文章编号：1673-3193(2024)01-0066-08

基于多阶段特征选择和CNN-GRU的网络入侵检测模型

王相月，赵利辉

（中北大学 软件学院，山西 太原 030051）

摘　要：针对网络入侵检测数据中冗余特征多导致入侵检测准确率低的问题，本文提出了一种基于多阶段特征选择和CNN-GRU的网络入侵检测模型。首先，针对数据集的特征冗余，结合皮尔逊相关系数(PCC)和随机森林(RF)构建PCC-RF特征选择算法进行多阶段特征选择，构造最优特征子集。其次，利用卷积神经网络(CNN)对空间特征的强大提取能力和门控循环单元(GRU)的优秀时序特征提取能力，构建CNN-GRU模型。最后，将最优特征子集输入到CNN-GRU模型中进行训练。使用UNSW-NB15数据集进行实验，实验结果表明：数据集在经过PCC-RF特征处理算法后，维度更低，效果更佳，本文所提模型检测准确率达到84.72%。

关键词：网络入侵检测；特征选择；卷积神经网络；门控循环单元

中图分类号：TP393.0　　**文献标识码：**A　　**doi:** 10.3969/j.issn.1673-3193.2024.01.009

引用格式：王相月，赵利辉．基于多阶段特征选择和CNN-GRU的网络入侵检测模型[J]．中北大学学报(自然科学版)，2024，45(1)：66-73.

WANG Xiangyue，ZHAO Lihui. Network intrusion detection model based on multi-stage feature selection and CNN-GRU[J]. Journal of North University of China(Natural Science Edition)，2024，45(1)：66-73.

Network Intrusion Detection Model Based on Multi-Stage Feature Selection and CNN-GRU

WANG Xiangyue，ZHAO Lihui

（School of Software，North University of China，Taiyuan 030051，China）

Abstract：A network intrusion detection model based on multi-stage feature selection and CNN-GRU is proposed to address the problem of low accuracy of intrusion detection due to redundant features of network intrusion detection data. Firstly，for the feature redundancy of the data set，the PCC-RF feature selection algorithm is constructed by combining Pearson correlation coefficient and random forest for multi-stage feature selection and constructing the optimal feature subset. Then the CNN-GRU model is constructed by using the powerful extraction capability of convolutional neural network for spatial features and the excellent temporal feature extraction capability of gated recurrent units. Finally，the optimal feature subset is input into the CNN-GRU model for training. Experiments are conducted by using the UNSW-NB15 dataset，and the experimental results show that the dataset，after the PCC-RF feature processing algorithm，has lower dimensionality and better results compared with other methods. The model detection

收稿日期：2023-04-13

作者简介：王相月(1998—)，男，硕士生，主要从事深度学习与网络安全方向的研究。

通信作者：赵利辉(1979—)，男，副教授，博士，主要从事网络测试与网络安全方向的研究。E-mail：leehwi@nuc.edu.cn。

accuracy reaches 84.72%.

Key words: network intrusion detection; feature selection; convolutional neural network; gated cycle unit

0 引 言

互联网被广泛运用在生产生活的各个方面，已经成为支撑当今社会快速高效运转的底层驱动力，对国家经济与社会发展都起到了重要作用。与此同时，各种网络安全问题层出不穷，严重危害着人们正常生产和生活秩序。针对网络安全问题的研究越来越重要，网络入侵检测系统（Network intrusion detection system）作为一种能高效、实时检测网络攻击的方法，对维护网络安全起到了巨大作用，具有很高的研究价值。

伴随着网络技术的不断发展，网络入侵检测也经历了多次迭代。基于数据来源的角度，可以将入侵检测系统划分为基于主机的入侵检测和基于网络的入侵检测。基于检测技术的角度，可以将入侵检测系统划分为基于误用的入侵检测和基于异常的入侵检测[1]。从数据来源和检测技术两个角度同时来看，基于网络的异常入侵不仅发生次数多、难以检测，而且危害巨大。现有基于网络的异常入侵检测方法中，既有基于机器学习的检测和基于深度学习的检测，也有基于强化学习等新技术的检测。

