, the detection accuracy of our proposed intrusion detection model is the highest, which is 1.53% higher than that of the IRES model [34]. Compared with the ResNet model with a low false alarm rate, its accuracy rate is increased by 5.82%; the false alarm rate is lower than that of the IRES model and LeNet model [35]. On the other hand, we can see that the depth of network structure is gradually deepened in the process of applying a convolutional neural network to intrusion detection, which is also a factor of a high false alarm rate of the comparison models. With the deepening of the network structure, the phenomenon of gradient vanish is obvious, but the special structure of our proposed convolutional neural network ensures that the problem of gradient vanish is improved in depth level.

4.4.3. Performance Analysis. In the experiment, the binary classification form is adopted, the input vector has 47 dimensional features, the intrusion attack vector is marked with 0, and the nonattack is marked with 1, which can improve the efficiency and the timeliness of the intrusion detection system. We can judge whether it is abnormal data through the model first and then classify the abnormal data by supervised learning. By using SMOTE sampling, it is divided into the training set and test set. There are 30 training datasets, with each dataset containing 3000 randomly selected samples, and 1500 test datasets, with each dataset containing 3000 randomly selected samples. In the process of feature extraction by a spatiotemporal network, the relationship between the number of iterations and the loss value is shown in Figure 5. After repeated iterations, the effect of the spatiotemporal model tends to be stable. The number of neurons in each layer was 47, 23, 23, and 23. In our improved LSTM model, three hidden layers are set up in LSTM. The number of neurons in the first layer is 128, that in the second layer is 256, and that in the third layer is 128. The batch size and epoch times are 200 and 800, respectively. When the learning

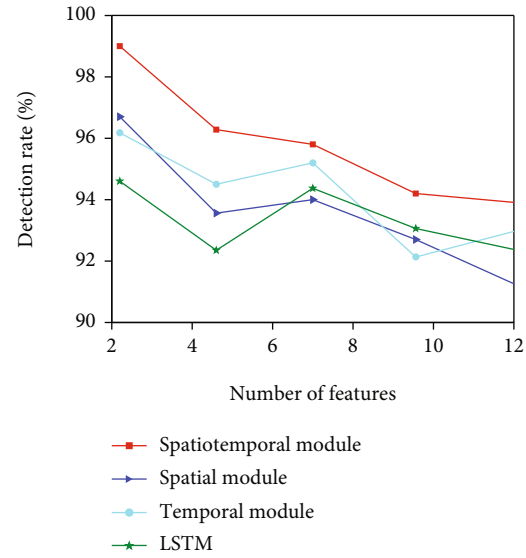


Figure 4: Detection rate for different modules.

Table 3: Comparison analysis.

Models	Accuracy (ACC)	DR	FNR	FPR
ResNet	91.36	1.35	0.25	0.27
LeNet	92.18	1.98	0.42	0.19
LSTM	97.22	2.25	023	0.22
IRES	96.25	3.59	0.28	0.25
Proposed	98.76	0.76	0.17	0.12

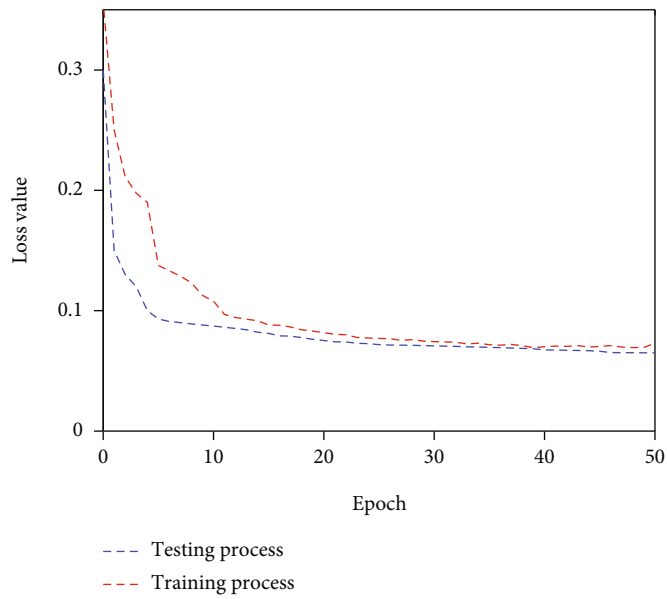


Figure 5: Relationship between the number of iterations and the loss value in training and testing processes.

rate of the model is too large or too small, the accuracy of the testing set is very low. Since the choice of the learning step has a great impact on the performance of the intrusion detection system, the time step is selected as 50 through repeated experiments.

