计算机科学
COMPUTER SCIENCE

基于CNN结合BiGRU的恶意流量分类算法研究

杨永平, 王思婷

引用本文

杨永平, 王思婷. 基于CNN结合BiGRU的恶意流量分类算法研究[J]. 计算机科学, 2024, 51(11A): 231100106-9.

YANG Yongping, WANG Siting. Study on Malicious Traffic Classification Algorithm Based on CNN Combined with BiGRU [J]. Computer Science, 2024, 51(11A): 231100106-9.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.231100106

基于 CNN 结合 BiGRU 的恶意流量分类算法研究

杨永平[1]　王思婷[2]
1 北京师范大学珠海分校信息技术学院　广东 珠海 519087
2 北京邮电大学国家移动安全重点实验室　北京 100876

摘　要　网络入侵检测是一项重要的网络安全技术，恶意流量识别分类是网络入侵检测的基础。利用端口检测技术、深度包检测技术、特征工程机器学习算法检测技术在当前网络环境下进行流量识别分类已失效或不易实施，因此文中提出了结合卷积神经网络和循环神经网络改进简化模型门控循环单元的恶意流量识别分类算法模型 CNNBiGRU，运用卷积神经网络 CNN 提取网络流结构特征和空间特征，双向门控循环单元 BiGRU 提取序列特征，符合网络流兼具空间结构和序列特征的特点。在 CIC-IDS2017 公开数据集上进行了测试和模型优化与参数选择，实验结果表明所提算法比经典机器学习算法在分类效果上有一定的优势而且不需要特征工程，与单一神经网络算法相比也具有更好的识别效果，与融合神经网络算法在同等准确率目标衡量下又有一定的学习迭代次数优势，具有更高的学习效率。

关键词：恶意流量分类；深度学习；卷积神经网络；双向门控循环单元
中图法分类号　TP391

Study on Malicious Traffic Classification Algorithm Based on CNN Combined with BiGRU

YANG Yongping[1] and WANG Siting[2]
1 School of Information Technology, Beijing Normal University, Zhuhai, Zhuhai, Guangdong 519087, China
2 National Key Laboratory of Mobile Security, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract　Network intrusion detection is an important network security technology, malicious traffic recognition and classification is the basis of network intrusion detection. In the current network environment, port detection technology, deep packet detection technology, and feature engineering machine learning algorithm detection technology for malicious traffic identification and classification have failed or are not easy to implement. This paper proposes a malicious traffic recognition classification algorithm model CNNBiGRU, which combines convolutional neural network and bidirectional gated recurrent unit. CNNBiGRU uses convolutional neural network CNN to extract network flow structure features and spatial features, and uses bidirectional gated recurrent unit BiGRU to extract sequence features, which is consistent with the characteristics of network flow with both spatial structure and sequence features. Tests and model optimization and parameter selection are performed on the CIC-IDS2017 dataset. The experimental results show that the proposed algorithm has certain advantages in classification effect and no feature engineering is required compared with the classical machine learning algorithm, and also has better recognition effect compared with the single-neural network algorithm. Compared with the fusion neural network algorithm, it maintains the same high detection result and has a little advantage in the number of learning iterations under the same accuracy target measurement.

Keywords　Malicious trafficclassification, Deep learning, Convolutional neural network, Bidirectional gated recurrent unit

1　引言

国家互联网应急中心 2021 年发布的《2020 年中国互联网网络安全报告》[1]显示，网络安全依然面临严峻形势，全年捕获恶意程序样本数量超过 4 200 万个，日均传播次数为 482 万余次，涉及恶意程序家族近 34.8 万个，APT 攻击行动高发，数据泄露风险突出，漏洞利用问题严重，日均漏洞利用攻击达 2 176.4 万次，Web 应用攻击、恶意爬虫攻击广泛存在，Dos 由于攻击成本低、攻击效果显著，仍是目前互联网用户面临的较常见、影响较严重的网络安全威胁。对网络流中的恶意流量进行识别分类是应对网络安全威胁的一种重要技术，网络流量分类就是将网络流量归类到预定义的类别。

用户访问网络所产生的数据流称为网络流量，进行网络流包分析时，分析对象有两种：数据包和数据流。针对数据包分析可以依据其所使用的协议、端口号、载荷内容、数据包大小、载荷包含的内容等，针对数据流的分析对象为多个关联的数据包或者是一次完整的会话和连接，比如 TCP 流包括 TCP 连接建立、TCP 数据传输、TCP 断开连接的一系列交互数据包，监测记录一个关联流里的数据包数量、方向、流量密度、持续时间、包间隔时间等统计信息，找出流特征以及包之间的变

基金项目：广东省教育厅科技项目(2020KTSCX175)；北京师范大学珠海分校校内教研项目(202041)
This work was supported by the Project of Department of Education of Guangdong Province(2020KTSCX175) and Beijing Normal University Zhuhai Campus Teaching and Research Project(202041).
通信作者：杨永平(yangyongping@bnu.edu.cn)

化、作用与联系。恶意流量分类指的是根据数据包特征和数据流的结构特征、时序特征、交互特征等将网络数据流量分为正常使用网络产生流量和网络异常流量，正常流量可以进一步识别其协议结构、应用领域、应用程序、处理技术，可用于进行网络流量监控与调节、流量计数计费等，网络异常流量识别则可判断其是否具有危害性，判断其恶意类别，并对可能的攻击行为和危害后果做出部署和响应，识别粒度随着技术和运算能力的提升可以更加细化。

传统的恶意流量分类方法主要有两种：一种是基于 IP 地址、端口号等数据包头信息进行的识别分类，适合早期互联网阶段，随着代理技术、封装技术、P2P 流量动态端口等技术的广泛应用，这种识别分类方法逐渐失效，识别的类别也不再能适应当前的分类目的，基于协议和端口的流量分类准确度大幅下降，Moore 等[2]的研究显示，基于端口的流量识别分类在最佳情况下已不足 30%，且这个比例还在不断降低；另一类是基于深度包检测（Deep Packet Inspection，DPI）的识别方法，该方法可以检测对比数据包载荷内容，通常采用预定义的各个类别的特征描述正则表达式匹配数据包负载以确定类别。DPI 可用于识别破解攻击、特征病毒、垃圾邮件等，在阻止缓冲区溢出攻击、Dos 攻击等方面也很有效，但深度包检测对数据包内容检测存在窥探用户隐私等法律风险，定义有效匹配规则难度大，规则组合运算复杂，难以自适应新协议。随着协议混淆、加密协议的大量应用，DPI 分类逐渐不可行[3]。