在使用机器学习进行入侵检测方面，任家东等[2]采用KNN离群点检测算法先从原始数据集中获取一个高质量的小型训练数据集，然后结合类别检测分组方法构建多层次的随机森林模型进行网络异常检测，提高了异常类型的检测准确率。高兵等[3]提出了基于麻雀搜索算法的改进粒子群优化（SSAPSO）算法对轻量级梯度提升机（Light GBM）进行参数寻优，在保证寻优精度的同时快速收敛，提高了网络入侵检测的准确率。Ahmed等[4]运用综合少数类过采样技术（SMOTE）来解决数据集的类别不平衡问题，使随机森林分类模型在网络入侵检测中的检测精度得到了进一步提高。Chen等[5]将袋装树运用于网络入侵检测中，提出的基于袋装树的入侵检测模型比其它使用机器学习分类器算法的入侵检测模型具有更高的准确率和更低的误报率。

机器学习需要大量的专业知识来进行人工特征选择等，而当前的网络入侵数据集数据量大、相关特征多，所以机器学习的入侵检测能力无法应对，而深度学习具有强大的特征学习能力，能解决这些问题。Diaba等[6]提出了一种混合深度学习算法，通过融合卷积神经网络和门控递归单元并行提取的特征再进行检测，有效提高了分布式拒绝服务攻击的检测率。Aldarwbi等[7]将流量特征转化为频率特征，并利用先进的基于音频识别的深度学习技术检测网络入侵，提高了入侵检测精度。Milosevic等[8]实现了用于入侵检测的深度神经网络，并调整超参数分析了不平衡数据中少数类的检测性能，大大提高了少数类的检测效果。Kurni等[9]针对大数据环境下深度入侵检测效果差的问题，提出了一种基于鳐鱼政治优化的深度超限网络，有效提高了入侵检测准确率。

使用机器学习和深度学习进行入侵检测的方法虽然具有一定的效果，但是使用特征多，在数据集较大时不能很好地平衡资源消耗与检测准确率。另外，单一的深度学习模型对入侵检测数据集特征提取不充分，检测效果欠佳。

本文提出基于多阶段特征选择和CNN-GRU的网络入侵检测模型，主要工作包括：1）提出PCC-RF特征选择算法在UNSW-NB15数据集[10]基础上进行多阶段特征选择，构建最优特征子集；2）构建CNN-GRU模型进行数据特征的学习；3）通过实验确定CNN-GRU模型的最佳参数，并验证了模型结构的合理性和检测效果的优越性。本研究为最优特征集的构建提供了一种新方法，证明了结合CNN和GRU的混合深度模型比单一的深度学习模型在入侵检测数据特征提取上具有更好的效果，为不断迭代的网络入侵提供了一种新的检测方式。

1 PCC-RF特征选择模型

1.1 皮尔逊相关系数

皮尔逊相关系数（Pearson correlation coefficient，PCC）r是用来衡量两个变量之间相关程度的值，其范围为（−1,1），绝对值越接近1，说明变量相关性越强。

长度为n的样本集$\{x_1,x_2,\cdots,x_i,\cdots,x_n\}$和$\{y_1,y_2,\cdots,y_i,\cdots,y_n\}$的相关系数计算公式为

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}},\quad(1)$$

式中：$\overline{X}$和$\overline{Y}$分别是样本集的均值。

1.2 随机森林

随机森林[11](Random forest，RF)是集成学习的一种，由多个决策树组合而成，既可以执行回归任务，也可以执行分类任务。随机森林分类过程如图1所示。

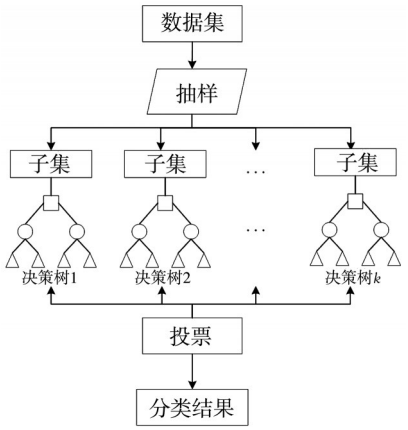

图1 随机森林的分类流程

Fig. 1 Classification process of random forest

1)假如总数据集R中共有N个样本，那么每次随机选择N个样本。选中的样本用来训练1个决策树，作为决策树根节点处的样本。

2)若每个样本有M个属性，在决策树的每个节点需要分裂时，随机从这M个属性中选取出m个属性，满足条件$m\ll M$。然后，从这m个属性中采用某种方法来选择1个属性作为该节点的分裂属性。