In this experiment, categorical_crossentropy, which is specially used for solving multiclassification problems, is selected as the optimization objective function, and the Adam optimization algorithm is adopted to carry out the backpropagation training of the model. As shown in Figure 6, in order to highlight the ability of spatiotemporal feature extraction, the model in this paper is compared with LeNet [35], ResNet [36], and LSTM [37] without feature extraction. The results show that the accuracy of LeNet is not significantly improved in high-dimensional data samples, and the accuracy rate of LSTM is significantly lower than that of our proposed model. With the increase in the number of samples, the accuracy of LSTM shows a downward trend, which shows that the LSTM and spatiotemporal structure have strong data feature extraction ability, and with the increase in time complexity, the spatiotemporal structure has more and more advantages than the spatial structure [10, 11, 38]. Therefore, the detection effect of the intrusion detection algorithm in this paper is obviously better than that of the LeNet, ResNet, and LSTM intrusion detection model. The experiment selects the existing LSTM model, MLP (multilayer perception), ELM (extreme learning machine), and the popular deep belief model (DBN) as comparison models [12]. ResNet has multiple hidden layers and is fully connected with the input layer, which can be processed by nonlinear activation function [39]; DBN is a probability generation model, which has the characteristics of multiple hidden layer models and can learn the essence of the dataset by learning a deep nonlinear network structure [40-46]; ELM is a simple forward propagation by setting the connection weights and thresholds of the input layer and hidden layer

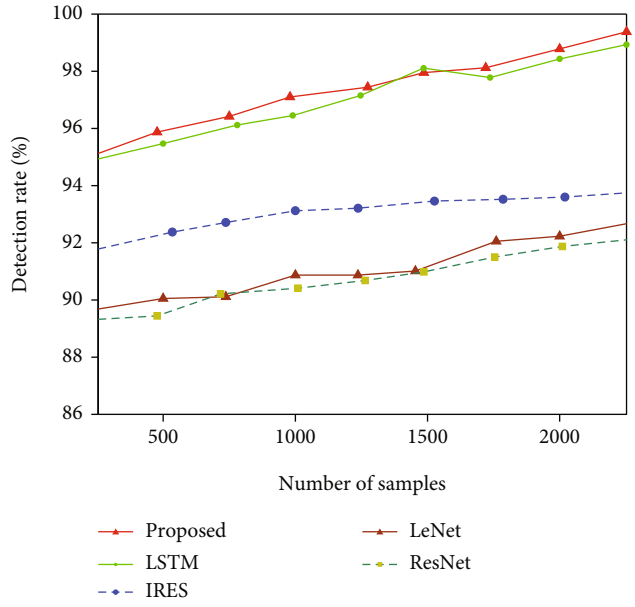


Figure 6: Comparison of detection rates for different models.

randomly and adjusting the weights without iteration, and the learning speed is improved. As can be seen from Table 3, the detection rate and false alarm rate of the proposed algorithm in this paper are improved compared with the current popular deep learning methods [47]. As shown in Figure 7, through the comparison of ROC curves, since the area of the ROC curve [48] of the proposed model is the largest, it can be proven that the proposed intrusion detection algorithm model not only improves the accuracy of intrusion detection but also significantly reduces the false alarm rate and achieves remarkable results in improving the model performance and detection efficiency.

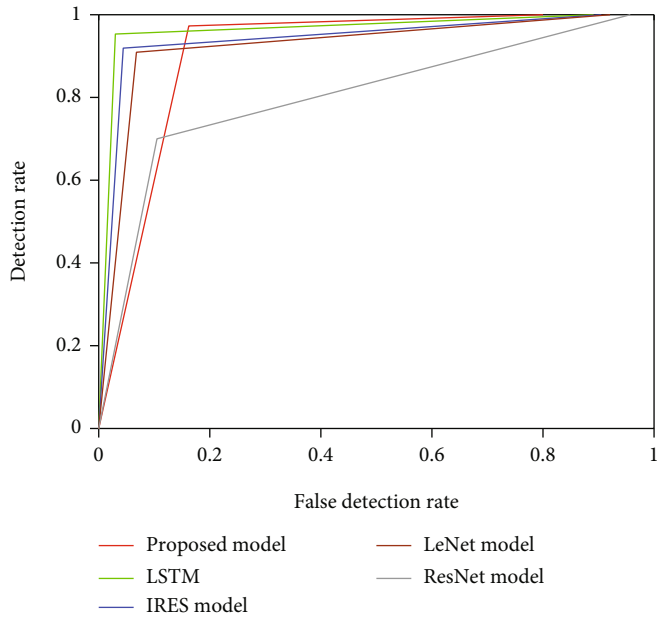


Figure 7: Comparison of ROC curves for different models.

