Computer Science and Information Systems 20(3):997–1014 https://doi.org/10.2298/CSIS221124004L

RESNETCNN:an Abnormal Network Traffic Flows Detection Model

Yimin Li[1,⋆], Dezhi Han[1], Mingming Cui[1], Fan Yuan[2], and Yachao Zhou[2]

[1] College of Information Engineering,
Shanghai Maritime University, Shanghai201306, China
{202130310117, mmcui}@stu.shmtu.edu.cn
dzhan@shmtu.edu.cn
[2] Hangzhou Anheng Information Technology Co.,
LTD, Hangzhou310051, China
{frank.fan, anna.zhou}@dbappsecurity.com.cn

Abstract. Intrusion detection is an important means to protect system security by detecting intrusions or intrusion attempts on the system through operational behaviors, security logs, and data audit. However, existing intrusion detection systems suffer from incomplete data feature extraction and low classification accuracy, which affects the intrusion detection effect. To this end, this paper proposes an intrusion detection model that fuses residual network (RESNET) and parallel cross-convolutional neural network, called RESNETCCN. RESNETCNN can efficiently learn various data stream features through the fusion of deep learning and convolutional neural network (CNN), which improves the detection accuracy of abnormal data streams in unbalanced data streams, moreover, the oversampling method into the data preprocessing, to extract multiple types of unbalanced data stream features at the same time, effectively solving the problems of incomplete data feature extraction and low classification accuracy of unbalanced data streams. Finally, three improved versions of RESNETCNN networks are designed to meet the requirements of different traffic data processing, and the highest detection accuracy reaches 99.98% on the CICIDS 2017 dataset and 99.90% on the ISCXIDS 2012 dataset.

Keywords: Intrusion detection, RESNETCNN, Deep learning.

1. Introduction

With the digitalization of the Internet [13], cyberspace security issues have become increasingly complex and diverse. On June 22, 2022, Northwestern Polytechnical University released a statement about a cyber attack in which multiple Trojan samples were found in the university's information network. In recent years, tens of thousands of malicious cyber attacks on domestic Chinese network targets, resulting in network devices have been controlled and high-value data have been stolen. Therefore, cybersecurity protection is extremely imminent.Intrusion detection is the detection of intrusions or intrusion attempts on a system through operational behavior, security logs, audit data, or other available network information, etc. When running on the inspected system, the Intrusion Detection System (IDS) [31] is responsible for scanning the current network activity of

⋆ Corresponding author

the system, monitoring and recording the network traffic of the system. Then IDS detects, records and judges the legitimacy of various processes running on the system, filtering abnormal network traffic from the host according to defined rules, and finally provides real-time alerts,which can effectively guarantee the network security.

Existing intrusion detection systems may suffer from overfitting, low classification accuracy, and high false alarm rate (FPR) when facing [32] with huge and diverse network data.Jun ho Bang et al.[1] used HsMM to model WSAN and developed an LTE signaling attack detection scheme. Compared with other detection schemes, the attack detection effect was better, but the accuracy was lower. M. Nazari et al.[26] proposed an ARIMA time series model and a new DoS and DDoS attack detection algorithm, which improved the classification performance of abnormal traffic, but had a high false positive rate.

Machine learning [42] has been widely used to identify various types of cyber attacks. Yang et al.[38] proposed an abnormal network traffic detection algorithm that integrates mixed information entropy and SVM to detect abnormal network traffic in cloud computing environment. K. Li [15] combined principal component analysis (PCA) and random forest (RF) algorithm to extract and combine features from different network layers, reducing redundancy and noise caused by multi-layer combination.However, most traditional machine learning methods are shallow learning, that usually emphasizes feature engineering and feature selection, which cannot solve classification problems for large-scale data in real network environments. And the classification accuracy of multiple classification tasks decreases with the dynamic growth of the dataset. Thus shallow learning is not suitable for intelligent analysis and high-dimensional massive data learning.

In the face of network traffic data with large scale and high dimensions, deep learning has greater advantages. Zhang [40] proposed an intrusion detection model based on deep hierarchical network, which combined CNN and LSTM (CNN_lSTM for short) and achieved good performance on CICIDS2017 data set. Zhong [41] proposed HELAD, a network abnormal traffic detection algorithm integrating multiple deep learning technologies. Although HELAD has better adaptability and detection accuracy, its bit error rate is slightly higher. At present, the intrusion detection system based on deep learning has the following two problems: 1. With the further increase of network traffic, the current intrusion detection system with high-speed detection capability is not very ideal in terms of packet capture capability and detection performance. 2. Data imbalance in real environment seriously affects the detection accuracy of most current intrusion detection systems.

To alleviate the above problems, an intrusion detection model, referred to as RESNETCNN is proposed that fuses RESNET and a parallel cross-convolutional neural network. The main contributions of this paper are as follows.

(1) By introducing the oversampling method to process the ISCXIDS 2012 dataset, we can extract multiple types of unbalanced data stream features simultaneously, which effectively solves the problems of incomplete data feature extraction and low classification accuracy in unbalanced dataset.

(2) The top layer of the proposed model adopts RESNET network structure and the bottom layer of the network adopts traditional convolutional neural network (CNN) structure. The combination of the top and bottom layers can effectively perform feature fusion and improve the detection accuracy of abnormal data streams in unbalanced datasets.

(3) Three improved versions of RESNETCNN are proposed. Simulation experiments

are conducted on the CICIDS 2017 and ISCXIDS 2012 datasets, and the corresponding detection accuracy can reach 99.98% and 99.90%, respectively.

The rest of the paper is organized as follows. The second part describes the evolution of the imbalanced dataset processing method, RESNET network model, and parallel cross-convolutional neural network. The third part describes the data processing method and the network structure of RESNETCNN, afterwards, ablation experiments are conducted on the CICIDS 2017 and ISCXIDS 2012 datasets to evaluate the effectiveness of the proposed model in the fourth part. Finally, the fifth part concludes the paper.

2. Related Work

This section discusses the related work from three aspects, namely, unbalanced dataset processing methods, RESNET network models, and the evolution of parallel cross-convolutional neural networks.