3)在单个决策树形成过程中，每个节点都按照步骤2)来分裂，直到没有属性可以分裂。

4)按照步骤1)~3)形成k个决策树，它们共同组合起来就是随机森林。

5)统计所有决策树的分类结果，出现次数最多的类别，就是随机森林模型的分类结果。

1.3 PCC-RF特征选择模型

首先，数据集中的特征有一部分具有很强的相关性，在进行后续学习时高相关特征不仅对训练效果没有明显提升，而且会造成资源浪费。PCC是一种计算特征相关性的有效方法，所以本文利用皮尔逊相关系数计算出各个特征之间的相关性值，删除大于阈值的高相关特征，这样能够有效提升训练效果，并减少计算开销。

其次，数据集中的噪声特征也会在训练时降低模型的训练效果，利用RF算法对数据进行训练可以获取各个特征的重要程度值，将大于阈值的特征定义为噪声特征，并将其去除，能够提高模型训练效果。

本文结合皮尔逊相关系数和随机森林提出PCC-RF算法进行多阶段特征选择。先利用PCC计算原始数据集中特征间的相关性，根据阈值去除高相关的特征，保留其余特征。然后，将保留的数据输入RF模型进行训练获取特征的重要程度，根据阈值保留部分重要特征构建最优特征子集。

PCC-RF特征选择模型负责读取数据，对数据集进行数据预处理，然后利用PCC-RF算法对原始数据集进行数据降维，构建出最优特征子集。

2 CNN-GRU模型

2.1 卷积神经网络

卷积神经网络(Convolutional neural network，CNN)的结构可以被分为输入层、卷积层、池化层和全连接层。CNN结构如图2所示。

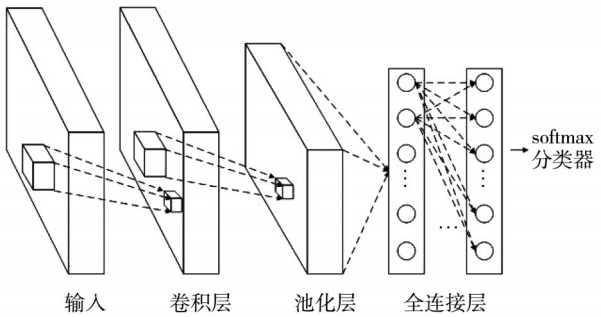

图2 卷积神经网络结构

Fig. 2 Convolutional neural network structure

卷积层的作用是使用过滤器对输入的数据进行特征提取，然后输出给池化层。池化层的作用是在不影响结果的情况下，减少卷积层输出数据的规模，也就是下采样。全连接层主要用来对卷积层和池化层输出的特征进行分类。

2.2 门控循环单元

门控循环单元[12](Gate recurrent unit，GRU)

是长短期记忆神经网络[13]（Long-short term memory，LSTM）的一个变体，与LSTM结构相似，拥有更新门和重置门，但GRU一个门控就可以实现LSTM需要多个门控才可以实现的遗忘和记忆操作，所以参数比LSTM少，但是效果却和LSTM一样。GRU结构如图 3 所示。

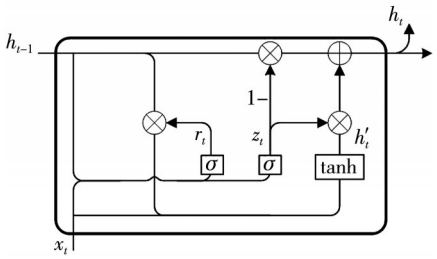

图 3　GRU结构

Fig. 3　Structure of GRU

GRU原理为

$$z_t=\sigma(\boldsymbol{W}_z\cdot[h_{t-1},x_t]), \tag{2}$$

$$r_t=\sigma(\boldsymbol{W}_r\cdot[h_{t-1},x_t]), \tag{3}$$

$$h'_t=\tanh(\boldsymbol{W}\cdot[h_{t-1}\cdot r_t,x_t]), \tag{4}$$

$$h_t=(1-z_t)h_{t-1}+z_t h'_t, \tag{5}$$

式中：z_t是更新门；r_t是重置门；h'_t是当前时间步t的记忆内容；h_t是当前时间步t的最终记忆内容；x_t是当前时间步t的输入数据；h_{t-1}保存的是前一个时间步$t-1$的信息；σ是sigmoid函数；$\boldsymbol{W}$是权重矩阵。