5. Conclusion

Aimed at the existing problems in network intrusion detection, this paper proposes an improved LSTM combined with spatiotemporal structure for intrusion detection. The unsupervised spatiotemporal encoder is used to intelligently extract the spatial characteristics of network traffic data samples. It can not only retain the nonlocal characteristics of the data samples but also extract the deep features of the data samples. The model is trained and tested by using the NSL-KDD dataset which is more in line with the data characteristics of the new era than the traditional KDDCUP99 dataset. The experimental results show that the proposed intrusion detection model has achieved remarkable results in improving the accuracy, performance, and efficiency of intrusion detection. The model only works well in the simulation dataset but needs to be tested in the actual network environment to verify the real performance of the model, which will be our working direction in the future.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work was supported in part by the Cultural Science Research of Jiangsu Province under Grant 18YB27.

References

[1] M. Weizhi, "Intrusion detection in the era of IoT: building trust via traffic filtering and sampling," *Computer*, vol. 51, no. 7, pp. 36–43, 2018.

[2] Z. Sun, Y. Xu, G. Liang, and Z. Zhou, "An intrusion detection model for wireless sensor networks with an improved V-detector algorithm," *IEEE sensors journal*, vol. 18, no. 5, pp. 1971–1984, 2018.

[3] S. Manimurugan, "Intrusion detection in cloud environment using hybrid genetic algorithm and back propagation neural network," *International Journal of Communication Systems*, vol. 56, no. 9, pp. 258–267, 2018.

[4] L. Li, H. Zhang, H. Peng, and Y. Yang, "Nearest neighbors based density peaks approach to intrusion detection," *Chaos Solitons & Fractals*, vol. 110, pp. 33–40, 2018.

[5] A. Karami, "An anomaly-based intrusion detection system in presence of benign outliers with visualization capabilities," *Expert Systems with Applications*, vol. 108, no. 6, pp. 36–60, 2018.

[6] H. Jun-Ho, "Implementation of lightweight intrusion detection model for security of smart green house and vertical farm," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, 2018.

[7] X. An, X. Zhou, X. Lü, F. Lin, and L. Yang, "Sample selected extreme learning machine based intrusion detection in fog computing and MEC," *Wireless Communications and Mobile Computing*, vol. 2018, 10 pages, 2018.

[8] R. Abdulhammed, M. Faezipour, A. Abuzneid, and A. AbuMallouh, "Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic," *Electronics Letters*, vol. 56, no. 19, pp. 23–39, 2018.

[9] G. Mamalakis, C. Diou, A. L. Symeonidis, and L. Georgiadis, "Of daemons and men: reducing false positive rate in intrusion detection systems with file system footprint analysis," *Neural Computing & Applications*, vol. 31, no. 3, pp. 7755–7767, 2018.

[10] Y. Jiang, Y. Zhang, C. Lin, D. Wu, and C.-T. Lin, "EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1752–1764, 2021.

[11] Y. Jiang, G. Xiaoqing, D. Wu et al., "A novel negative-transfer-resistant fuzzy clustering model with a shared cross-domain transfer latent space and its application to brain CT image segmentation," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 1, pp. 40–52, 2021.

[12] Y. Jiang, K. Zhao, K. Xia et al., "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," *Journal Medical Systems*, vol. 43, no. 5, 2019.

[13] K. Leyli, S. Erkay, and A. Halit, "Intrusion detection over encrypted network data," *The Computer Journal*, vol. 63, no. 4, pp. 604–619, 2020.

[14] D. He, Q. Qiao, Y. Gao et al., "Intrusion detection based on stacked autoencoder for connected healthcare systems," *IEEE Network*, vol. 33, no. 6, pp. 64–69, 2019.

[15] M. F. Umer, M. Sher, and Y. Bi, "A two-stage flow-based intrusion detection model for next-generation networks," *Plos One*, vol. 13, no. 1, 2018.

[16] X. Tan, S. Su, Z. Huang et al., "Wireless sensor networks intrusion detection based on SMOTE and the random forest algorithm," *Sensors*, vol. 133, no. 26, pp. 64–69, 2019.

[17] J. Liang, M. Ma, M. Sadiq, and K.-H. Yeung, "A filter model for intrusion detection system in vehicle ad hoc networks: a hidden Markov methodology," *Knowledge-Based Systems*, vol. 23, no. 16, pp. 214–219, 2020.

[18] M. K. Prasath and B. Perumal, "A meta-heuristic Bayesian network classification for intrusion detection," *International Journal of Network Management*, vol. 29, no. 3, pp. e2047.1–e2047.12, 2019.

