Network Anomaly Traffic Detection Algorithm Based on RIC-SC-DeCN

Xingyu Gong, Ke Cao, Na Li, and Pengtao Jia

College of Computer Science and Technology, Xi’an University of Science and Technology, Xi’an 710054, China

Correspondence should be addressed to Ke Cao; 208694911@qq.com

Received 27 November 2021; Accepted 4 May 2022; Published 24 May 2022

1. Introduction

With the increasing development of computer networks, the Internet has been integrated into all aspects of people’s lives. Shopping, medical [1], communication [2], and other aspects of data are also transmitted and interacted through the network. At the same time, network security incidents occur frequently, and various network attacks often cause abnormal changes in network traffic. The network abnormal traffic detection technology proposed to solve the above network security problems is one of the important means to ensure network security.

In the detection of network abnormal traffic, the network traffic data attack forms are diverse and have the problem of noise feature. Therefore, reasonable feature selection for data can not only reduce the resource consumption of processing, storing, and transmitting data but also improve the accuracy of anomaly detection [3, 4]. In recent years, many scholars have conducted in-depth research on the feature selection of network traffic data. The study in [5] uses the chi-square test for feature selection, which can effectively handle high-dimensional data of network traffic. The work in [6] proposes an adaptive binning feature selection algorithm based on information gain for the problem of long training and detection time for anomaly detection systems based on deep learning. The research in [7] proposes a random forest algorithm based on recursive feature elimination (RFE-RF) to address the low accuracy of high-dimensional data in classification tasks.

In addition, in the detection of network abnormal traffic, the pooling operation in the convolutional neural network used can easily lead to the loss of original feature information [8, 9]. In response to this problem, [10] proposes a deep convolutional neural network model based on
deconvolution feature extraction, which can automatically extract rich implicit features from bottom-level boundaries to high-level objects. The work in [11] proposes an improved one-dimensional convolutional neural network, which removes the pooling operation in the convolutional neural network and preserves the complete traffic information as much as possible. However, this will slow down the feature extraction and the network. The training time of the model also becomes longer.

For the feature selection problem mentioned above, the research in [5, 6] excessively removes redundant information and reduces the detection accuracy, and the work in [7] has a high computational cost for feature selection. In addition, for the information loss problem, the network anomaly detection algorithm proposed in [10, 11] improves the information loss problem, but the disadvantage is that the training time of the model becomes longer. In view of the above-mentioned problems of unsatisfactory feature selection and information loss, this paper proposes a network abnormal traffic detection algorithm based on RIC-SC-DeCN. Firstly, for the problem of data redundancy and high dimensionality, a RIC feature selection mechanism is proposed, which reduces data redundancy through the maximum information correlation feature selection algorithm and recursive feature elimination method. Secondly, aiming at the problem of information loss of network traffic data, a SC-DeCN model is proposed, which combines the deconvolutional neural network to reconstruct the input signal to reduce the information loss. Finally, the RIC mechanism and the SC-DeCN model are merged to form a network abnormal traffic detection algorithm based on RIC-SC-DeCN.

Our contributions are as follows:

1. A RIC feature selection mechanism is proposed. Through the maximum information correlation feature selection algorithm and recursive feature elimination method to reduce data redundancy, network traffic features can be extracted more efficiently.

2. A SC-DeCN model is proposed. The combination of shallow convolutional neural network information and deep deconvolutional neural network information makes feature extraction more accurate, thereby reducing information loss.

3. Combine the proposed RIC feature selection mechanism with the SC-DeCN model to form a network abnormal traffic detection algorithm based on RIC-SC-DeCN. Compared with other classification algorithms, the method proposed in this paper greatly reduce the training time of the model and improve the accuracy of anomaly detection.

The remainder of this paper is organized as follows. Section 2 introduces the work related to network abnormal traffic detection; Section 3 mainly describes the RIC-SC-DeCN algorithm; Section 4 introduces the experimental dataset, and describes the experimental results and analysis in detail; the conclusion and future work part are in Section 5.