2.3　CNN-GRU模型

在对数据集进行特征提取时，CNN具有强大的空间特征提取能力，而GRU具有结构简单和时间特征提取能力强的优点。为了充分提取数据集中的特征信息，本文将CNN和GRU结合起来构建了CNN-GRU模型，依次提取数据集的空间特征和时间特征。

CNN-GRU模型主要由CNN、GRU和全连接神经网络组成，包含1个输入层、1个卷积层、1个池化层、1个Flatten层、1个GRU层和1个使用softmax作为激活函数的全连接层，如图 4 所示。

卷积层使用一维CNN，参数filters为16，kernel_size为3，strides为2，padding为same，使用rule作为激活函数。GRU层有64个神经元。模型使用交叉熵计算损失，使用adam优化器进行优化。

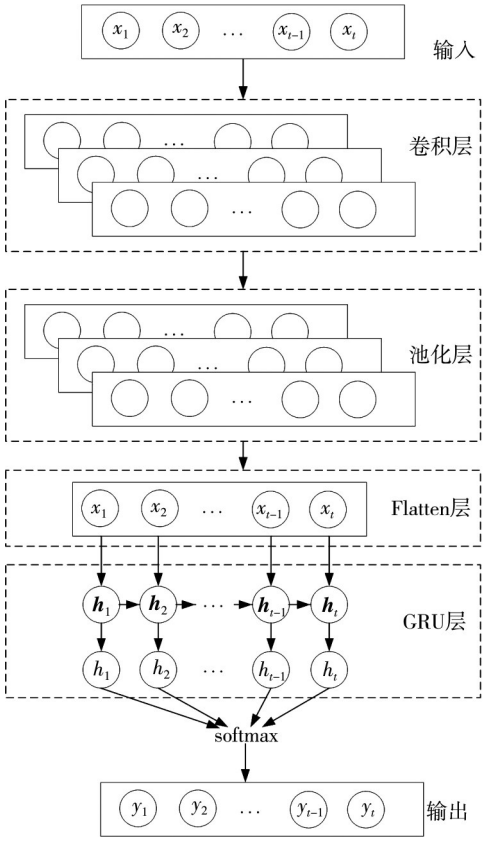

图 4　CNN-GRU模型

Fig. 4　CNN-GRU model

深度学习模型使用浮点数计算量FLOPs（FLoating-point OPerations）来表示时间复杂度，而模型的空间复杂度用模型的参数量表示。本文模型中的时间复杂度和空间复杂度主要受模型中单个CNN、单个GRU和最后的全连接层影响。利用tensorflow框架的profiler对模型的浮点数计算量和参数量进行计算，本文模型的时间复杂度约为86 300 FLOPs，空间复杂度约为38 000。由此可知，本文模型的时间复杂度和空间复杂度都较小，能够很好地节约资源，提高入侵检测速度。