[19] Z. Liu, M. Zhou, W. Nie, L. Xie, and Z. Tian, "Indoor intrusion detection based on fuzzy membership-aided Dempster-Shaper theory," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 4, pp. 3687–3696, 2020.

[20] G. Yang, X. Yu, L. Xu, Y. Xin, and X. Fang, "An intrusion detection algorithm for sensor network based on normalized cut spectral clustering," *PLoS ONE*, vol. 14, no. 10, 2019.

[21] Z. Tang, S. Wang, J. Huo, H. Guo, H. Zhao, and Y. Mei, "Bayesian framework with non-local and low-rank constraint for image reconstruction," *Journal of Physics Conference Series*, vol. 787, 2017.

[22] D. B. Gothawal and S. V. Nagaraj, "Anomaly-based intrusion detection system in RPL by applying stochastic and evolutionary game models over IoT environment," *Wireless Personal Communications*, vol. 110, no. 3, pp. 789–795, 2019.

[23] H. Deng, "An improved two-steps saliency detection algorithm based on binarized normed gradients and nuclear norm model in video sequences," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 4, pp. 841–852, 2018.

[24] S. Priyanga, M. R. G. Raman, S. S. Jagtap, N. Aswin, K. Kirthivasan, and V. S. S. Sriram, "An improved rough set theory based feature selection approach for intrusion detection in SCADA systems," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 5, pp. 3993–4003, 2019.

[25] K. R. C. Boni, L. Xu, Z. Chen, and T. D. Baddoo, "A security concept based on scaler distribution of a novel intrusion detection device for wireless sensor networks in a smart environment," *Sensors*, vol. 20, no. 17, 2020.

[26] X. Tian, H. Li, and H. Deng, "Object tracking algorithm based on improved Siamese convolutional networks combined with deep contour extraction and object detection under airborne platform," *Journal of Imaging Science and Technology*, vol. 12, no. 6, pp. 2369–2375, 2020.

[27] B. S. Bhati, G. Chugh, F. Al-Turjman, and N. S. Bhati, "An improved ensemble based intrusion detection technique using XGBoost," *Transactions on Emerging Telecommunications Technologies*, vol. 18, no. 11, pp. 3909–3917, 2020.

[28] X. Tian, H. Li, and H. Deng, "Object tracking algorithm based on improved context model in combination with detection mechanism for suspected objects," *Multimedia Tools and Applications*, vol. 78, no. 4, pp. 259–268, 2019.

[29] G. Spathoulas, G. Theodoridis, and G. P. Damiris, "Using homomorphic encryption for privacy-preserving clustering of intrusion detection alerts," *International Journal of Information Security*, vol. 62, no. 33, 2020.

[30] G. Folino, F. S. Pisani, and L. Pontieri, "A GP-based ensemble classification framework for time-changing streams of intrusion detection data," *Soft Computing*, vol. 24, no. 23, pp. 17541–17560, 2020.

[31] K. Vieira, F. L. Koch, J. B. M. Sobral, C. B. Westphall, and J. L. de Souza Leao, "Autonomic intrusion detection and response using big data," *IEEE Systems Journal*, vol. 14, no. 2, pp. 1984–1991, 2019.

[32] D. Zheng, Z. Hong, N. Wang, and P. Chen, "An improved LDA-based ELM classification for intrusion detection algorithm in IoT application," *Sensors*, vol. 20, no. 6, 2020.

[33] H. Zhang, L. Huang, C. Q. Wu, and Z. Li, "An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," *Computer Networks*, vol. 177, 2020.

[34] R. H. Dong, X.-Y. Li, Q.-Y. Zhang, and H. Yuan, "Network intrusion detection model based on multivariate correlation analysis – long short-time memory network," *IET Information Security.*, vol. 14, no. 2, pp. 166–174, 2020.

[35] V. Dutta, M. Choraś, M. Pawlicki, and R. Kozik, "A deep learning ensemble for network anomaly and cyber-attack detection," *Sensors*, vol. 20, no. 16, 2020.

[36] N. Ding, H. X. Ma, H. Gao, Y. H. Ma, and G. Z. Tan, "Real-time anomaly detection based on long short-term memory and Gaussian mixture model," *Computers & Electrical Engineering*, vol. 79, no. 5, 2019.

[37] A. Diro and N. Chilamkurti, "Leveraging LSTM networks for attack detection in fog-to-things communications," *IEEE Communications Magazine.*, vol. 56, no. 9, pp. 124–130, 2018.

[38] L. Nicholas, S. Y. Ooi, Y. H. Pang, S. O. Hwang, and S.-Y. Tan, "Study of long short-term memory in flow-based network intrusion detection system," *Journal of Intelligent and Fuzzy Systems*, vol. 35, pp. 5947–5957, 2018.