本文提出了基于卷积神经网络 CNN 和双向长短期记忆时间递归循环神经网络 BiGRU 融合的深度学习恶意流量分类算法，不需要定义和提取网络流的特征，直接使用原始数据流包作为分析网络的输入，获取每个数据流的前 N 个包的 M 个字节构成数据流代表，转换成标准的矩阵数据。CNN 结合 BiGRU 提取网络流的空间分布特征和时间序列特征，使用接近真实流量的网络流数据进行算法测试，效果良好，算法在保持高准确率的同时还具有一定的学习效率优势。

2 相关工作

传统的流量分类方法进行网络流量分类存在算法描述能力不足、分类误报率高、无法适应新协议、不能处理加密流量等问题，使用基于特征的机器学习算法进行流量分类可以大幅提高流量分类的准确性，而且能有效地处理加密流量的分类问题。使用采集的历史数据提取特征并训练机器学习模型使之成为具有分类能力的分类器，基于特征学习的机器学习分类步骤如图 1 所示。

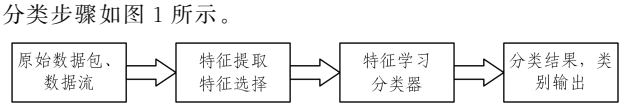

图 1 基于数据包、数据流机器学习流量分类的步骤

Fig. 1 Steps of machine learning traffic classification based on packet and data flow

常见的机器学习分类算法有支持向量机（Support Vector Machines，SVM）[4]、贝叶斯、C4.5 决策树、随机森林（Random forest，RF）、K 近邻（K Nearest Neighbor，KNN）等，根据执行任务目标和具备条件不同可以选择有监督、无监督、半监督、混合式等类别算法。流特征一般包括流持续时间、正向包数量、反向包数量、正向平均包大小、反向平均包大小、流字节率、流包率、包间平均间隔时间等[5]，流特征不需要数据包的内容就可以描述，因此可以用于识别分类加密流量。文献[4]对比了多种机器学习算法在同一异常流量数据集的分类识别情况，无监督学习的分类算法效果较差，不同的算法擅长的流量类别有所不同，在特征工程较好的情况下，机器学习算法的识别准确率可以达到 99.5%。基于机器学习的流量分类算法的主要困难是需要足够的特征工程经验，实施专门的特征提取和筛选工作，不同的分类目标和来源数据需要进行不同的分类标签定义、特征定义和特征选取，标签粒度质量和准确性直接关系到分类结果的可靠性，而且这些机器学习算法对复杂的特征关系的分类表达能力有限，不同的数据和应用领域需要实施不同的特征工程，针对复杂分类问题泛化能力不足，对未知流量进行识别分类，效果难以保证效果的稳定性。

深度学习（Deep Learning，DL）指一种多层的机器学习（Machine Learning，ML）模式，与预定义特征和标签的机器学习方法相比，深度学习不需要分析数据结构，定义数据特征，能够自主进行数据特征学习，学习样本数据的内在规律和表示层次。深度学习方法主要有深度神经网络（DNN）、卷积神经网络（CNN）、循环神经网络（RNN）、生成对抗网络（GAN）等，其中 CNN 在空间领域分析[6]、RNN 在时间序列分析[7]中具有良好的效果。基于深度学习的流量分类方法可以学习数据包和数据流的统计特征、序列特征、字节特征、网络节点联通图及交互特征等，比传统的机器学习方法学习提取的特征更加多样化。加密流量同样具备包特征、流特征、统计特征、网络特征，因此深度学习流量分类利用流相关信息可以分类加密流量，且与非加密流量分类并无方法和流程的明显区别[8]，依赖于内容的更精细化的分类由于加密导致信息被隐藏，还不能达到预期的效果。

基于深度学习的方法可自主学习原始输入数据的特征并建立和输出分类之间的非线性关系，避免了特征定义选取复杂以及描述不准确、不完整的风险。Lotfollahi 等[9]提出的 DeepPacket 是基于数据包原始字节的深度学习方法，将数据包原始字节流数据作为输入，无需进行人工提取特征，分类模型为一维卷积神经网络（1DCNN）和稀疏自动编码机（SAE），取得了 98%的分类准确率；Lopez-Martin 等[10]提出利用数据流前 20 个数据包的端口号、包负载长度、包间隔时间、窗口大小等属性构成 20×6 的矩阵，输入到卷积神经网络（CNN）和长短期记忆循环神经网络（LSTM）的组合模型中，准确率达到 96%以上；Wang 等[11]通过直接向 CNN 输入没有提取特征的原始网络流量用于恶意流量分类，将流序列当成固定大小的二维图像作为模型输入，借助 CNN 算法在图像识别中的良好效果[12]，结果显示该模型能够达到较高准确率，其中 CNN 模型使用了类似于 LeNet-5[13]的结构；Lopez-Martin 等[10]针对融合 CNN 和 RNN 的方法展开了研究，分析了模型结构、特征选取、数据包个数对分类结果的影响，结果显示，模型 CNN + RNN-2a 表现最佳。结构由两个 CNN 层加一个 LSTM 层，然后连接两个全连接层构成，其模型结构为 Conv(32,4,2,1,V)-BN-Conv(64,4,2,1,V)-BN-LSTM(100)-DR(0.2)-FC(100)-DR(0.4)-FC(108)，激活函数为 ReLU(Rectified Linear Units)，损失函数为 Softmax 交叉熵，使用 Adam

的批量随机梯度下降(Stochastic Gradient Descent，SGD)进行优化，Wang 等[14]设计实现基于此算法思想的 HAST-IDS 入侵检测系统，取得了良好的效果；Liu 等[15]提出了融合 CNN 和 LSTM 的网络入侵检测方法，与单纯基于 CNN 的方法、基于 LSTM 的方法进行对比，DoS，Normal，Probe，U2R 和 R2L 这 5 种攻击类型的检测准确率均更高；Deng 等[16]基于 CNN CBAM-BiGRU Attention 多融合算法对加密的恶意流量进行识别，二分类识别率达到 99.95%，多分类识别率达到 99.39%，F1 值相比单一的 CNN 或 RNN 显著提高。