2.1. Evolution of Methods for Processing Unbalanced Data Sets

The number of samples of each category in real intrusion detection datasets is often unbalanced, and the methods to deal with this problem mainly include feature selection-based, and resampling-based methods. The feature selection-based approach [35] first identifies highly relevant features from the source and target domains, then removes irrelevant or redundant features, and finally applies the highly relevant features to the target problem.This method has two problems: first, it does not consider the correlation between features and data; second, the selected feature data usually tend to be in the majority class with large amount of data. The resampling-based methods change the data distribution through specific data sampling to equalize the data distribution of different classes in the dataset. The common methods are divided into two categories: random downsampling and random oversampling. Random oversampling is performed by randomly sampling a small number of samples in the dataset, and then adding the sampled samples to the original dataset. Random oversampling can restore the data of the real scene and improve the detection rate of the small number of data in the unbalanced abnormal traffic.

2.2. Evolution of the RESNET Model

ImageNet classification with deep convolutional neural network(Alexnet) [14] unveiled the dominance of neural networks in computer vision by incorporating a variant of the neural network model in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. Visual geometry group(VGG) [27] investigated the effect of convolutional network's depth on the accuracy of large-scale image recognition, where (3×3) Convolutional filter architecture was used to evaluate the depth of the network. And the results showed that pushing the network's depth to 16-19 weight layers achieves a significant improvement over existing technology configurations.Going deeper with convolutions (InceptionV1) [29] is a 22-layer deep network, which won the first place in the 2014 ILSVRC challenge. It increases the depth and width of the network while keeping the computational budget the same. Its disadvantage is that the network is difficult to change, and the cost will increase four times if the number of filter sets is doubled.

Batch normalization: accelerating deep network training by reducing internal covariate shift(InceptionV2) [12] improves InceptionV1 with a Batch Normalization(BN) layer, achieving a top-5:4.9% result in the ILSVRC competition, but also increases the weight consumption by 25% and the computational consumption by 30%.Rethinking the inception architecture for computer vision (InceptionV3) [30] improves InceptionV2 by reducing the initial module and parallelizing the network structure. RESNET [10] improves InceptionV3 by introducing a residual structure that smoothly propagates the gradient from backward to forward, which extends the RESNET structure to thousands of layers and further alleviates the gradient disappearance problem in deep neural networks.

2.3. Evolution of Parallel Cross Convolutional Neural Networks

There is an imbalance of different categories of data in intrusion detection dataset. In order to improve the detection accuracy of few categories of data, many researchers have proposed parallel cross-convolutional neural network. It draws temporal, spatial, and semantic features of anomalous traffic dataset through feature fusion of two or three layers of different convolutional neural networks, greatly improving the classification accuracy of anomalous traffic dataset.

[39] proposed parallel cross-convolutional neural network based on deep hierarchical network [40],called PCCN. This network model fused fully convolutional networks for semantic segmentation(FCN) [25] and convolutional neural network (CNN). Although its overall accuracy reached 0.9991 at CICIDS 2017, there are problems of excessive amount of parameters, long training time ,and low classification accuracy. An anomaly network traffic detection model integrating temporal and spatial features(ITSN) [22] and an efficient hybrid parallel deep learning model(HPM) [4] are both parallel deep learning models that fuse long short-term memory(LSTM) [11] and CNN,which have the advantage of introducing LSTM to extract spatial features.Compared with PCCN, the classification accuracy of minority classes is improved, but the classification accuracy of GlodenEye, Hulk, slowhttp and slowloris in CICIDS 2017 dataset is lower.A network abnormal detection method(PCCS) [34] combines single-stage headless face detector algorithms(SSH) and parallel crossover neural network, and consists of two parallel convolutional network layers. It has the advantage that the overall accuracy is improved compared to ITSN, reaching 0.9996, but the classification accuracy is lower on GlodenEye, Hulk, slowhttp, and slowloris in the CICIDS 2017 dataset.

RESNETCNN adopts a parallel crossover network structure with RESNET at the top and CNN at the bottom. RESNET absorbs semantic features and CNN absorbs high-resolution features.After three stages of feature fusion, the overall accuracy reaches 99.96% at CICIDS 2017, and the classification accuracy of GlodenEye, Hulk, slowhttp, and slowloris in the dataset is greatly improved.Compared with the traditional work, RESNETCNN has fewer parameters, high classification accuracy and fast speed, which is suitable for the classification of large anomalous traffic datasets.

3. Data Pre-processing

This section mainly discusses the algorithm of data preprocessing and verifies the validity of random oversampling method.

The ISCXIDS 2012 dataset suffers from data imbalance in different categories, and the classification accuracy of different abnormal traffic categories varies greatly. Thus, a random oversampling method is used with the following algorithm.

First, the data classified by the confusion Matrix is marked as X_{ij} (i=1 . . . m, j=1 $\cdots$ n).

Step1 calculates the ratio of the number of incorrect classifications to the number of correct classifications for each category in the confusion Matrix (misclassification ratio).

$$ratio = \frac{\sum_{j=1}^{n} X_{ij}(j \neq i)}{X_{ii}} \tag{1}$$

Observe whether the proportion of misclassified data in most minority classes is less than the proportion of misclassified data in most classes, and if so, proceed to step2.

Step2 random oversampling, that is, the minority class data in the dataset is amplified according to the maximum data volume of the majority class.

(1) Integrates all categories of data in anomalous traffic datasets.

(2) Classifies large data sets into m classes based on labeled features and store them in a collection of lists.

(3) lists stores the m-class data set, expands the data amount of minority class to the maximum data amount Max of majority class, and then stores it in the excell table.

The pseudo code for the oversampling algorithm is shown in Algorithm1. Description of the argument in Algorithm 1 is shown in table 1.