2. Related Work

Feature selection techniques have been widely used in many fields. Among them, in the detection of network abnormal traffic, many scholars use feature selection methods to reduce high-dimensional to low-dimensional data, thereby reducing the time for model training and detection. Aiming at the multiobjective feature selection problem in intrusion detection systems, Zhu et al. [12] proposed a population evolution strategy based on a special control method and predefined multiobjective search, which can distinguish anomaly types. Zhao et al. [13] considered the redundancy between features and the influence between features and classes and proposed a new Redundancy Penalized Feature Mutual Information Algorithm (RPFMI). Sumaiya Thaseen et al. [14] utilized chi-square feature selection and an ensemble of Support Vector Machines (SVM), Modified Naive Bayes (MNB), and LP Boost classifiers to build an intrusion detection model. Ran [15] proposed an MMR-based network traffic feature selection method, which reduced the dimension of network traffic data. However, none of the methods proposed above make a reasonable feature selection for network traffic datasets. Redundant features will be removed excessively, resulting in the loss of important information.

Deep learning models have powerful feature extraction capabilities and do not require expensive manual feature engineering, so they are widely used in network abnormal traffic detection tasks. Aceto et al. [16] designed a traffic classifier based on automatically extracted features, using deep learning as a feasible strategy for network abnormal traffic detection. Erxue [17] utilizes word embeddings and text convolutional neural networks to extract effective information from the payload, synthesizes statistical features and payload features, and proposes a new TR-IDS intrusion detection system.

Due to the powerful feature capturing capabilities of convolutional neural networks, many scholars have conducted in-depth research and proposed various variants [18–21]. Basumalik et al. [18] improved a convolutional neural network to identify vulnerable points in PMU networks. Zhang et al. [19] proposed a fraud detection model based on convolutional neural networks, which constructed an input feature ranking layer to reorganize the original transaction features to form different convolutional patterns. Gu et al. [20] proposed an improved bidirectional linear convolutional neural network model. In order to improve the performance of network intrusion detection, Tian et al. [21] adopted the method of Faster Regional Convolutional Neural Network (Faster R-CNN) to complete network anomaly detection.

In addition, there are many research works that integrate different types of neural networks to improve the accuracy of network abnormal traffic detection [22, 23]. Kim and Cho [22] proposed a network abnormal traffic detection method based on a hybrid algorithm of convolutional neural network and long-short-term memory network, using a combination of CNN and LSTM to learn and classify traffic packets in time and space, preserving the order of feature
sequences characteristics to more accurately identify network traffic with hierarchical spatiotemporal characteristics. Ding and Peng [23] proposed a multilayer network structure CNN algorithm to convert the network traffic data in KDD99 into data that can be input by a convolutional neural network. Convolution kernels of different scales are used to extract different levels of features from a large number of high-dimensional unlabeled raw data, which greatly improves the classification performance. Javed et al. [24] proposed a convolutional neural network with a multistage attention mechanism that converts data traffic into vectors for anomaly detection. Experimental results show that the proposed method achieves good performance on both single-source and mixed multisource anomaly types. Khan et al. [25] proposed a new deep learning model (TSDL) that introduced a low-cost DSAE method that is able to learn useful feature representations from large amounts of unlabeled data and automatically and efficiently sort.

3. Network Anomaly Traffic Detection Algorithm

This section provides an overview of our proposed feature selection mechanism and neural network model, which is the main part of this paper.

3.1. The Overall Framework of Network Anomaly Traffic Detection Algorithm. The overall framework of the RIC-SC-DeCN network abnormal traffic detection algorithm proposed in this paper is shown in Figure 1. The algorithm includes three modules: data preprocessing, feature selection, and abnormal detection. The main task of the data preprocessing module is to standardize the features and one-hot encoding the labels and input the processed data to the feature selection module. In the feature selection module, the RIC feature selection mechanism is adopted. Firstly, remove features with zero variance; Second, use the feature selection with maximum information correlation and select some features with the best phenotype according to the correlation between the feature and the phenotype. Finally, the optimal feature subset is selected by recursive feature elimination method, and the feature subset is input to the anomaly detection module. The anomaly detection module inputs the optimal feature subset into the convolutional neural network to automatically learn the traffic features and then deconvolutes the deep information of the convolutional, through fusing the shallow information in the convolutional neural network with the deep information in the deconvolutional neural network, the SC-DeCN model is constructed, and the network abnormal traffic is detected by this model.