LSTM 在 RNN 模型的基础上进行改进，可以学习较远距离关键字段，GRU 与 LSTM 相似，但神经单元结构更加简单，速度更快，运算量更少。本文提出了一种 CNN 结合 BiGRU的方法对恶意流量进行分类，并使用公开数据进行测试。CNN 从空间范围学习原始流量特征，GRU 学习流量的时间序列方面的特征，算法模型不需要进行特征提取与选择，直接输入原始流量数据，此模型能保证获得良好准确率并具有更好的收敛能力。

3 神经网络算法融合的深度学习流量分类算法模型设计

融合卷积神经网络和循环神经网络的流量分类算法模型的流程包括数据流量的获取、数据预处理、卷积网络层、循环网络层以及分类结果[17]，CNN 部分采用了 LeNet-5 结构，RNN 部分采用了 BiGRU 结构，CNN 的输出作为 BiGRU 的输入，算法网络结构如图 2 所示。

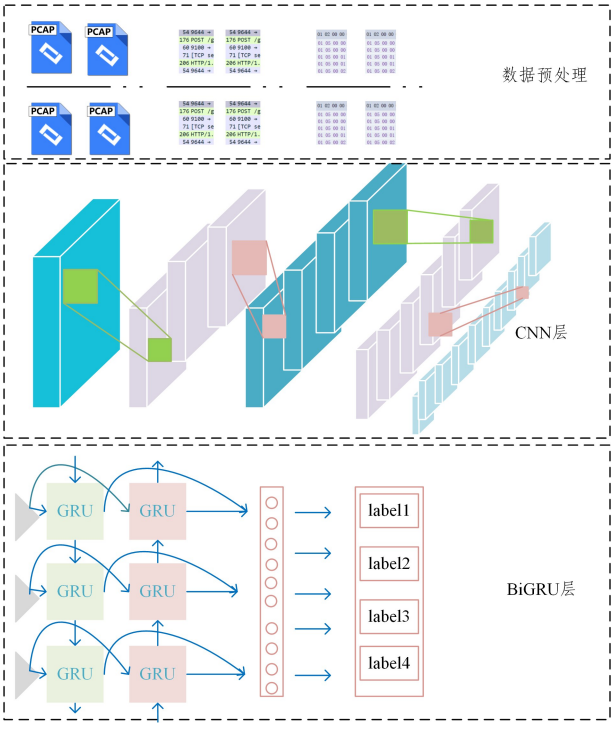

图 2 CNN 结合 BiGRU 的恶意流量分类网络结构图

Fig. 2 Structure diagram of CNN combined with BiGRU for malicious traffic classification

3.1 流数据获取与预处理

网络通信数据由不同协议层次的数据包构成，每个数据包分为数据头和载荷部分。通常互联网获取数据为 IP 层以上数据，数据流由一组地址和逻辑关联的数据包构成，对恶意流量识别和类别判定主要根据其流特征进行。网络流量数据的获取可以在真实网络环境进行捕获，但是由于隐私问题，不适合公开研究和发布。流量生成模型系统可以作为一种网络流量数据采集替代方案，可以使用 BRUTE，iPerf，Ostinato，Nmap，Nessus，IP scanner，Port scanner，Backtrack OS，LOIC 等工具生成特定的网络流量数据。捕获到的网络数据包文件以 pcap 等文件格式存在，基于源地址、源端口、传输协议、目的端口、目的地址五元组，可以划定网络流的边界，采集到的数据用于特征工程或者原始流量的分类。nfdump2，YAF3，CICFlowMeter，pkt2flow 等工具可以作为流划分工具。

文献[10]显示，不需要完整数据流，根据数据流的开始部分包就可以自动提取到流的大部分特征信息，获得良好的学习识别效果，因此可以将每个流序列裁切，保留前 N 个数据包，每个数据包保留 P 个字段，构成一个二维向量继而成为当前网络流的二维图像数据。数据预处理阶段要去掉不完整数据和没有标签的数据流，以及有标签没有数据流的多余的标签信息，将数据流提取的矩阵数据每个字节归一化到 0～1 之间。数据预处理主要完成数据流划分检查、数据清洗、数据裁切、数据归一化的工作，类别混淆乱序化。

3.2 卷积神经网络 CNN

卷积神经网络 CNN 针对图像类二维数据的识别具有良好的效果，和全连接神经网络相比，具有参数少、计算快的特点，并有效减少了过拟合。一个卷积神经网络主要由数据输入层、卷积层、激励层、池化层、全连接层[18]组成，输出特征向量。

3.2.1 数据输入层

输入归一化的二维数据，如果数据量特别大，可以分批处理，每批处理固定条数。

3.2.2 卷积层

使用卷积核进行卷积运算实现特征提取，卷积计算特点是局部关联性和窗口滑动计算，卷积运算输出特征矩阵，运算式如式(1)所示：

$$\boldsymbol{y}_{i,j}=f\left(\sum_{m=-\frac{k}{2}}^{\frac{k}{2}}\sum_{n=\frac{k}{2}}^{\frac{k}{2}}\boldsymbol{w}_{m,n}\boldsymbol{x}_{i+m,j+n}+\boldsymbol{w}_b\right) \tag{1}$$

其中，$\boldsymbol{x}$ 是输入数据；$\boldsymbol{y}$ 为卷积输出，表示数据特征；$\boldsymbol{w}$ 是卷积核(也叫滤波器，卷积核一般为正方形)；k 为卷积核的大小；i 和 j 表示输出位置；m 和 n 表示卷积核的位置；$\boldsymbol{w}_b$ 为 $\boldsymbol{w}$ 的偏置项；f 表示激活函数。卷积运算感知局部特征，网络层数越深感知越全局和抽象。卷积运算过程窗口滑动的权值参数共享机制使得 CNN 参数数量相比全连接网络大幅度降低。