Algorithm 1 Resampling of a Small Number of Data Sets

Input: all data[];
Output: Expanded data;

```
temp_list=[]; # Store the final processed data
label_list=[]; # Store raw data label set
MAX;
length;# The length of raw data
for i in range(0, m) do
        #Generate m empty list sets to temporarily store the generated m feature sets
        temp_list.append([]);
end for
for i in range(0, length) do
        #Get the index number of the tag
        data_index = label_list.index(all data[i]);
        temp_lists[data_index].append(all data[i]);
end for
#Expand the small number of data sets to at least Max, and then store them.
for i in range(0, temp_lists.length) do
    while len(temp_lists[i]) < MAX do
                temp_list[i].expend(temp_list[i])
    end while
end for
#Save the descended data to the specified csv file
```

Table 1. Description of the argument in Algorithm 1

Argument	Description
all Data[]	Enter the csv file data in the ISCXIDS 2012 data set
temp_list[]	Generate m empty list sets to temporarily store the generated m feature sets
Label_list	Raw data label set
Max	Maximum data size for most classes
Length	The number of rows of all Data[]

We found that in the intrusion detection dataset ISCXIDS 2012, the misclassification ratio of most minority data is lower than that of most majority data, namely, the dataset is unbalanced. So we conducted random oversampling on ISCXIDS 2012. After random oversampling, we trained on the RESNETCNN3 model, and the accuracy increased by 0.0001. The results are shown in Fig. 1.

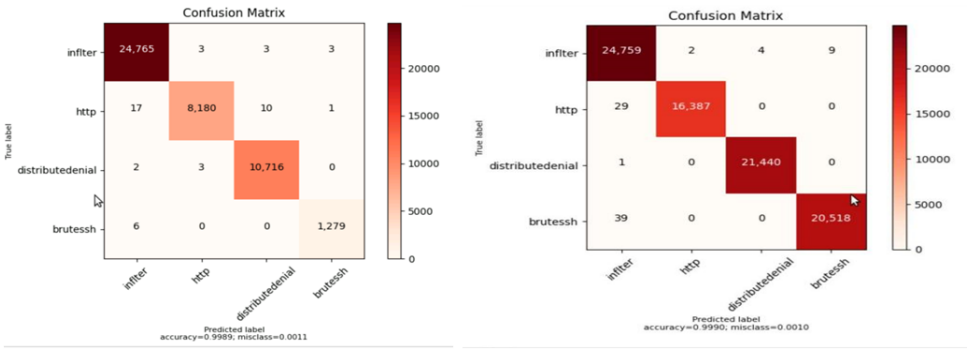

Fig. 1. (1) Data Set before Random Oversampling (2)Data Set after Random Oversampling

4. System Model

This section details the network structures of RESNETCNN and three improved versions of RESNETCNN.

4.1. Network Structure of RESNETCNN

As shown in Fig. 2,the basic idea of RESNETCNN is mainly inspired by the residual structure RESNET, and the model is divided into two layers, top branch and bottom branch. Top branch adopts the first three layers of RESNET version 18, while bottom branch adopts the traditional CNN structure, which consists of three convolutional pooling layers. In order to improve the detection accuracy of a few classes, top branch and

bottom branch are fused into three stages: the first two stages are add feature fusion and the third stage is concatenateenate feature fusion.

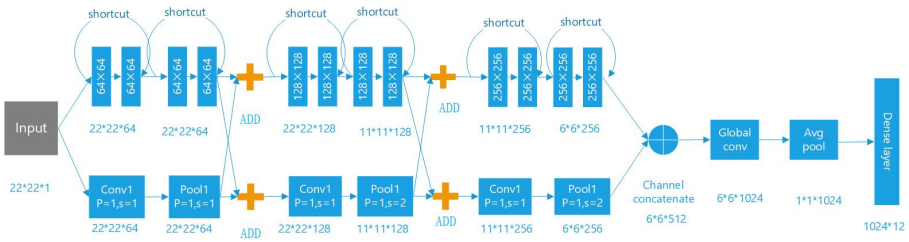

Fig. 2. Network Structure of RESNETCNN

1) TOP BRANCH

RESNET structure is applied to RESNETCNN upper layer. To solve the problem of network degradation, Kai-Ming He proposed the residual learning. Generally speaking, the deeper the network layer, the stronger the network fitting ability, and the smaller the network training error should be. But in fact, the deeper the network layer, the easier it is to overfit, and the network optimization is more difficult.RESNET takes a cue from LSTM, introduces a gating unit, and adds a computational path, shortcut mapping, to the traditional forward propagation.Shortcut mapping is beneficial to gradient propagation.

$$\frac{\partial H}{\partial X} = \frac{\partial F}{\partial X} + \frac{\partial X}{\partial X} = \frac{\partial F}{\partial X} + 1 \tag{2}$$

The constant mapping of Eq. (2) allows the gradient to propagate unimpeded from back to front, which enables the RESNET structure to expand to thousands of layers (1202 layers).

As shown in Fig. 3,the residual structure is stacked in two ways, Basic block and Bottleneck block. Basic block uses two 3×3 convolutional stacking mode. First input parameter x, then perform 3×3 convolution, relu function, and 3×3 convolution to get F(x), finally calculate H(x)=F(x)+x. Fitting F(x) makes the convolutional block easier to learn constant mapping. When F(x)=0, H(x)=x,in which case the precision of the deep network is higher than that of the shallow network and the network achieves constant mapping.

$$F(x) = W_2 \times Relu(W_1 \times x) \tag{3}$$

$$H(x) = F(x) + x = W_2 \times Relu(W_1 \times x) + x \tag{4}$$

Unlike the former,the number of channels in multiple network layers is like a bottleneck. The input channels change from large to small, and then from small to large.Bottleneck block uses 1×1 convolution kernel. The first 1×1 decreases the number of channels by 1/4 and the second 1×1 increases the number of channels by 4 times, which is more conducive to extracting advanced features by first descending and then ascending, and at the same time reduces computation and saves training time.