3　基于多阶段特征选择和CNN-GRU的网络入侵检测模型

3.1　模型结构

基于多阶段特征选择和CNN-GRU的网络入侵检测模型（PCC-RF-CNN-GRU）如图 5 所示，主要由PCC-RF特征选择模型和CNN-GRU模型组成。

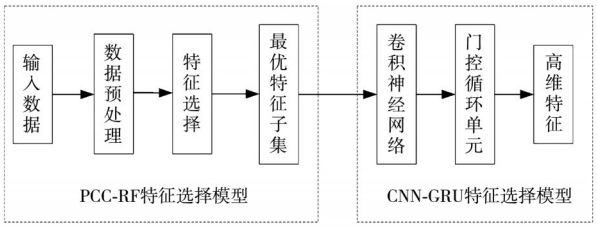

图5 PCC-RF-CNN-GRU模型

Fig. 5 PCC-RF-CNN-GRU model

3.2 入侵检测流程

基于多阶段特征选择和CNN-GRU的网络入侵检测流程如图6所示。

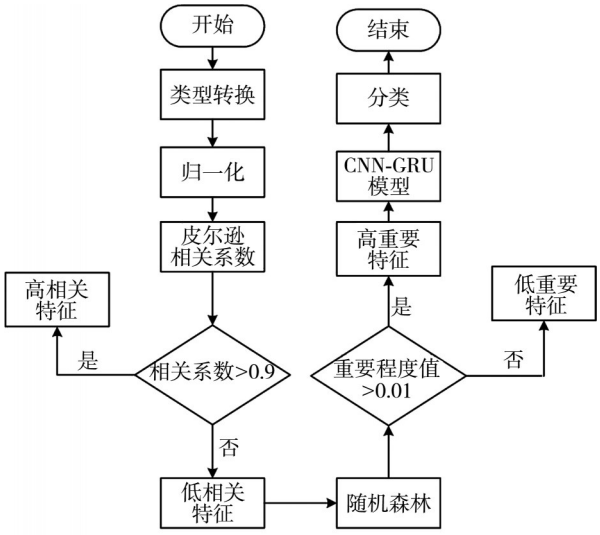

图6 检测流程

Fig. 6 Testing process

3.2.1 数据预处理

1）类型转换。通过分析发现，数据集中含有4个字符型特征，分别是proto、service、state和attack_cat，它们无法被直接使用，需要进行类型转换，利用Python的类别型变量category将其替换成数值特征。

2）归一化。进行归一化操作能够加快梯度下降求最优解的速度，提高模型的精度。本文使用标准差标准化进行归一化，标准差标准化可以使转换后的数据符合标准正态分布，公式为

$$x^{*}=\frac{x-\mu}{\sigma}, \tag{6}$$

式中：μ是全体样本数据的均值；σ是全体样本数据的标准差。

3.2.2 特征选择

利用PCC-RF算法进行多阶段特征选择，并通过仿真实验确定最佳相关系数阈值为0.9和最佳重要程度阈值为0.01，至此将数据集从44维降低到21维，构建出最优特征子集。

3.2.3 高维特征提取

通过构建CNN-GRU模型将CNN和GRU的优势结合起来，将最优特征子集输入到CNN-GRU模型中提取出高维特征，用于后续分类。

3.2.4 分类

先对获取的高维特征进行降维操作，然后使用一层全连接神经网络对提取的高维特征进行分类，检测各种网络攻击。

4 实验与分析

为了验证模型的效果，实验使用sklearn和keras框架。实验环境为Window 10操作系统，Intel(R) Core(TM) i7-6500U CPU @2.50 GHz，AMD Radeon (TM) R7 M360，RAM 16 GB。

4.1 数据集与评估标准

4.1.1 数据集

实验使用新南威尔士大学UNSW-NB15入侵检测数据集，原始数据集包含被分为训练集和测试集的两个CSV文件，因此，本文使用官方训练集和测试集分别用作模型的训练和测试。

该数据集总共包含10种类别，训练集包含82 332条数据，测试集包含175 341条数据，数据集样本分布状况如表1所示。数据集包含45个特征，其中数值型特征41个，符号型特征4个。该数据集包含真实的网络活动，被作为基准数据集广泛应用于各种网络入侵检测研究之中。

表1 UNSW-NB15数据集样本分布状况

Tab. 1 Sample distribution status of the UNSW-NB15 dataset

数据集	类别									
	Normal	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Reconnaissance	Shellcode	Worms
训练集	37 000	677	583	4 089	11 132	6 062	18 871	3 496	378	44
测试集	56 000	2 000	1 746	12 264	33 393	18 184	40 000	10 491	1 133	130