3.2.3 激励层

位于卷积层之后，激活函数将神经网络非线性化，函数连续可导。常用的激活函数有 σ 即 Sigmoid 函数、tanh 函数、ReLU 函数，如图 3 所示。激活函数作为连接上一层输出和下一层输入的连接函数，其导数对应激活函数的影响梯度，通过函数图像对比可知，sigmoid 和 tanh 在靠近饱和区时梯度接近于 0，ReLU 在负半区梯度为 0，sigmoid 函数和 tanh 由于包含幂运算运算速度较慢。Relu 函数计算速度快、收敛快，卷积神经网络通常首选 ReLU 激活函数，迭代速度快，也可以用 tanh 作为候选替代激活函数。激励层和卷积层合并到一起成为实际算法中的卷积层。

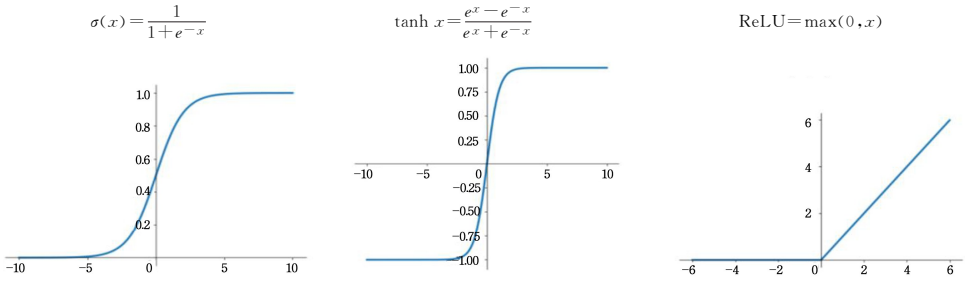

图 3 Sigmoid 函数，tanh 函数，ReLU 函数

Fig. 3 Sigmoid, tanh and ReLU functions

3.2.4 池化层

池化层处于连续的卷积层之间，用于压缩数据，减少参数数量，减小过拟合；保持特征不变的前提下，降低特征的维度，将最重要的特征抽取出来。池化层所用的方法有最大池化 Max-pooling 、平均池化 Average-pooling、随机池化 Stochastic-Pooling、全局平均池化 global average pooling，通常用最大池化来保存变化性特征，在一个高度为 h、宽度为 w 的矩阵中选取最大值作为矩阵区域代表。最大池化计算式如式(2)所示：

$$maxplooling[i,j]=\max(\boldsymbol{x}[i*stride\_h:i*stride\_h+h,j*stride\_w:j*stride\_w+w]) \tag{2}$$

其中，$maxpooling$ 表示池化后的结果，$\boldsymbol{x}$ 表示输入矩阵，$stride\_h$ 和 $stride\_w$ 表示垂直方向和水平方向的步长，h 和 w 表示池化窗口的大小。

3.2.5 全连接层

通常在神经网络尾部，输出结果之前，将卷积输出的二维特征图转化成一维向量，也可用全局平均池化(GAP)代替全连接层或将其添加到全连接层前面，GAP 在保持特征的同时可大幅减少全连接层权重参数。

3.3 循环神经网络 RNN

RNN 描述了序列中前组对后组的影响，当前隐藏状态不仅受到当前输入的影响，还受到前一输入的隐藏状态的影响。其关系见式(3)、式(4)：

$$h_t=f(U\cdot X_t+W\cdot h_{t-1}+b) \tag{3}$$

$$O_t=g(V\cdot h_{t+}c) \tag{4}$$

其中，$\boldsymbol{X}$ 为输入，h 为隐藏状态，O 为输出，f 和 g 为非线性函数，W 为序列上一时刻对下一时刻的影响参数。双向 RNN 结构中，增加了后序序列隐状态对当前隐状态的影响，如图 5 所示，BiRNN 网络适用于对上下文相关的序列进行学习和评估，能有效提升效果，网络结构如图 4 所示。基本神经元的 RNN 网络主要受到相邻神经元的影响，容易出现梯度消失和梯度爆炸问题，难以学习到序列中步间距离较大的依赖关系。

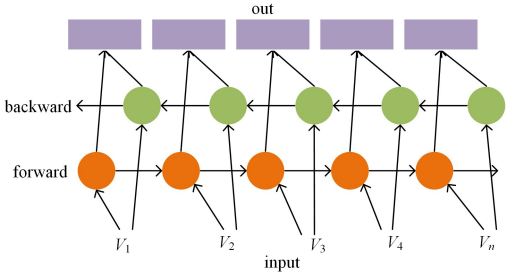

图 4 双向 RNN 网络结构

Fig. 4 Bidirectional RNN network architecture

长短期记忆网络(LSTM)[19]是一种改进型的 RNN，神经元组成包含遗忘门、输入门、输出门和记忆单元，使用 tanh 和 sigmoid 激活函数进行控制，LSTM 通过门控机制选择性地保留重要信息，并遗忘非重要信息，可有效避开 RNN 的弱点，针对序列性数据如文本、语音等类型数据的识别与预测有良好的效果。门控循环单元(GRU)[20]是一种简化的 LSTM，GRU 神经元中有重置门(resetgate)和更新门(updategate)，重置门控制如何将新的输入信息与记忆信息相结合，更新门控制记忆信息保存到当前步的量，通过参数控制可以将较远距离的信息传到当前位置而更近的信息可被舍弃。其与 LSTM 模型效果接近而参数更少、计算效率更高，LSTM 和 GRU 与 RNN 网络结构保持一致，GRU 神经元结构如图 5 所示。

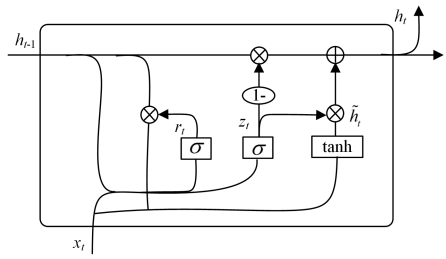

图 5 GRU 神经元节点

Fig. 5 GRU neuron node

GRU 神经元节点函数关系如式(5)－式(8)所示，分别是更新门、重置门、候选隐藏状态和隐藏状态。

$$z_t=\sigma(W_z\cdot[h_{t-1},x_t]+b_z) \tag{5}$$

$$r_t=\sigma(W_r\cdot[h_{t-1},x_t]+b_r \tag{6}$$

$$\tilde{h}_t=\tanh(W_h\cdot[r_t\times h_{t-1},x_t]+b_h \tag{7}$$

$$h_t=(1-z_t)\times h_{t-1}+z_t\times\tilde{h}_t \tag{8}$$

3.4 算法处理和优化

深度学习算法容易出现过拟合、学习速度慢、泛化不佳等问题，本文采用 Mini-batch，RMSprop，Dropout，CrossEntropy Loss，softmax 来提升效果和性能。