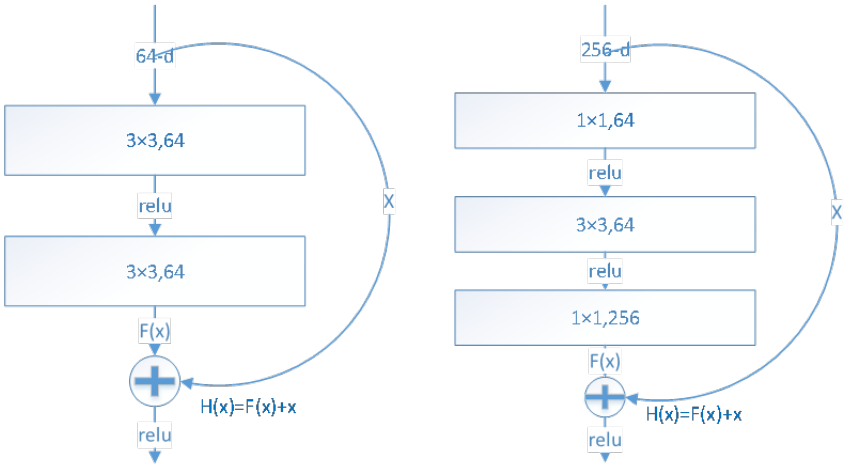

Fig. 3. Basic Block and Bottleneck Block

The upper layer is structured with RESNET18, which consists of 6 Basic blocks with 3×3 convolutional kernels, to learn the semantic features of unbalanced anomalous traffic.

2) BOTTOM BRANCH

The lower layer of RESNETCNN adopts CNN structure.CNN is a mathematical model that imitates the structure and function of biological neural network. CNN can be divided into three categories: convolutional layer, pooling layer, and fully-connected layer. The main role of convolutional layer and pooling layer are feature extraction and downsampling respectively. The pooling layer retains the most significant features and discards other useless information to reduce the operation. The introduction of pooling layer also ensures the translation invariance, i.e., the same image after flipping and deforming can get similar results. The fully-connected layer is mainly responsible for classifying the features derived from the previous convolution and pooling layers. After each neuron feedback a different weight, it will adjust the weight and network to get the classification results.

The lower layer of RESNETCNN consists of three sets of convolutional pooling layers stacked, which are used to extract high-level traffic features of unbalanced a.nomalous traffic.

3) FEATURE CROSS FUSION

Feature fusion is an important tool to improve classification performance. Low-level features have higher resolution, and contain more location and detail information. But they are less semantic and more noisy because they go through fewer convolutional layers. High layer features have stronger semantic information, but have very low resolution and poor perception of details. How to fuse the two is the key to improve the classification model. According to the order of fusion and prediction, they are divided into early fusion and late fusion. The former first fuses the features of multiple layers first, and then trains the predictor on the fused features. Two classical fusion methods are 1) concatenate: series

feature fusion, where two features are directly connected. If the dimension of two input features x and y are p and q, and the dimension of output feature z is p+q; 2) add:parallel strategy fusion, combining these two feature vectors into a complex vector. After features x and y are input,$z = x + i \times y$ will be output,where i is an imaginary unit. Late fusion improves detection performance by combining detection results from different layers in two ways: 1) multi-scale features are first predicted separately, and then the results are fused. 2) multi-scale features are first pyramid fused, and then the results are predicted. RESNETCNN uses concatenate and add of the early fusion.

In the first stage, the upper layer passes through two Basic blocks with 64 channels and 3×3 convolutional kernels in RESNET18, and the lower layer passes through two convolutional pooling layers with 64 channels, 3×3 convolutional kernels, padding=1 and stride=1 in CNN. Add fusion is used for feature fusion in the first and second stages. Compared with the first stage, in the second stage, the stride of the second block of the upper layer is changed to 2, the stride of the second pooling layer of the lower layer is changed to 2, and the size of the convolutional kernel is reduced by 1/2. The purpose of this approach is to extract richer semantic information through downsampling. The third stage of feature fusion uses concatenate fusion, where the number of channels is expanded twice. It realizes the complementary advantages between different features, and is more conducive to fully learning the intrinsic features of traffic data, thus weakening the impact of data imbalance on traffic data feature learning. In the fourth stage, after full convolution, average pooling and fc layer, the information extracted from the feature layer is classified, and the final accuracy reaches 99.96% on the CICIDS 2017 dataset.

4.2. Three Improved Versions of RESNETCNN

Version 1 is RESNETCNN1 (MINIRESET50 Cross Convolutional Neural Network), as shown in Fig. 4, which contains top branch and bottom branch. Top branch is the RESNET50 structure, consisting of six Bottleneck blocks, and bottom branch is the CNN structure, consisting of six convolutional and pooling blocks. The output of the two branches is fused with concatenate channels, and finally the classification features are output through the fully connected layer. Version 2 is RESNETCNN2 (RESNET50 Cross Convolutional Neural Network),as shown in Fig. 5. The top branch adopts RESNET50 structure and the bottom branch is CNN structure. Different from version 1, Version 2 performs feature fusion in three stages. The top two Bottleneck block and the lower two CNN convolutional pooling layers in the first and second stages perform ADD fusion and extract semantic features of the data. In the third stage, two Bottleneck blocks in the upper layer and two CNN convolutional pooling layers in the lower layer perform concatenate channel fusion to improve the detection performance of few classes of data in imbalanced datasets. Version 3 is RESNETCNN3 (RESNET Cross Convolutional Neural Network),as shown in Fig. 6,which differs from versions 1 and 2 in that (1) the top branch adopts the structure of RESNEXT [37]. RESNEXT introduces a grouped convolution with a grouping number of 32, where each block is divided into 32 groups and the convolutional kernel size is 4; (2) feature fusion is performed in two stages, each stage the upper three Bottleneck blocks and the lower two CNN convolutional pooling blocks are fused. The upper and lower layers of the first stage are add fused to extract low-level traffic features. In the second stage, the upper and lower layers perform concatenate fusion to extract high-level traffic features, and finally the final result is obtained through the classification layer.