4. RIC Feature Selection Mechanism

For the problem that most network traffic data is high-dimensional and accompanied by noise, this paper proposes a feature selection mechanism of Recursive Information Correlated (RIC), which extracts key features from the original network traffic data and reduces the impact of noise features on network abnormal traffic detection. The RIC feature selection mechanism includes three steps: first, removing irrelevant features; second, performing feature selection with maximum information correlation; finally, using recursive feature elimination to select features.

4.1. Remove Irrelevant Features. The network traffic dataset contains features with a variance of 0, and these features have no effect on judging whether the network traffic is abnormal. Therefore, these irrelevant features should be removed. Given a network traffic data \( x \), the network traffic feature is used to \( f \) represent, and the network traffic feature set is used to \( S \) represent, then the expressions of features sets \( S \) and feature \( f \) are shown in the following:

\[
S = \{f_1, f_2, f_3, \cdots, f_k\},
\]

\[
f = \{x_1, x_2, x_3, \cdots, x_m\}.
\]
Here, \( f_b \) represents the \( b \) dimension feature in the network traffic dataset, \( x_{m} \) represents \( m \) pieces of network traffic data, and the feature set after removing irrelevant features is \( S_1 = \{ f_1, f_2, f_3, \ldots, f_d, d < b \} \).

4.2. Feature Selection of Maximum Information Correlation. According to the correlation between the feature and the label, select the \( t \) features that are most related to the label and delete the features that are not related to the label. Use the idea of an F-test to select the \( t \) features with the most correlated features and labels. The method is to calculate the F value of each feature, query the \( F \) distribution table, and select the \( t \) features that are most relevant to the type of network traffic attack.

There are \( k \) kinds of attacks on the network traffic data, the network data of the \( j \) attack type is \( n_j \), and the expression of the average value \( \bar{x} \) of the network traffic data is as follows:

\[
x = \frac{\sum_{i=1}^{d} \sum_{j=1}^{n_j} x_{ij}}{n}
\]

Here \( x_{ij} \) represents the \( i \) network traffic data of the \( j \) label type and \( n \) is the total number of network traffic data. The sum of the squares of the overall differences of all attack types of network traffic data \( \text{SST} \) is expressed as follows:

\[
\text{SST} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2.
\]

The sum of squares of differences within each attack type of network traffic data \( \text{SSE} \) is expressed as follows:

\[
\text{SSE} = \sum_{j=1}^{k} n_j (x_{ij} - \bar{x}_j)^2.
\]

Here, \( \bar{x}_j \) represents the mean value of the \( j \) class label, as follows:

\[
\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j}
\]

The sum of squares of differences between attack types of network traffic data \( \text{SSB} \) is expressed as follows:

\[
\text{SSB} = \text{SST} - \text{SSE} = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{x})^2.
\]

The expression of the correlation function \( F(t) \) between the feature of the network traffic data and the attack type is the ratio of the intraclass difference \( \text{MSB} \) and the interclass difference \( \text{MSE} \), and \( F(t) \) is expressed as follows:

\[
F(t) = \frac{\text{MSB}}{\text{MSE}}.
\]

The intraclass difference \( \text{MSB} \) is expressed as in (9), and the interclass difference \( \text{MSE} \) is expressed as in (10).

\[
\text{MSB} = \frac{\text{SSB}}{k-1}
\]

\[
\text{MSE} = \frac{\text{SSE}}{n-k}
\]

For a feature set \( S_1 \) of network traffic, calculate the F-value of each feature and attack type. Select the \( t \) features most relevant to the attack type according to the F-test distribution table. After feature selection with maximum information correlation, the selected feature subset is \( S_2 = \{ f_1, f_2, f_3, \ldots, f_t, t < d \} \).

4.3. Recursive Feature Elimination. Recursive feature elimination is used to select the optimal feature subset, random forest (RF) is used as the classifier, each feature \( X \) and phenotype \( Y \) are trained, and the classifier is used to evaluate the influence of feature \( X \) on phenotype \( Y \). Find the optimal dataset after multiple cross-validations. The optimal subset of features finally selected is \( S_3 = \{ f_1, f_2, f_3, \ldots, f_t, t < d \} \).