3.4.1 Mini-batch 随机梯度下降

批量梯度下降算法需要完整数据集进行一轮训练才能进行一次权值更新，收敛较慢；随机梯度下降算法会导致收敛到局部最优而非全局最优；Mini-batch 随机梯度下降算法具有较快的收敛速度并能获得全局最优，每批次选取一个 batch-size 大小的样本数进行训练。

3.4.2 RMSprop 梯度下降优化算法

传统梯度下降法中学习率存在过大或过小的问题，RM-

SProp 算法是梯度下降法的一种改进算法，可以实现学习率自适应调整，变化较大的梯度分量上的学习率会自动减小，变化较小的梯度分量上的学习率会自动增大，在变化的位置保持注意力。

3.4.3 Dropout

使用 dropout 将网络神经元隐藏层激活函数以一定的概率置为 0，迫使网络去学习更加鲁棒的特征，减少特征之间的依赖关系，使模型泛化能力更强，减少对局部特征的依赖。

3.4.4 Softmax

softmax 可以将一个数值向量归一化为一个概率分布向量，且各个概率之和为 1，用作算法网络最后一层的多分类输出，softmax 函数表示为式(9)：

$$\mathrm{Softmax}(z_i)=\frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{9}$$

交叉熵损失函数 *CrossEntropy Loss* 能够衡量同一个随机变量中的两个不同概率分布的差异程度，在机器学习中表示为真实概率分布与预测概率分布之间的差异。信息熵表达式如式(10)所示：

$$H(X)=-\sum_{i=1}^{n} p(x_i)\log(p(x_i)) \tag{10}$$

其中，$p(x_i)$表示 X 分类为x_i类别的概率。相对熵也称 KL 散度，用于衡量两个概率分布的差异，p 对 q 的相对熵计算式如式(11)所示：

$$D_{\mathrm{KL}}(p\parallel q)=\sum_{i=1}^{n}\log\left(\frac{p(x_i)}{q(x_i)}\right) \tag{11}$$

如果 $p(x_i)$等于 $q(x_i)$，则相对熵为 0。交叉熵为信息熵与相对熵的和，交叉熵的表达式如式(12)所示：

$$H(p,q)=\sum_{i=1}^{n} p(x_i)\log\left(\frac{1}{q(x_i)}\right)=-\sum_{i=1}^{n} p(x_i)\log(q(x_i)) \tag{12}$$

三者的关系是 $D_{\mathrm{KL}}(p\parallel q)=H(p,q)-H(p)$，如果 p 表示真实分布，q 表示预测分布，则交叉熵值越小，预测值越接近真实值，模型预测效果就越好。

损失函数表示的是真实值与神经网络预测值之间的误差，softmax 交叉熵损失函数$L_i=-\sum_i c_i\log(p_i)$表示第 i 个标签的预测损失，一轮预测的偏差用损失的平均值表示，如式(13)所示：

$$Loss=\frac{1}{N}\sum_i L_i=-\frac{1}{N}\sum_i\sum_{j=1}^{m} c_{ij}\log(p_{ij}) \tag{13}$$

其中，m 表示标签类别数，$c_{ij}\log(p_{ij})$表示预测正确的类别，分类标签采用独热编码方便计算。

3.5 算法评价

流量分类算法的有效性可以使用学习训练和预测结果进行评价，对未知数据的预测需要结合其他方法验证。学习训练设定为一个次数或者达到分类损失 Loss 和模型参数不再显著变化为止，划分完整数据集的一部分进行预测测试，训练与测试数据有多种划分方式，通常需要多轮次、交叉综合评估。存在预测结果与标签相符合和不符合的情况，预测与标签符合称为正确，预测与标签不符合称为错误，因此先定义二元分类的混淆矩阵，如表 1 所列，实际类别就是标签类别。

表 1 二元分类混淆矩阵

Table 1 Binary classification confusion matrix

二元混淆矩阵		预测类别	
		Positive(Yes)	Negative(No)
实际类别	Positive(Yes)	TP(true positive)	FN(false negative)
	Negative(No)	FP(falsepositive)	TN(true negative)

TP 和 TN 为预测与实际类别相同，是预测的正确集；FP 表示误报；FN 表示漏报。常用的效果度量指标在此基础上产生，本文采用以下 4 个指标：准确率 ACC(accuracy)、精确度 PRE(precision)、召回率 REC(recall)、F 综合分数(F-Measure)。计算公式为：

$$ACC=\frac{TP+TN}{FP+FN+TP+TN} \tag{14}$$

$$PRE=\frac{TP}{TP+FP} \tag{15}$$

$$REC=\frac{TP}{FN+TP} \tag{16}$$

$$F\text{-}Measure=\frac{(1+\beta^2)\times PRE\times REC}{\beta^2\times(PRE+REC)} \tag{17}$$

样本不均衡时，依据 ACC 进行评价有失偏颇，综合评价指标 F-Measure 综合考虑了 PRE 和 REC，综合反映了查全率和查准率，为了减少样本不均衡的影响，实际应用中主要用 $F1$ 值综合评价训练和测试效果。当 $\beta=1$ 时，*F-Measure* 变为 $F1=2\frac{PRE\times REC}{PRE+REC}$，$F1$ 为准确率和召回率的调和平均值，趋近 1 时表示效果较好。

4 实验与评估

本节评估算法实际需要的运行环境框架、数据、参数学习以及预测测试，确定框架下的最优模型，并与其他方法进行比较。

4.1 网络流量实验数据

采用开放的数据集有利于算法的对比评价，而且数据质量经受过一定的检验，可靠性较高，公开的网络流量数据集一般包括原始数据和标签两部分。知名的异常和攻击检测数据集包括 KDD Cup99 数据集、NSL-KDD 数据集、ISCX2012 数据集、UNSW-NB15 数据集、CIC-IDS2017 数据集等，包含正常流量和攻击流量的数据集可直接使用，几种常用的网络公开数据集如表 2 所列。