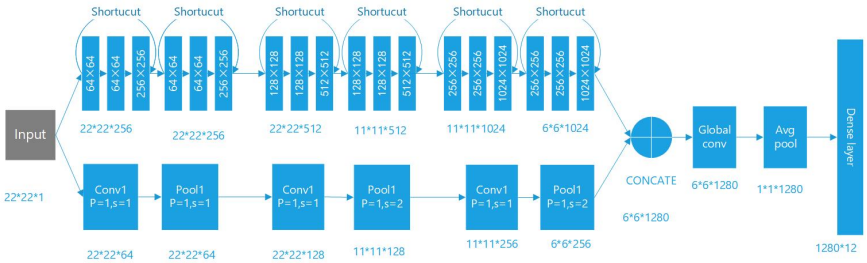

Fig. 4. Network Structure of RESNETCNN1

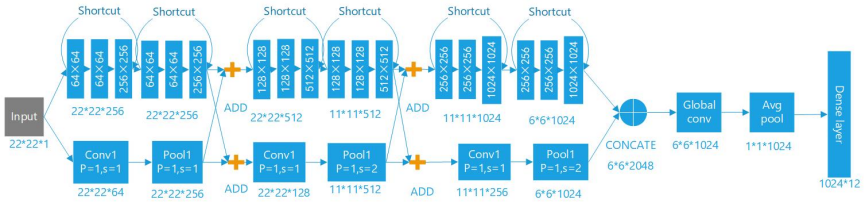

Fig. 5. Network Structure of RESNETCNN2

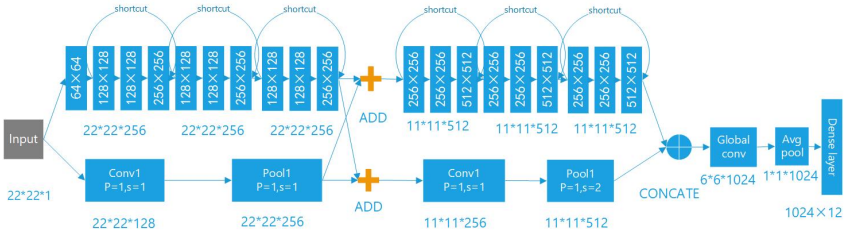

Fig. 6. Network Structure of RESNETCNN3

5. Experimental Results and Analysis

This section first describes the environment required for the experiments, including the hardware environment, the software environment, and the datasets. In addition, the evaluation metrics used in the experiments and the configuration of parameters during model training are presented. In the last part of this section, the content of the experiment is introduced in detail and the experimental results are analyzed.

5.1. Experimental Environment

The experimental environment is shown in Table 2.

Table 2. Experiment Environment

Equipment	Example
CPU	11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz
GPU	Quadro P4000
Memory	64GB
Hard Disk	2T
OS	Ubuntu 16.04
Compile software	PyCharm 2021.2.3
Python	3.8
Database	ISCXIDS 2012,CICIDS2017

5.2. Hyperparameter Setting

In RESNETCNN, we fix the size of the convolutional kernel to 3×3. The model uses a batch size of 256 during training, where the momentum is fixed at 0.9 and the weight decay is set to 1×10-4 to prevent overfitting and the model falling into local optima. Also, we use a cross-entropy loss function to continuously optimize the model parameters and an Adam optimizer to accelerate the convergence of the network. Setting the learning rate too large or too small can affect the convergence of the model and cause the model to miss the optimal point. Therefore, a total of 10 iterations are designed in this paper. The learning rate settings for each iteration are shown in Table 3.

Table 3. Learning Rate Setting

Epoch	0	1	2	3	4	5	6	7	8	9
Learning Rate	0.0001	0.0001	0.0001	0.0001	0.0001	0.00002	0.00002	0.00002	0.000004	0.000004

5.3. Evaluation Indicators

The horizontal axis in the confusion matrix [33] is a count of the number of categories predicted by the model, and the vertical axis is a count of the number of true labels of the data.The diagonal line, represents the number of model predictions that agree with the data labels, so the sum of the diagonals divided by the total number of test sets is the accuracy rate. The larger the number on the diagonal, the better, and the darker the color in the visualization results, indicating the more accurate the model's prediction in that category.

True Positive (True, TP): predicts the positive class as the number of positive classes, True Negative (True Negative, TN): predicts the negative class as the number of negative classes, False Positive (False Positive, FP): predicts the negative class as the number of positive classes, and False Negative (False Negative, FN): predicts the positive class as the number of negative classes.

1. Precision

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

The fraction of predicted positive cases that are determined to be positive, as a percentage of all predicted positive cases.

2. Recall

$$Recallrate = \frac{TP}{TP + FN} \tag{6}$$

The fraction of cases predicted to be positive and that are indeed positive, as a percentage of all classes that are indeed positive.

3. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

All positive and negative cases with correct predictions, as a percentage of all samples.

4. F1_Score

$$F1\_score = 2 \times \left(\frac{Precision \times Recall}{precision + Recall} \right) \tag{8}$$

F1_Score, also known as the balanced F_score, is the summed average of Precision and Recall, which combines the results of Precision and Recall. F1_Score ranges from 0 to 1, with 1 and 0 representing the best and the worst output of the model respectively. The two metrics, Precision and Recall, are commonly used to evaluate the analytical effectiveness of two classification models. However, when these two metrics are in conflict, it is difficult to compare between models. For example, we have two models A and B. Model A has a higher recall than model B, but model B has a higher precision than model A. This is where F1_Score is used to judge the overall performance of the two models.

5.4. Experimental Analysis

To further explore the detection performance of RESNETCNN for unbalanced data streams in complex environments, we conducted ablation experiments on the CICIDS 2017 dataset. Fig. 7 and Table4-8 show the experimental results, where the data of PCSS are cited from

Tables 3, 4, 5, 6, and 7 in PCSS . From Fig. 7, it can be seen that the RESNETCNN series network model proposed in this paper outperforms the PCCN, PCSS, and ITSN models in the four evaluation indexes of average recall, average F1_score, average precision, and average accuracy. It shows that the PCSS and RESNETCNN are the same in average accuracy, but the average precision of RESNETCNN is lower than that of PCSS.In comparison, the three improved versions of RESNETCNN in the experiments are better than PCSS.Compared with RESNETCNN2 and RESNETCNN3 , RESNETCNN1 shows an increasing trend in average accuracy, average precision, and average F1_score, but the average recall of RESNETCNN1 is better than that of PCSS. RESNETCNN1 outperforms the other two models in average recall, indicating that multiple feature fusion will reduce the callback rate while improving the accuracy.