The RFC feature selection mechanism proposed in the paper includes three steps: Step 1, remove irrelevant features, remove those features whose feature columns are all zero or zero variance, and obtain feature subset \( S_1 \); Step 2, carry out maximum information correlation feature selection, according to the correlation between features and phenotype, select the \( t \) features with the best phenotype, and obtain the feature subset \( S_2 \); Step 3, use the recursive feature elimination method to select the optimal feature subset \( S_3 \).

4.4. SC-DeCN Network Anomaly Traffic Detection Model. The convolutional neural network feature extraction process focuses on the distribution of data information, which can effectively distinguish normal and abnormal data distribution and is widely used in network abnormal traffic detection. However, in the process of pooling, it will cause a loss of original information. In order to solve this problem, the paper proposes a SC-DeCN model for network abnormal traffic detection. The model mainly consists of three parts: (1) A CNN module is designed, which mainly includes convolutional layers and pooling layers, and the main task is to automatically extract network traffic features from the optimal feature subset after feature selection. (2) A DeCN module is designed to deconvolute the deep information of the CNN module to extract richer feature information. (3) Combined with the shallow information of the CNN module and the deep information of the DeCN module, the SC-DeCN model is constructed to detect network abnormal traffic and determine which kind of attack the network abnormal traffic belongs to. The model diagram is shown in Figure 2.

4.4. SC-DeCN Network Anomaly Traffic Detection Model. The convolutional neural network feature extraction process focuses on the distribution of data information, which can effectively distinguish normal and abnormal data distribution and is widely used in network abnormal traffic detection. However, in the process of pooling, it will cause a loss of original information. In order to solve this problem, the paper proposes a SC-DeCN model for network abnormal traffic detection. The model mainly consists of three parts: (1) A CNN module is designed, which mainly includes convolutional layers and pooling layers, and the main task is to automatically extract network traffic features from the optimal feature subset after feature selection. (2) A DeCN module is designed to deconvolute the deep information of the CNN module to extract richer feature information. (3) Combined with the shallow information of the CNN module and the deep information of the DeCN module, the SC-DeCN model is constructed to detect network abnormal traffic and determine which kind of attack the network abnormal traffic belongs to. The model diagram is shown in Figure 2.

(1) The 1st layer is the input layer, the input is network traffic data, and the input of the model is the 20-dimensional network traffic feature after feature selection. (2) The 2nd and 3rd layers are convolutional layers. In traditional neural networks, different neurons
between the two layers connect with each other, resulting in a large number of parameters to learn and bringing some difficulty in training. The convolutional layer has a local connection and weight sharing, which greatly reduces the number of parameters. The expression of convolution layer \( a_i \) is shown in the following equation:

\[
a_i = g(a_{i-1} \ast w_i + b_i)
\]

(11)

Among them, \( w_i \) is the weight of the convolution kernel of the \( i \) layer, \( a_{i-1} \) is the output of the convolution layer of the \( i-1 \) layer, \( b_i \) is the bias vector of the \( i \) layer, \( \ast \) is the convolution operation, and \( g \) is the activation function. Each layer in the convolution is top-down, and the expression of the shape feature \( o \) of the feature map is shown in the following:

\[
o = \left\lfloor \frac{c - e + 2p}{s} \right\rfloor + 1.
\]

(12)

The input feature map of the previous layer is \( c \), the size of the convolution kernel is \( e \), the stride is \( s \), and the number of rows and columns filled with 0 in the feature shape is \( p \).

(3) The 4th layer is the pooling layer, and the pooling window of the pooling layer is 2. The pooling layer is equivalent to further sampling the features and also reduces the dimension of the features. The expression of the pooling layer \( h_j \) is shown in the following equation:

\[
h_j = \beta_j \text{down}(h_{j-1}) + b_j.
\]

(13)

Here, \( \beta_j \) represents the activation value, “down” represents downsampling, \( h_{j-1} \) is the output of the \( j-1 \) layer of the pooling layer, and \( b_j \) represents the output of the \( j \) layer.

(4) The 5th, 6th, and 7th layers are deeper convolutional and pooling layers.