4.1.2 评估标准

为了评估本文模型的效果，使用准确率*Accuracy*、精准度*Precision*、召回率*Recall*和*F*1值作为评估指标。准确率可计算分类正确的样本

与总样本的比值；精准度表示预测为正的样本与其中实际正样本的比例；召回率表示预测为正的样本与实际正样本的比值；$F1$值是召回率和精准度的调和平均。

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

$$Precision=\frac{TP}{TP+FP},\tag{8}$$

$$Recall=\frac{TP}{TP+FN},\tag{9}$$

$$F1=\frac{2Recall\cdot Precision}{Recall+Precision},\tag{10}$$

式中：TP为正确预测为正样本的数量；FP为错误预测为正样本的数量；TN为正确预测为负样本的数量；FN为错误预测为负样本的数量。

4.2　实验结果与分析

4.2.1　参数实验

在PCC-RF特征选择算法中，特征相关系数阈值和特征重要程度阈值直接决定最优特征子集的构建，本文通过对比多组实验效果确定最佳参数，实验结果如表2和表3所示。通过分析特征相关系数，以0.5～0.9为范围，每次递增0.1，进行多组实验。由表2可知，实验准确率随阈值变化呈现先增后减的趋势，在阈值为0.9时具有最高的准确率82.39%，故确定最佳相关系数阈值为0.9。

表2　相关系数阈值实验结果

Tab. 2　Experimental results of correlation coefficient threshold

阈值	*Accuracy*/%	*Recall*/%	*Precision*/%	*F1*/%
0.5	54.66	53.28	55.65	54.52
0.6	57.36	56.01	57.82	56.95
0.7	64.78	63.72	65.85	64.01
0.8	78.74	76.24	79.25	78.53
0.9	**82.39**	**80.82**	**82.73**	**80.46**
1.0	80.85	79.27	81.81	79.25

表3　重要程度阈值实验结果

Tab. 3　Experimental results of importance thresholds

阈值	*Accuracy*/%	*Recall*/%	*Precision*/%	*F1*/%
0.001	72.34	69.16	73.49	71.22
0.005	78.56	75.81	79.14	78.24
0.01	**81.82**	**78.9**	**82.39**	**81.54**
0.05	74.91	71.48	75.19	74.63
0.1	53.12	49.89	53.91	52.63

根据随机森林获取的特征重要性状况，由小到大依次选择0.001，0.005，0.01，0.05和0.1作为重要程度阈值。由表3可知，实验准确率随重要程度阈值增大出现先增后减的趋势，在阈值为0.01时具有最高的准确率81.82%，所以确定最佳重要程度阈值为0.01。

对于深度学习的过程，超参数对实验效果具有很大的影响，尤其是CNN的过滤器数量和GRU的单元数，所以针对这两个参数，本文通过多组实验来寻找最佳参数，以确保获取最佳检测效果。

为了确定最佳CNN过滤器数量，在1～128之间以2次幂的频率逐步增加CNN过滤器数量进行多组仿真实验，实验结果如表4所示。由表4可知，随着CNN过滤器数量不断增加，实验准确率出现先增后减的趋势，并在CNN过滤器数量为16时表现出最佳准确率83.06%，因此，确定最佳CNN过滤器数量为16。

表4　不同CNN过滤器数量的实验结果

Tab. 4　Experimental results with different number of CNN filters

过滤器数	*Accuracy*/%	*Recall*/%	*Precision*/%	*F1*/%
1	71.26	64.01	72.83	69.70
2	73.78	65.71	73.91	70.43
4	71.73	61.28	74.32	69.62
8	81.19	75.92	82.83	80.58
16	**83.06**	**79.55**	**84.98**	**82.93**
32	83.44	78.13	83.96	81.11
64	80.83	77.11	81.40	79.51
128	79.06	75.27	80.60	78.10

为了确定最佳GRU单元数，在1～256之间以2次幂的频率增加GRU单元数进行多组仿真实验，实验结果如表5所示。随着GRU单元数增加，实验准确率出现先增加后减小的趋势，并在GRU单元数为64时表现出最佳准确率84.17%，因此，确定最佳GRU单元数为64。

表5　不同GRU单元数的实验结果

Tab. 5　Experimental results for different number of GRU units

GRU单元数	*Accuracy*/%	*Precision*/%	*Recall*/%	*F1*/%
1	79.96	71.57	80.31	70.66
2	81.32	72.18	81.83	71.38
4	79.69	71.70	79.62	69.52
8	80.27	70.94	82.31	72.72
16	81.81	75.10	83.25	74.04
32	78.71	69.95	79.09	68.98
64	**84.17**	**78.78**	**85.93**	**81.98**
128	83.06	73.55	84.98	78.93
256	82.86	72.01	83.48	77.52