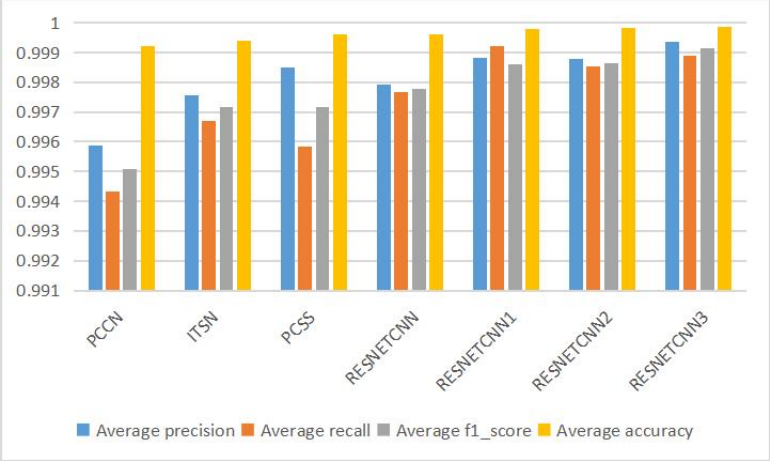

Fig. 7. Experimental Comparison of Various Abnormal Traffic Detection Algorithms

In Table 4, RESNETCNN3 has the highest overall accuracy, reaching a classification accuracy of 0.999876. Meanwhile, it can be seen that the overall accuracies of RESNETCNN, RESNETCNN1, RENETCNN2, and RENETCNN3 show a stepwise trend of increasing, indicating that the three improved versions proposed in this paper in the network model with RESNET50 to replace RESNET18 is effective.Tables 5 to 8 show the classification precision, F1_score, recall, and False-positive rate of 12 types of anomalous traffic data in the CICIDS 2017 dataset for PCCN, ITSN , PCCS, RESNETCNN, RESNETCNN1 , RESNETCNN2, and RESNETCNN3 on eight network models. In Table 6 and 7, except for the third category of anomalous traffic with recall and f1_score that reaches the optimum on PCSS, the other categories of anomalous traffic all achieve the optimum on RESNETCNN3. In Table 5 and 8, RESNETCNN3 achieves the optimum on precision, F1_score, recall, and False_positive rate.In order to prevent overfitting, the RESNETCNN3 model is validated on the ISCXIDS 2012 dataset, with an overall accuracy of 99.9%.

Table 4. Accuracy Compare

Label	PCCN	ITSN	PCSS	RESNETCNN	RESNETCNN1	RESNETCNN2	RESNETCNN3
Accuracy	0.9992	0.9995	0.9996	0.9996	0.9997	0.99982	**0.99987**

Table 5. Precision Compare

Label	0	1	2	3	4	5	6	7	8	9	10	11
PCCN	1	1	0.9955	0.9989	0.9759	0.9928	1	1	1	0.9999	1	0.9981
ITSN	1	1	0.9946	0.9996	0.9803	0.9943	1	1	1	0.9999	0.9998	0.9986
PCSS	1	0.9998	0.9985	0.9997	0.9948	0.9981	0.9992	0.9995	1	1	0.9976	0.9948
RESNETCNN	1	1	0.9949	0.9997	0.9897	0.9934	1	1	1	1	0.9998	0.9976
RESNETCNN1	1	1	0.9985	0.9998	0.9941	0.9948	1	1	1	0.9999	1	0.9986
RESNETCNN2	1	1	0.9981	0.9999	0.9933	0.9957	1	1	1	1	0.9995	0.9991
RESNETCNN3	**1**	**1**	**0.999**	**0.9999**	**0.9978**	**0.9967**	**1**	**1**	**1**	**1**	**1**	**0.9991**

Table 6. F1_score Compare

Label	0	1	2	3	4	5	6	7	8	9	10	11
PCCN	1	1	0.9841	0.9994	0.9798	0.9886	1	1	1	0.9999	0.9998	0.9988
ITSN	1	1	0.9919	0.9997	0.9846	0.9912	1	1	1	0.9999	0.9999	0.9993
PCSS	0.9927	0.9998	**0.9984**	0.9998	0.993	0.9969	0.9991	0.9972	0.9943	1	0.9987	0.9959
RESNETCNN	1	1	0.9931	0.9997	0.9889	0.9934	1	1	1	0.9999	0.9997	0.9988
RESNETCNN1	1	1	0.997	0.9999	0.9911	0.996	1	1	1	0.9999	0.9999	0.9993
RESNETCNN2	1	1	0.9977	0.9999	0.9911	0.996	1	1	1	1	0.9996	0.9995
RESNETCNN3	**1**	**1**	0.9982	**0.9999**	**0.9945**	**0.9976**	**1**	**1**	**1**	**1**	**0.9999**	**0.9995**

Table 7. Recall Compare

Label	0	1	2	3	4	5	6	7	8	9	10	11
PCCN	1	1	0.973	0.9999	0.9838	0.9843	1	1	1	1	0.9996	0.9995
ITSN	1	1	0.9893	0.9998	0.989	0.9881	1	1	1	0.9999	1	1
PCSS	0.9855	0.9999	**0.9983**	0.9999	0.9912	0.9957	0.999	0.9949	0.9887	1	0.9998	0.9971
RESNETCNN	1	1	0.9912	0.9998	0.9882	0.9934	1	1	1	0.9999	0.9996	1
RESNETCNN1	1	1	0.9954	1	0.9982	0.9972	1	1	1	1	0.9998	1
RESNETCNN2	1	1	0.9973	1	0.989	0.9962	1	1	1	1	0.9998	1
RESNETCNN3	**1**	**1**	0.9973	**1**	**0.9912**	**0.9986**	**1**	**1**	**1**	**1**	**0.9998**	**1**

Table 8. False Positive Rate Compare

Label	0	1	2	3	4	5	6	7	8	9	10	11
PCCN	0	0	0.0001	0.0008	0.0001	0.0001	0	0	0	0	0	0
ITSN	0	0	0.0001	0.0003	0.0001	0.0001	0	0	0	0	0	0
PCSS	0	0.0001	0	0.0002	0	0	0	0	0	0	0.0001	0
RESNETCNN	0	0	0.0001	0.0002	0.0001	0.0001	0	0	0	0	0	0
RESNETCNN1	0	0	0	0.0001	0	0	0	0	0	0	0	0
RESNETCNN2	0	0	0	0.0001	0	0	0	0	0	0	0	0
RESNETCNN3	**0**	**0**	**0**	**0.0001**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

To understand the errors in the experimental results of the RESNETCNN family of models, the heat map shown in Figure 8 is generated according to the results of the experiments performed on the RESNETCNN3 model. The diagonal numbers represent the number of correct predictions and the remaining numbers are the number of incorrect predictions.It can be seen that the RESNETCNN3 network model proposed in this paper achieves a high detection success rate for monitoring twelve types of abnormal traffic in the CICIDS 2017 dataset, which proves that our proposed method is effective.