(5) Layers 8, 9, and 10 are deconvolution layers. Unlike the convolutional layer, its network signals are down-top and automatically use the deconvolution network to extract image high-level features, which can generally reflect the essence of the sample more than the original dataset. The expression of the shape feature \( o_2 \) of the feature map in the deconvolution network is shown in the following equation:

\[
o_2 = \left[ \frac{c_2 + 2p_2 - e_2 - (e_2 - 1) \ast (l - 1)}{s_2} \right] + 1.
\]

(14)

The feature shape of the input feature map of the previous layer is \( c_2 \), the size of the convolution kernel is \( e_2 \), the step size is \( s_2 \), \( l - 1 \) is the number of spaces inserted, and the number of rows and columns filled with 0 in the feature shape for \( p \).

(6) The 11th layer combines the information of layer 4 and layer 10 as the input of layer 12 (Flatten layer). Combine the shallow features of convolutional neural networks with the deep features of deconvolutional neural networks.

(7) Layers 13 and 14 are fully connected layers with softmax layers, and the expression is shown in the following equation:

\[
y = \text{softmax}(w \ast o_3 + b).
\]

(15)

Among them, softmax represents the activation function, \( w \) represents the weight of the fully connected layer, \( o_3 \) represents the feature map after connection through the skip level, \( b \) represents the bias of the fully connected layer, and the softmax layer classifies the network traffic data to determine whether the network traffic data belongs to what kind of attack.

Finally, the RIC feature selection mechanism in Section 3 and the SC-DeCN model in this section are combined to form the RIC-SC-DeCN network abnormal traffic detection algorithm in this paper.
5. Experimental Simulation Results and Analysis

The experimental operating system environment is Windows10, the computer hardware cup is i5-11260H, 16GB memory, and it is programmed in Python 3.7 software environment.

5.1. Introduction and Analysis of Dataset. The CIC-IDS-2017 dataset is proposed by the Communications Security Establishment (CSE) & the Canadian Institute for Cybersecurity (CIC), covering all 11 necessary criteria for common security network events. The dataset contains normal traffic and 14 common attacks, and each record has 78 features. Because CIC-IDS-2017 does not divide the training set and test set, 80% and 20% are used to divide it into a training set and test set in this experiment. In addition, this ratio has been used by many researchers recently.

5.2. Data Preprocessing. The data set CIC-IDS-2017 contains a large number of outliers and missing values, which will cause the data model to lose a lot of useful information and make the rules contained in the model more difficult to grasp. Therefore, these data need to be processed before analysis and modeling.

1. First, the outliers and missing values in the network traffic dataset are processed. Due to a large amount of data in the network traffic data set, the method adopted in this paper is to directly delete the data rows with null values and outliers.
2. Converts the text type in each connection in the raw data to numeric form.
3. Data standardization, the paper adopts the minimum and maximum normalization processing, which is a linear change of the original data, and maps the value to the [0, 1] interval. The conversion formula is in the following equation:

\[ x^* = \frac{x - \text{min}}{\text{max} - \text{min}} \]  

where max is the maximum value of the sample data, min is the minimum value of the sample data, and max – min is the extremely poor.

4. One-hot encoding is performed on the label, because label data is an attack type of network traffic, which is a discrete type, does not have seriality, and cannot directly compare the size. Using one-hot encoding can better represent the relationship between labels.

5.3. Model Parameter Settings. After many experiments to adjust the parameters, this experiment selects the following hyperparameters for model training:

1. **Batch.** If the value of the batch is too small, the fast calculation will not be possible. On the contrary, the weights cannot be updated well, and the experimental results show that a batch of 256 works best.
2. **Epoch.** If the epoch is set too large, the network training will overfit, resulting in lower test accuracy. The experimental results show that an epoch of 4 has the best effect.
3. **Dropout.** In this experiment, the dropout is 0.5. The experiment effect is the best.

5.4. Evaluation Indicators. This experiment uses AR (Accuracy Rate), RR (Recall Rate), FAR (False Alarm Rate), F1-Score, and Time as indicators for performance evaluation. AR is used to evaluate the overall performance of the system, RR represents the ratio of anomaly examples detected by network anomaly traffic detection models, FAR is the ratio of
misclassification, and F1-Score is a comprehensive indicator of precision and recall.

6. Experimental Results and Analysis

6.1. RIC Feature Selection Mechanism Experiment. In order to verify the effectiveness of the RIC feature selection mechanism proposed in this paper, the MSCNN model is used as the basic classification model, and the MIR mechanism, the RFE mechanism, the CHI mechanism, and the RIC mechanism proposed in this paper are used to select the features in the CIC-IDS-2017 dataset.