4.2.2　消融实验

为了证明CNN-GRU模型结构的优势，使用CNN-GRU、CNN和GRU分别在最优特征子集上进行实验，如表6所示。实验结果表明，本文的CNN-GRU模型比CNN模型和GRU的模型具

有更高的检测准确率和$F1$值，证明了本文模型比CNN模型和GRU模型具有明显优势。

表6 本文模型与CNN和GRU模型的对比

Tab. 6 Comparison of the proposed model with CNN and GRU

模型	*Accuracy*/%	*Precision*/%	*Recall*/%	*F*1/%
CNN	79.89	43.05	53.60	40.75
GRU	78.72	42.62	43.25	39.44
本文模型	**84.72**	**85.62**	**86.18**	**83.05**

4.2.3 对比实验

随机森林、决策树(Decision tree，DT)、逻辑回归(Logistic regression，LR)和K近邻(K-Nearest neighbor，KNN)都是常见机器学习算法[14-15]。将本文所提模型与这些机器学习算法对比，结果如表7和图7所示。RF准确率是77.06%，DT准确率是44.20%，LR准确率是77.62%，KNN准确率是74.92%，它们的准确率都低于本文方法的84.72%，证明本文所提模型在入侵检测方向比传统机器学习算法具有优势。

表7 本文所提模型与机器学习算法对比

Tab. 7 Comparison of the proposed model with machine learning algorithm

模型	*Accuracy*/%	*Precision*/%	*Recall*/%	*F*1/%
RF	77.06	73.47	79.11	78.27
DT	44.2	59.56	49.32	41.83
LR	77.62	74.01	78.32	76.87
KNN	74.92	71.88	75.39	73.03
本文模型	**84.72**	**85.62**	**86.18**	**83.05**

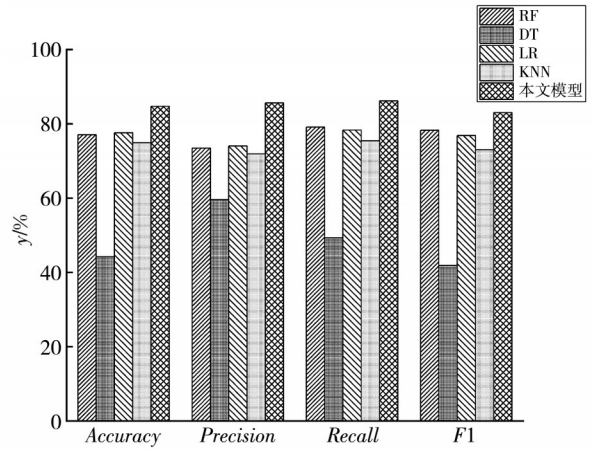

图7 本文所提出的模型与机器学习算法对比

Fig. 7 Comparison of the proposed model with machine learning algorithm

关于网络入侵检测的研究，部分学者基于UNSW-NB15数据集进行了仿真实验。本文通过仿真现有相关模型，并与之对比，结果如表8和图8所示。Jiang等[16]提出的OSS-SMOTE-CNN-BiLSTM模型准确率达到77.16%，Wang等[17]提出的SDAE-ELM模型准确率达到72.38%，Dina等[18]的CTGANSamp-DT模型准确率达到67.31%。由表8和图8可知这3个模型的准确率和$F1$值均低于本文模型，故本文模型具有较好地检测效果。

表8 本文所提模型与其他模型的对比结果

Tab. 8 Comparison of the proposed model with other models

模型	*Accuracy*/%	*Precision*/%	*Recall*/%	*F*1/%
文献[16]	77.16	82.63	79.91	81.52
文献[17]	72.38	69.94	87.42	77.71
文献[18]	67.31	63.69	67.31	64.77
本文模型	**84.72**	**85.62**	**86.18**	**83.05**