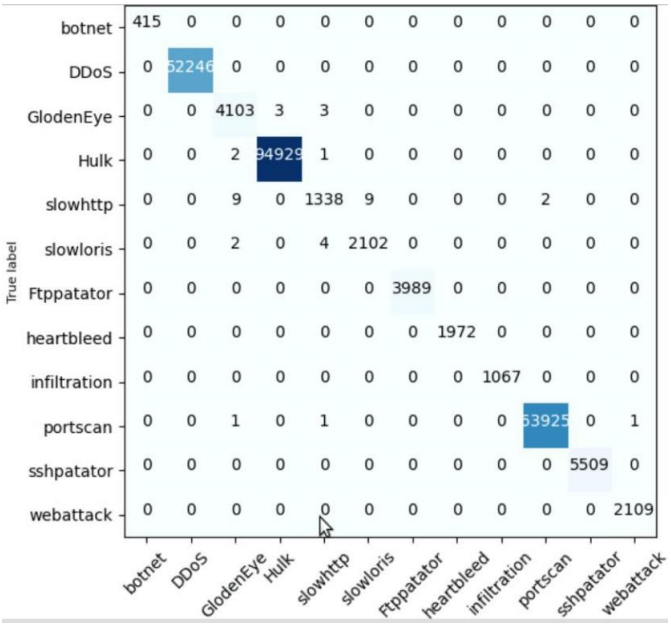

Fig. 8. Confusion Matrix of RESNETCNN3

6. Conclusion

In this paper, an intrusion detection model (RESNETCCN) is proposed that fuses RESNET and parallel cross-convolutional neural networks, and three improved versions of the RESNETCNN network model are designed. In addition, a data oversampling method is introduced to improve the detection accuracy of imbalance data in the ISCXIDS 2012 dataset. The experimental results show that the four RESNETCNN network models proposed in this paper can effectively handle the unbalanced abnormal traffic data and provide an effective solution for network security intrusion detection systems.

Although the RESNETCNN network model achieves extremely high classification accuracy, the current network environment is complex and changing, resulting in the RESNETCNN model based on closed-set protocols cannot meet the new network anomaly traffic detection requirements. In our future work, we will introduce more new ideas such

as blockchain cryptography [8], [18], [9], [19], [16], alliance chain [36], [7], [20], visual Q&A [5], [28], transformer [21], panoramic image [17], reinforcement learning [3], internet of things [23], [24], shared data [6] in our model.We will continue to explore network intrusion detection methods in more areas such as unsupervised and semi-supervised [2] areas for network anomalous traffic data detection. In addition, we also try to introduce new evaluation metrics and establish systematic evaluation methods of intrusion detection.

References

1. Bang, J.h., Cho, Y.j., Kang, K.: Anomaly detection of network-initiated lte signaling traffic in wireless sensor and actuator networks based on a hidden semi-markov model 65, 108–120 (2017)
2. Cai, S., Han, D., Li, D.: A feedback semi-supervised learning with meta-gradient for intrusion detection. IEEE Systems Journal (2022)
3. Cai, S., Han, D., Li, D., Zheng, Z., Crespi, N.: An reinforcement learning-based speech censorship chatbot system. The Journal of Supercomputing 78(6), 8751–8773 (2022)
4. Cai, S., Han, D., Yin, X., Li, D., Chang, C.C.: A hybrid parallel deep learning model for efficient intrusion detection based on metric learning. Connection Science 34(1), 551–577 (2022)
5. Chen, C., Han, D., Chang, C.C.: Caan: Context-aware attention network for visual question answering. Pattern Recognition 132, 108980 (2022)
6. Cui, M., Han, D., Wang, J., Li, K.C., Chang, C.C.: Arfv: An efficient shared data auditing scheme supporting revocation for fog-assisted vehicular ad-hoc networks. IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY 69(12), 15815–15827 (2020)
7. Gao, N., Han, D., Weng, T.H., Xia, B., Li, D., Castiglione, A., Li, K.C.: Modeling and analysis of port supply chain system based on fabric blockchain. COMPUTERS & INDUSTRIAL ENGINEERING 172(A) (2022)
8. Han, D., Pan, N., Li, K.C.: A traceable and revocable ciphertext-policy attribute-based encryption scheme based on privacy protection. IEEE Transactions on Dependable and Secure Computing 19(1), 316–327 (2022)
9. Han, D., Zhu, Y., Li, D., Liang, W., Souri, A., Li, K.C.: A blockchain-based auditable access control system for private data in service-centric iot environments. IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS 18(5), 3530–3540 (2022)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
11. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation 9(8), 1735–1780 (1997)
12. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning. pp. 448–456. PMLR (2015)
13. Ji, S.: Research on network traffic intrusion detection based on deep learning (2020)
14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Communications of the ACM 60(6), 84–90 (2017)
15. Li, B., Zhang, S., Li, K.: Towards a multi-layers anomaly detection framework for analyzing network traffic 29 (2017)
16. Li, D., Han, D., Weng, T.H., Zheng, Z., Li, H., Liu, H., Castiglione, A., Li, K.C.: Blockchain for federated learning toward secure distributed machine learning systems: a systemic survey. Soft Computing 26(9), 4423–4440 (2022)
17. Li, D., Han, D., Zhang, X., Zhang, L.: Panoramic image mosaic technology based on sift algorithm in power monitoring. In: 2019 6th International Conference on Systems and Informatics (ICSAI). pp. 1329–1333. IEEE (2019)