From Table 1, it can be concluded that the feature subsets selected by the four feature selection mechanisms are different. In the experiments in this section, the four feature subsets in Table 1 are used as input, and the MSCNN model is used for anomaly detection. The AR, RR, and F-Score of the above four feature selection mechanisms are shown in Figure 3, and the FAR is shown in Figure 4.

As can be seen from Figure 3, the feature subset selected by the RIC mechanism has the best performance in AR, RR, and F-Score on the MSCNN model compared to the MIR mechanism, the RFE mechanism, and the CHI mechanism. As can be seen from Figure 4, the FAR of the RIC mechanism on the CIC-IDS-2017 dataset is only 3.89%, which is 0.5% lower than that of the RFE mechanism, with the lowest FAR in the comparative experiment. The specific experimental data are shown in Table 2. In this table, select-time represents the time required for feature selection, Model-time represents the time required for model training, and Total-time represents the time required for feature selection and model training.

As can be seen from Table 2, compared with the MIR mechanism, the RIC mechanism improves AR, RR, and F-Score by 3.10%, 3.02%, and 3.11%, and the detection time increase is 49.71%. The RIC mechanism improves the AR, RR, and F-Score by 0.65%, 0.5%, and 0.61% compared with the RFE mechanism and shortens the time by 70.38%. The RIC mechanism improved AR, RR, and F-Score by 2.09%, 1.98%, and 2.09% compared with the CHI mechanism. The detection time increased by 46.93%.

Both the MIR mechanism and the CHI mechanism belong to the filtering feature selection method. The advantage of this method is that the computational cost of feature selection is short, and the time complexity is $T(n) = n$. The RFE mechanism should use the performance of the classifier as the evaluation criterion of the feature subset and traverse all possible combination feature subsets to select the optimal feature subset. The time complexity is $T(n) = n^2$, and the computational cost of feature selection is relatively large, not suitable for large data samples. The time complexity of the RIC mechanism in removing irrelevant features in the first step is $n$, the time complexity of the maximum information correlation in the second step is $n$, and the time complexity of the recursive feature elimination method in the third step is $(n/2)^2$. The overall time complexity is $(n/2)^2 + 2n$. On the CIC-IDS-2017 dataset, the feature selection computation time of the RIC mechanism has increased compared with the MIR mechanism and the CHI mechanism, but it has improved in AR, RR, and F-Score. Compared with the RFE mechanism, the RIC mechanism is greatly shortened in time.

In addition, due to the maximum information correlation feature selection in the RIC feature selection mechanism,
compared with the RFE mechanism, the correlation between features and labels can be considered, and the phenotype-related features can be better selected. In AR, RR, F-Score is also better. Therefore, compared with the MIR mechanism, the RFE mechanism, and the CHI mechanism, the RIC feature selection mechanism proposed in this paper can better select the optimal feature subset for network anomaly traffic detection.

6.2. Experiment of Network Anomaly Traffic Detection Based on SC-DeCN. In order to verify the effectiveness of the SC-DeCN model proposed in the paper in the detection of network abnormal traffic, experiments were carried out on the CIC-IDS-2017 dataset, using 78 original features after data preprocessing as the input of the model, and combined with the LSTM model, CNN-LSTM model, DSAE model, MSCNN model, and PCCN model are compared experimentally. The AR, RR, and F-Score of the above six network abnormal traffic detection models are shown in Figure 5, and the FAR is shown in Figure 6.

As can be seen from Figures 5 and 6, among the six classification models, the performance indicators of the LSTM model and the DSAE model are lower, so these two models perform poorly in the detection of network abnormal traffic. The variant of the CNN model performs better than other models, among which the MSCNN model performs worse than PCCN and SC-DeCN in all aspects, and SC-DeCN performs best overall. The specific data are shown in Table 3.

In Table 3, compared with the best classification model PCCN in the comparative experiment, the SC-DeCN model has improved by 1.14%, 0.86%, and 0.88% in AR, RR, and F-Score, respectively, and FAR decreased by 0.51%. Among the six classification models, LSTM has the worst classification effect, has not only lower AR but also has the longest model training time.