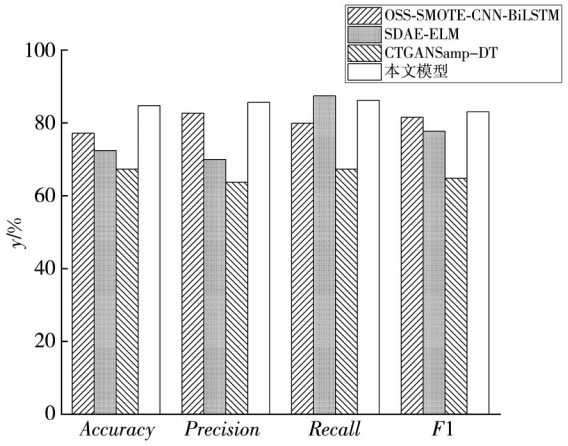

图8 本文所提出的模型与其他模型的对比

Fig. 8 Comparison of the proposed model with other models

5 结 论

本文提出的PCC-RF特征选择算法在特征选择阶段减少了高相关特征和噪声特征的数量，提高了模型的的训练速度和资源消耗。本文设计了CNN-GRU模型提取数据的高维特征，提高了入侵检测的效果。本文所提模型的检测准确率达到84.72%。

参考文献：

[1] 蹇诗婕，卢志刚，杜丹，等. 网络入侵检测技术综述[J]. 信息安全学报，2020，5(4)：96-122.
JIAN Shijie, LU Zhigang, DU Dan, et al. A review of network intrusion detection techniques[J]. Journal of Cyber Security, 2020, 5(4): 96-122. (in Chinese)

[2] 任家东，刘新倩，王倩，等. 基于KNN离群点检测和随机森林的多层入侵检测方法[J]. 计算机研究与发展，2019，56(3)：566-575.
REN Jiadong, LIU Xinqian, WANG Qian, et al. A multilayer intrusion detection method based on KNN

outlier point detection and random forest[J]. Journal of Computer Research and Development, 2019, 56(3): 566-575.(in Chinese)

[3] 高兵，郑雅，秦静，等. 基于麻雀搜索算法和改进粒子群优化算法的网络入侵检测算法[J]. 计算机应用，2022，42(4)：1201-1206.
GAO Bing, ZHENG Ya, QIN Jing, et al. Network intrusion detection algorithm based on sparrow search algorithm and improved particle swarm optimization algorithm[J]. Journal of Computer Applications, 2022, 42(4): 1201-1206.(in Chinese)

[4] AHMED H A, HAMEED A, BAWANY N Z. Network intrusion detection using oversampling technique and machine learning algorithms[J]. PeerJ Computer Science, 2022, 8: e820.

[5] CHEN P T, LI F, LI J T. Research on Intrusion Detection Model Based on Bagged Tree[C]//2021 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS). 2021: 579-582.

[6] DIABA S Y, ELMUSRATI M. Proposed algorithm for smart grid DDoS detection based on deep learning[J]. Neural Networks, 2023, 159: 175-184.

[7] ALDARWBI M Y, LASHKARI A H, GHORBANI A A. The sound of intrusion: A novel network intrusion detection system[J]. Computers and Electrical Engineering, 2022, 104: 108455.

[8] MILOSEVIC M S, CIRIC V M. Extreme minority class detection in imbalanced data for network intrusion[J]. Computers & Security, 2022, 123: 102940.

[9] KURNI M, MD M S, YANNAM B B, et al. MRPO-Deep maxout: Manta ray political optimization based Deep maxout network for big data intrusion detection using spark architecture[J]. Advances in Engineering Software, 2022, 174: 103324.

[10] MOUSTAFA N, SLAY J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)[C]//2015 Military Communications and Information Systems Conference (MilCIS). 2015: 1-6.

[11] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[12] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. 2014, arXiv. 1412. 3555.

[13] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[14] QUINLAN J R. C4. 5: programs for machine learning[M]. New York: Elsevier, 2014.

[15] HOSMER D W, LEMESHOW S. Applied logistic regression[M]. Hoboken: Wiley, 1989.

[16] JIANG K Y, WANG W Y, WANG A L, et al. Network intrusion detection combined hybrid sampling with deep hierarchical network[J]. IEEE Access, 2020, 8: 32464-32476.

[17] WANG Z D, LIU Y D, HE D D, et al. Intrusion detection methods based on integrated deep learning model[J]. Computers & Security, 2021, 103: 102177.

[18] DINA A S, SIDDIQUE A B, MANIVANNAN D. Effect of balancing data using synthetic data on the performance of machine learning classifiers for intrusion detection in computer networks[J]. IEEE Access, 2022, 10: 96731-96747.