18. Li, H., Han, D., Tang, M.: A privacy-preserving charging scheme for electric vehicles using blockchain and fog computing. IEEE SYSTEMS JOURNAL 15(3), 3189–3200 (2021)
19. Li, H., Han, D., Tang, M.: A privacy-preserving storage scheme for logistics data with assistance of blockchain. IEEE INTERNET OF THINGS JOURNAL 9(6), 4704–4720 (2022)
20. Li, J., Han, D., Wu, Z., Wang, J., Li, K.C., Castiglione, A.: A novel system for medical equipment supply chain traceability based on alliance chain and attribute and role access control. Future Generation Computer Systems 142, 195–211 (2022)
21. Li, M., Han, D., Li, D., Liu, H., Chang, C.C.: Mfvt: an anomaly traffic detection method merging feature fusion network and vision transformer architecture. EURASIP Journal on Wireless Communications and Networking 2022(1), 1–22 (2022)
22. Li, M., Han, D., Yin, X., Liu, H., Li, D.: Design and implementation of an anomaly network traffic detection model integrating temporal and spatial features. Security and Communication Networks 2021 (2021)
23. Liu, H., Han, D., Cui, M., Li, K.C., Souri, A., Shojafar, M.: Idenmultisig: Identity-based decentralized multi-signature in internet of things. IEEE Transactions on Computational Social Systems pp. 1–11 (2023)
24. Liu, H., Han, D., Li, D.: Fabric-iot: A blockchain-based access control system in iot. IEEE Access 8, 18207–18218 (2020)
25. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440 (2015)
26. Nezhad, S.M.T., Nazari, M., Gharavol, E.A.: A novel dos and ddos attacks detection algorithm using arima time series model and chaotic system in computer networks 20(4), 700–703 (2016)
27. Sercu, T., Puhrsch, C., Kingsbury, B., LeCun, Y.: Very deep multilingual convolutional neural networks for lvcsr. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP). pp. 4955–4959 (2016)
28. Shen, X., Han, D., Guo, Z., Chen, C., Hua, J., Luo, G.: Local self-attention in transformer for visual question answering. APPLIED INTELLIGENCE (2022)
29. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2015)
30. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2818–2826 (2016)
31. Tian, Q., Han, D., Li, K.C., Liu, X., Duan, L., Castiglione, A.: An intrusion detection approach based on improved deep belief network. APPLIED INTELLIGENCE 50(10), 3162–3178 (2020)
32. Tian, Q., Han, D., Li, K.C., Liu, X., Duan, L., Castiglione, A.: An intrusion detection approach based on improved deep belief network. Applied Intelligence 50(10), 3162–3178 (2020)
33. Visa, S., Ramsay, B., Ralescu, A.L., Van Der Knaap, E.: Confusion matrix-based feature selection. MAICS 710, 120–127 (2011)
34. Wang, Z., Han, D., Li, M., Liu, H., Cui, M.: The abnormal traffic detection scheme based on pca and ssh. Connection Science 34(1), 1201–1220 (2022)
35. Wasikowski, M., Chen, X.w.: Combating the small sample class imbalance problem using feature selection. IEEE Transactions on knowledge and data engineering 22(10), 1388–1400 (2009)
36. Xiao, T., Han, D., He, J., Li, K.C., de Mello, R.F.: Multi-keyword ranked search based on mapping set matching in cloud ciphertext storage system. CONNECTION SCIENCE 33(1), 95–112 (2021)
37. Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1492–1500 (2017)

38. Yang, C.: Anomaly network traffic detection algorithm based on information entropy measurement under the cloud computing environment 22(4), S8309–S8317 (2019)
39. Zhang, Y., Chen, X., Guo, D., Song, M., Teng, Y., Wang, X.: Pccn: parallel cross convolutional neural network for abnormal network traffic flows detection in multi-class imbalanced network traffic flows. IEEE Access 7, 119904–119916 (2019)
40. Zhang, Y., Chen, X., Jin, L., Wang, X., Guo, D.: Network intrusion detection: Based on deep hierarchical network and original flow data. IEEE Access 7, 37004–37016 (2019)
41. Zhong, Y., Chen, W., Wang, Z., Chen, Y., Wang, K., Li, Y., Yin, X., Shi, X., Yang, J., Li, K.: Helad: A novel network anomaly detection model based on heterogeneous ensemble learning (2020)
42. Zhou, Z.H.: Machine learning. Prentice Hall, Springer Nature (2021)

Yimin Li received the B.S degree from Hubei University of Economics,Wuhan,China.She is currently working toward the M.S. degree with Shanghai Maritime University, Shanghai, China. Her main research interests include intrusion detection, cloud computing security, machine learning, and deep learning.

Dezhi Han received the B.S. degree from Hefei University of Technology, Hefei, China, the M.S. and Ph.D. degrees from Huazhong University of Science and Technology, Wuhan, China. He is currently a Professor of computer science and engineering with Shanghai Maritime University, Shanghai, China. His specific interests include storage architecture, blockchain technology, cloud computing security, and cloud storage security technology.

Mingming Cui received the B.S. degree in Computer Science and Technology from the Anhui University of Finance and Economics, Bengbu, China. She is currently pursuing the Ph.D. degree in Information management and information systems from Shanghai Maritime University, Pudong, China. She is currently a Visiting Ph.D. student in the Nanyang Technological University, Singapore. Her current research interests include cryptology, blockchain, data privacy protection, network security, VANETS security, and Internet of things.

Yuan Fan received the B.S. degree from Nanjing University of Posts and telecommunications, Nanjing, China, and the M.S. degree from San Jose State University, California, the U.S.A. He is the Chairman of DBAPPSecurity Co., Ltd. His specific interests include cybersecurity and data security. In recognition of his great contribution, he has been granted the special allowance of the State Council, selected into the National Million Talent Project, he is also a member of the 10th National Committee of China Association for Science and Technology.

Yachao Zhou received the B.S. degree from Beijing University of Posts and Telecommunications, Beijing, China, the M.S. degree from Tsinghua University, Beijing, China, and Ph.D. degree from Dublin City University, Dublin, Ireland. She is the chief scientist of DBAPPSecurity Co., Ltd. Shanghai headquarters. Her specific interests include cloud computing, deep packet inspection and IoT security.

Received: November 24, 2022; Accepted: February 10, 2023.