各数据集中普遍存在冗余和重复数据，要得到较为可靠的结果需要对数据进行再处理。ISCX-IDS2012，UNSW-NB15，CIC-IDS2017 数据集规模和数据特征、特点较适合本文算法处理领域，但 ISCX-IDS2012 和 UNSW-NB15 协议覆盖和标签全面性不及 CIC-IDS2017，且包和流特征没有全覆盖。CIC-IDS2017 数据集由加拿大网络安全研究所(CIC)创建，数据较新、类别较多，流量比例接近普通使用环境，包含了包和流两种特征标签，环境为普通网络环境。与其他经典数据集相比有更多的应用层攻击行为，其包含种类齐全的加密流量及标签，数据集产生的环境包括真实网络的设备和拓扑结构，仿真度高，因此本文选择更有代表性的 CIC-IDS2017 来研究常态下的网络流量异常分类。CIC-IDS2017 数据集特征分类 CSV 文件保存特征信息和标签信息[1)]，原始流量以 pcap

1) https://www.unb.ca/cic/datasets/ids-2017.html

格式保存，共包括5天产生的数据，大小超过50GB，定义了超过80个特征和14个类别标签。该数据集包括加密和不加密的普通流量，DoS，DDoS，Brute Force，XSS，SQL Injection，Infiltration，Port scan and Botnet等攻击类别。B-Profile system (Benign Profile Agent)系统产生普通流量，涵盖HTTP，HTTPS，FTP，SSH和SMTP等协议的用户行为；暴力攻击流量使用Patator生成FTP和SSH两种协议；DoS(Denial of Service)攻击由GoldenEye，Hulk，Slowhttptest和slowloris 4种工具生成；心血漏洞(Heartbleed)利用OpenSSL库的心跳窃取服务器的内部数据，由Heartleech生成；Web攻击包括SQL注入(SQL Injection)、跨站脚本(XSS)和HTTP暴力攻击；渗透攻击(Infiltration)数据集使用Metasploit产生；僵尸网络代码由Ares工具生成；DDoS(Distributed Denial of Service)数据集使用LOIC(Low Orbit Ion Canon)工具向受害服务器发送UDP，TCP，HTTP请求；PortScan产生端口扫描数据。

表2 入侵检测公开数据集对比

Table 2 Comparison of public data sets for intrusion detection

数据集	年份	数量	数据形式	来源	协议或攻击类别	评价
KDD CUP99	1999	500万	包	实验环境，DARPA网络	Normal\Dos\R2L\U2R\Probe	测试集和训练集概率分布不同；时间过久；主要是网络层，无https加密流量；覆盖面不足，有大量的冗余和重复数据
NSL-KDD	2009	15万	混合(有些流、有些包)	实验环境，DARPA网络	Normal\Dos\R2L\U2R\Probe	改进了KDD CUP99，数据集划分比例平衡，去除了冗余和重复，但仍然缺乏当前真实网络类似的类别
ISCX-IDS2012	2012	245万	混合(有些流、有些包)	实验环境，物联网环境	Normal\DoS\DDoS\Bruteforce\Infiltration	攻击种类多，但无https协议
UNSW-NB15	2015	254万	包	实验环境，IXIA获取网络流量，Argus Bro-IDS等工具筛选	Normal\DoS\Fuzzers\Analysis\Backdoors\Exploits\Generic\Reconnaissance\Shellcode\Worms	攻击种类多，但无https协议
CIC-IDS2017	2018	283万	双向流，包	实验环境，B-Profile系统和其他攻击工具	普通和各层次攻击14个类别标签	覆盖齐全，有网络层、应用层的各种攻击，数据存在重复和冗余
BoT-IoT	2019	7300万	包	真实数据，物联网设备数据	DDoS\DoS\OS\servicescan\keyrecord\dataexplore	数据非通用网络产生而是由物联网产生

4.2 流调整、合并与清洗

数据集已经提供了CICFlowMeter进行的流划分及对应标记，网络流量数据使用不同的流拆分工具得到的流划分会有一些区别，因此利用pkt2flow按照五元组(源IP，目的IP，源端口，目的端口和传输层协议)对CIC-IDS2017数据进行流拆分并和获取划分进行比较。拆分流存在长度不同、超时不同、流截断等情况的区别，对照原始数据，定义流合并和流去重，数据清洗删除不合逻辑以及过短的的一些流拆分结果。数据集搭建环境使用工具模拟生成，其MAC地址、IP地址、TTL值会形成标签暴露，将数据源和标签联系造成过拟合，因此将这几个字段进行匿名化处理，结合CICFlowMeter定义的标签，将条目过少、类别性质相近的标签进行合并，最终获得的标签流量总条数为1 594 756条，共分为7个标签类别，其中普通流量1 357 468条，暴力攻击6 952条，DoS攻击23 996条，Web攻击2 032条，僵尸网络1 235条，DDoS攻击44 655条，端口扫描158 418条，各类别占比如图6所示，训练集和测试集按照7∶3的比例划分。

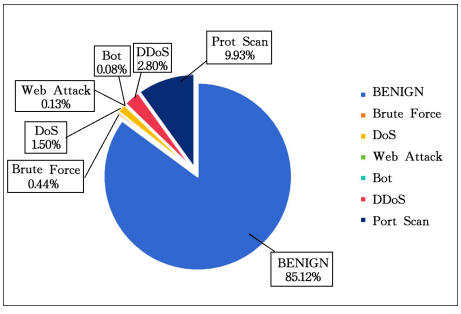

图6 数据集网络流量类别比例

Fig. 6 Proportion of dataset network traffic categories

4.3 算法实验过程

深度学习框架采用Tensorflow 1.14.0和Keras 2.2.4，操作系统Ubuntu 18.10，CPU为Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz，内存16GB。

网格化搜索算法最优参数设置如表3所列，搜索比较部分包括CNN核、CNN层数、CNN激励函数、GRU神经元数等，分类数为7，最优结构参数采用了一个较小的、与完整数据集同分布比例的数据集进行。

表3 网格化搜索算法网络结构参数设置

Table 3 Network structure parameter settings of meshing search algorithm

选项	可选值
CNN网络数	1,2
CNN卷积核	3,5,7
CNN神经元	32,64,96,128,160,256,320,384,512
CNN层数	1,2
激活函数	ReLU,tanh
GRU神经元	50,100,150,200,250