Although the DSAE model has a short model training time, the AR is too low. The variants of the CNN model, the MSCNN model, and the CNN-LSTM model perform relatively well. Among them, MSCNN has a short training time due to its simple model, but other performances are relatively poor compared to CNN-LSTM. The CNN-LSTM model outperforms the MSCNN model and the LSTM model because it combines the powerful feature extraction capabilities of the CNN model and the LSTM model has certain advantages in dealing with time series problems.

The PCCN model, as a parallel cross convolutional neural network, can extract richer feature information after fusing the branches of the two convolutional neural networks. Compared with the SC-DeCN model, the PCCN model does not perform deconvolution, and its effect in the detection of network abnormal traffic is worse than that of the SC-DeCN model. The experimental results show that the SC-DeCN model proposed in this paper has obvious advantages in the process of abnormal network traffic detection. It can be shown that the information loss problem caused by the pooling operation in the convolutional neural network can be effectively solved by combining the deconvolution model.

6.3. Network Anomaly Traffic Detection Based on RIC-SC-DeCN. In order to verify the effectiveness of the combination of the RIC feature selection mechanism proposed in the paper and the SC-DeCN model, the paper uses the SC-DeCN model to conduct comparative experiments, and it is verified on the public dataset CIC-IDS-2017. The experimental results are shown in Table 4.

It can be seen from Table 4 that compared with the SC-DeCN algorithm, RIC-SC-DeCN has an AR increase of 1.13%, RR increased by 1.27%, and F-Score increased by 1.44%. FAR decreased by 1.27% and shortened by 45.50% in time. The experimental results show that the RIC-SC-DeCN model proposed in this paper is better than the SC-DeCN model in the process of network abnormal traffic detection. It can greatly shorten the time used for model training and improve accuracy.

7. Conclusions

In the detection of network abnormal traffic, in view of the problem of noise characteristics in network traffic data and the problem of information loss caused by the pooling operation of convolutional neural networks used in existing research, this paper proposes a new method based on RIC-SC-DeCN network abnormal traffic detection algorithm. First, the important features in the network traffic are extracted through the proposed RIC feature selection mechanism, which reduces the impact of noise features and greatly shortens the training time; second, the extraction results are used as the input of the convolutional neural network, automatically learn the deep features of the data; then, combine the deconvolution network to reconstruct the original input signal, and fuse the deep features learned by the convolutional neural network to solve the information loss problem of the pooling operation in the convolutional neural network. The experimental results show that, compared with other classification algorithms, the method proposed in this paper has the best overall performance in terms of accuracy and efficiency.

The maximum information correlation feature selection method used in this paper is unstable in feature selection. In response to this problem, in future research, it can be
considered to add stability evaluation criteria to the correlation function, and at the same time, the number of feature subsets and classification error rate indicators are added to the evaluation indicators to form multiobjective evaluation criteria so as to enhance the stability of selecting feature subsets and improve the performance of the network abnormal traffic detection algorithm.

Data Availability
All the data used to support the findings of the study are included within the paper.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (No. 62002285).