经过10次独立训练测试，并取平均值进行比较，选取的CNN卷积网络参数设置如表4所列，选取的BiGRU神经元个数为50。

表4 CNN参数设置

Table 4 CNN parameter settings

层数	种类	过滤器/神经元	卷积核	步数	填充
1	二维卷积+ReLU	192	7	1	Same
2	二维最大池	2	--	2	Same
3	二维卷积+ReLU	384	5	1	Same
4	二维最大池	2	—	2	Same
5	全局池化	—	—	—	None
6	全连接层	256	—	—	None

算法网络 Minibatch 设为 128，权重更新采用 RMSprop 优化，各层 dropout 设为 0.25，CNN 激活函数为 ReLU。设置 Epoch 为 10，其训练和测试的 loss 曲线和 accuracy 如图 7、图 8 所示。

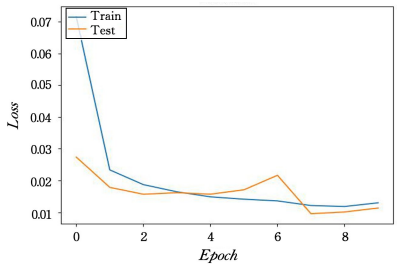

图 7　loss 值变化曲线

Fig. 7　Curves of loss value changes

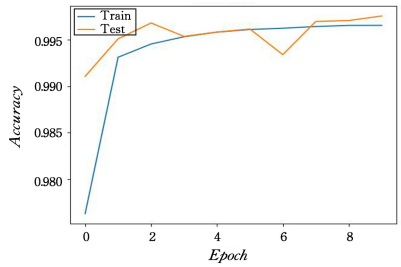

图 8　ACC 值变化曲线

Fig. 8　Curves of ACC value changes

对候选的测试标签数据进行分类测试，7 种类别的预测结果如表 5 所列，并给出了包括 ACC，PRE，REC，F1 在内的 4 个分类评价指标。

从测试结果可以看出，多数标签类型的流量识别效果较好，但有的类别如 Bot 识别准确率偏低。攻击流量之间互相分类混淆的情况较少，主要的误报和漏报发生在普通流量和各种攻击流量之间。由于普通流量占比较高，因此误报漏报情况对普通流量占比较小，普通流量预测的高准确率一定程度上决定了数据的整体准确率较高，对数据集识别分类的效果主要参考精确度、召回率、F1 等指标。

表 5　针对数据的分类别分类结果统计及指标

Table 5　Statistics and indicators of classification results by data categories

标签	预测标签							ACC/%	PRE/%	REC/%	F1
	普通流量	暴力攻击	DoS	Web 攻击	僵尸网络	DDoS	端口扫描				
普通流量	407271	0	1	1	2	16	12	99.95	99.97	99.99	0.9998
暴力攻击	0	2085	0	0	0	0	0	100.00	100.00	100.00	1.0000
DoS	1	0	7198	0	0	0	0	100.00	99.99	99.99	0.9999
Web 攻击	0	0	0	609	0	0	0	100.00	99.84	100.00	0.9992
僵尸网络	44	0	0	0	326	0	0	99.99	99.39	88.11	0.9341
DDoS	14	0	0	0	0	13397	0	99.95	99.88	99.90	0.9989
端口扫描	62	0	0	0	0	0	47617	99.98	99.97	99.87	0.9992

4.4　结果和对比

整体检测效果如表 6 所列，对表 3 中的数据进行类别算术平均和加权平均，加权平均包含了分类数据的数量信息。

表 6　完整测试数据分类结果

Table 6　Complete test data classification results

(%)

汇总	ACC	PRE	REC	F1
算术平均	99.98	99.86	98.24	0.9900
加权平均	99.95	99.95	99.95	0.9995

文献[21]对 CIC-IDS2017 数据集进行了各种机器学习分类测试，涵盖了 ANN，DT，KNN，NB，RF，SVM，CNN 等经典机器学习算法和深度学习算法，与本文提出的 CNNBiGRU 算法的检测结果对比如表 7、图 9 所示。

表 7　文献[21]所列分类算法与本文算法的效果对比

Table 7　Comparison of the effect of classification algorithms in reference [21] and the proposed algorithm

(%)

	ACC	PRE	REC	F1
CNNBiGRU	99.95	99.97	99.97	0.9997
ANN	99.31	99.50	99.31	0.9922
DT	99.49	99.49	99.49	0.9949
KNN	99.49	99.50	99.49	0.9949
NB	98.86	99.01	98.86	0.9885
RF	99.54	99.56	99.54	0.9955
SVM	99.72	99.27	96.72	0.9789
CNN	99.50	99.46	99.50	0.9947

特征工程统计类机器学习算法使用的是标记的特征信息进行分类预测，CNNBiGRU 算法使用的是原始流量输入进行的分类测试。实验对比结果表明，本文算法在准确率 ACC、精确度 PRE、召回率 REC、综合 F1 值几个方面都占据了优势，处于领先的地位。

4.4.1　七分类的全体数据分类对比

文献[10]和文献[14]中的 CNN 结合 RNN 系列算法属于神经网络融合类算法，综合考虑了网络流空间特征和时间特征，本文算法与典型的融合算法 HDST_IDS 和 CNNBiLSTM 进行比较，另外将文章的算法去除 GRU 模块进行对比，效果对比如图 9 所示。3 种融合算法的准确率和 F1 值比较接近，均达到了 99.9%以上的高准确率，分类效果高于单一 CNN 的神经网络算法。

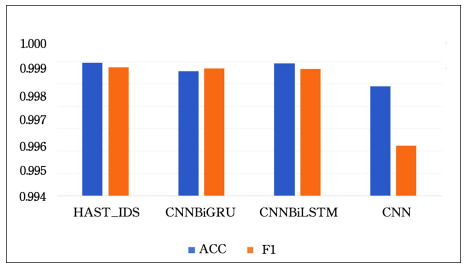

图 9　本文算法与文献[10,14]算法效果对比

Fig. 9　Comparison of the effect of the proposed algorithm and algorithms in literatures [10,14]

以学习准确率 99%为目标衡量算法效率，本文算法使用较少的权值迭代步数就可以达到目标准确率，文献[14]算法

的效率为本文算法的 69%，文献[10]算法效率为本文算法的 78%，本文算法的消融算法纯 CNN 效率为本文算法的 90%。如图 10 所示，在此指标下本文算法相比另外 3 种算法具有一定的效率优势。