References
[1] A. Mubashar, K. Asghar, A. R. Javed et al., “Storage and proximity management for centralized personal health records using an IPFS-based optimization algorithm,” Journal of Circuits, Systems, and Computers, vol. 31, no. 01, 2022.
[2] A. Naeem, A. R. Javed, M. Rizwan, S. Abbas, J. C.-W. Lin, and T. R. Gadekallu, “DARE-SEP: a hybrid approach of distance aware residual energy-efficient sef for WSN,” IEEE Transactions on Green Communications and Networking, vol. 5, no. 2, pp. 611–621, 2021.
[3] F. Iglesias and T. Zseby, “Analysis of network traffic features for anomaly detection,” Machine Learning, vol. 101, no. 1–3, pp. 59–84, 2014.
[4] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, “Network anomaly detection: methods, systems and tools,” IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 303–336, 2014.
[5] T. Y. Hu and S. Liu, “Research on intrusion detection technology based on chi-square test and LDOF algorithm,” Journal of Qi Lu University of Technology, vol. 33, no. 3, pp. 62–69, 2019.
[6] J. M. Liu and Y. B. Gao, “Adaptive bin feature selection for a fast network intrusion detection system,” Journal of Xidian University, vol. 48, no. 1, pp. 176–182, 2021.
[7] G. W. Wu, J. H. Liang, W. Wang, and C. S. Li, “Random Forest algorithm based on recursive feature elimination methods,” Statistics and Decision-making, vol. 21, pp. 60–63, 2017.
[8] M. D. Zeiler and R. Fergus, “Visualizing and Understanding Convolutional Neural Networks,” in Proceedings of the European Conference on Computer Vision (ECCV), pp. 818–833, Zurich, Switzerland, September 2014.
[9] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440, Boston, MA, USA, June 2015.
[10] M. X. Hang, W. Chen, and R. J. Zhang, “Abnormal traffic detection based on improved one-dimensional convolutional neural network,” Computer Applications, vol. 41, no. 2, pp. 433–440, 2021.
[11] E. H. Lv, X. S. Wang, and Y. H. Cheng, “Deep convolutional neural network learning based on deconvolution feature extraction,” Control and Decision, vol. 33, no. 3, pp. 447–454, 2018.
[12] Y. Zhu, J. Liang, J. Chen, and Z. Ming, “An improved NSGA-III algorithm for feature selection used in intrusion detection,” Knowledge-Based Systems, vol. 116, no. 15, pp. 74–85, 2017.
[13] F. Zhao, J. Zhao, X. Niu, S. Luo, and Y. Xin, “A filter feature selection algorithm based on mutual information for intrusion detection,” Applied Sciences, vol. 8, no. 9, Article ID 103390, 1535 pages, 2018.
[14] I. Sumayia Thaseen and C. Aswani Kumar, “Intrusion detection model using fusion of chi-square feature selection and multi class SVM,” Journal of King Saud University - Computer and Information Sciences, vol. 29, no. 4, pp. 462–472, 2017.
[15] J. Y. Ran, Research on Analysis Method of Network Traffic Abnormal Information, University of Electronic Science and Technology of China, China, 2020.
[16] G. Aceto, D. Ciouonzo, A. Montieri, and A. Pescapé, “Mobile encrypted traffic classification using deep learning: experimental evaluation, lessons learned, and challenges,” IEEE Transactions on Network and Service Management, vol. 16, no. 2, pp. 445–458, 2019.
[17] M. Erxue, L. Jun, L. Qiang, and I. D. S. Tr, “Anomaly-Based Intrusion Detection through Text-Convolutional Neural Network and Random Forest,” Security & Communication Networks, vol. 1, pp. 1–9, 2018.
[18] S. Basumallick, R. Ma, and S. Eftekharnejad, “Packet-data anomaly detection in PMU-based state estimator using convolutional neural network,” International Journal of Electrical Power & Energy Systems, vol. 107, no. 5, pp. 690–702, 2019.
[19] Z. Zhang, X. Zhou, X. Zhang, L. Wang, and P. Wang, “A model based on convolutional neural network for online transaction fraud detection,” Security and Communication Networks, vol. 2018, Article ID 5680264, 9 pages, 2018.
[20] Z. J. Gu, J. T. Hao, and J. X. Zhou, “Malicious network traffic classification algorithm based on an improved bilinear convolutional neural network,” Information Network Security, vol. 20, no. 10, pp. 67–74, 2020.
[21] G. F. Tian, Z. L. Shan, Z. H. Liao, and Y. L. Wang, “Network intrusion detection model based on Faster R-CNN deep learning,” Journal of Nanjing University of Technology (Natural Science Edition), vol. 45, no. 01, pp. 56–62, 2021.
[22] T.-Y. Kim and S.-B. Cho, “Web traffic anomaly detection using C-LSTM neural networks,” Expert Systems with Applications, vol. 106, pp. 66–76, 2018.
[23] C. Ding and H. Peng, “Minimum redundancy feature selection from microarray gene expression data,” Journal of Bioinformatics and Computational Biology, vol. 03, no. 02, pp. 185–205, 2005.
[24] A. R. Javed, M. Usman, S. U. Rehman, M. U. Khan, and M. S. Haghhighi, “Anomaly detection in automated vehicles using multistage attention-based convolutional neural network,” IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4291–4300, 2021.
[25] F. A. Khan, A. Gumaei, A. Derhab, and A. Hussain, “TSDL: a two-stage deep learning model for efficient network intrusion detection,” IEEE Access, vol. 7, Article ID 30373, 2019.