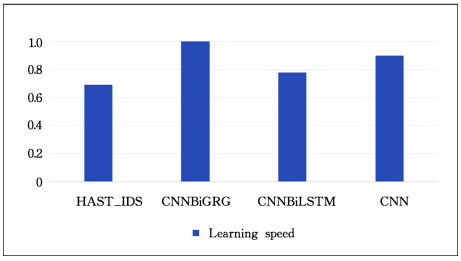

图 10 本文算法与文献[10]、文献[14]以及 CNN 的训练速度对比

Fig. 10 Comparison of the training speed of the proposed algorithm and algorithms in reference [10,14] and CNN

4.4.2 四分类的子集数据分类对比

使用星期五数据进行重新测试，验证本文算法对于模拟真实环境数据的有效性。经过预处理获得流数量 442433 条，普通流 53.66%，僵尸网络 0.29%，DDos 10.15%，端口扫描 35.90%，共 4 种类别流量。学习准确率、F1 值对比结果如图 11 所示，学习准确率 98%目标下速度对比如图 12 所示。

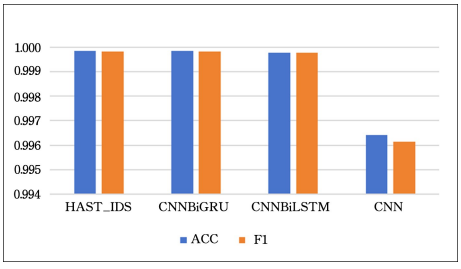

图 11 星期五数据测试效果对比

Fig. 11 Friday data test effect comparison

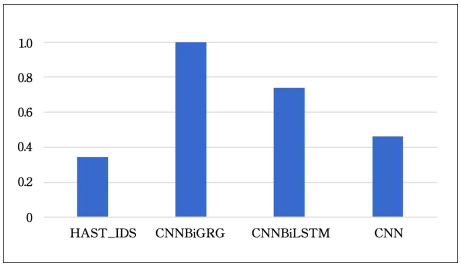

图 12 星期五数据 98%准确率学习速度对比

Fig. 12 Friday data learning speed comparison with accuracy of 98%

结束语 针对传统的网络恶意流量分类方法如端口协议分类法和深度包检测网络流量分类法不能解决的动态、混淆网络、隐私风险和加密网络等情况，本文提出了一种融合 CNN 和 RNN 网络的恶意流量分类算法 CNNBiGRU 来解决恶意流量分类问题，并对网络中神经元和层次采取了适当的过程优化，完善了算法网络结构。使用包含普通流量和攻击流量、种类齐全的数据集 CIC-IDS2017，并进行数据质量分析、去重和预处理[21-22]，通过网格化搜索寻找网络各层和组合的最优参数，并在 Tensorflow 和 Keras 框架下进行了神经网络融合算法的分类训练和分类测试。相比传统的经典机器学习算法，CNNBiGRU 的分类效果具有一定的优势，无需特征工程；本文算法对全局数据进行七分类测试，其准确率和 F1 值与 CNN 和 LSTM 融合算法相当，高于单独的 CNN 算法，且本算法具有训练达成目标迭代更快的优势；采用仅星期五数据四分类进行测试，本算法也具有类似的结论。如果要提高对占比较小的异常流量的识别率，可以通过引入生成对抗网络、类别权重调整、分类组合数据集、特征重要度区分等方法解决类别不平衡问题，提高异常流量在模型中的影响力，在相同策略下对算法进行测试及对比。

参 考 文 献

[1] China Internet Network Security Report 2020 [R]. https://www.cert.org.cn/publish/main/upload/File/2020%20Annual%20Report.pdf.

[2] MOORE A W, PAPAGIANNAKI K. Toward the accurate identification of network ap-plications[C]// PAM 2005: Proceedings of the 2005 International Workshop on Passive and Active Network Measurement, LNCS 3431. Berlin: Springer, 2005: 41-45.

[3] GU Y, LI D, GAO K G. Research on Network traffic Classification based on Machine Learning and Deep Learning[J]. Telecommunication Science, 2021, 37(3): 105-113.

[4] KONG L, HUANG G, WU K. Identification of Abnormal Network Traffic Using Support Vector Machine[C]// 2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies(PDCAT). 2017: 288-292.

[5] IMAN S, LASHKARI H, GHORBANI A, et al. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization[C]// International Conference on Information Systems Security and Privacy. 2018: 108-116.

[6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[7] GRAVES A, MOHAMED A R, HINTON G E. Speech recognition with deep recurrent neural networks[C]// 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. 2013: 6645-6649.

[8] REZAEI S, LIU X. Deep learning for encrypted traffic classification: An overview[J]. IEEE Communications Magazine, 57(5): 2019: 76-81.

[9] LOTFOLLAHI M, JAFARI SIAVOSHANI M, SHIRALI HOSSEIN ZADE R, et al. Deep packet: a novel approach for encrypted traffic classification using deep learning[J]. Soft Computing, 2020, 24(3): 1999-2012.

[10] LOPEZ-MARTIN M, CARRO B, SANCHEZ-ESGUEVILLAS A, et al. Network traffic classifier with convolutional and recurrent neural networks for Internet of Things[J]. IEEE Access, 2017(5): 18042-18050.

[11] WANG W, ZHU M, ZENG X W, et al. Malware traffic classification using convolutional neural network for representation learning[C]// 2017 International Conference on Information Networking(ICOIN). Da Nang, Vietnam, 2017: 712-717.

[12] CIREGAN D, MEIER U, SCHMIDHUBER J. Multi-column

deep neural networks for image classification[C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence,RI,USA,2012:3642-3649.

[13] LECUN Y,JACKEL L D,BOTTOU L,et al. Learning Algorithms for Classification:A Comparison on Handwritten Digit Recognition[C]// Neural Networks:The Statistical Mechanics Perspective. 1995.

[14] WANG W,SHENG Y Q,WANG J L,et al. HAST-IDS:Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection[J]. IEEE Access,2018(6):1792-1806.

[15] LIU Y F,CAI S,YANG H X,et al. Network Intrusion Detection Method Integrating CNN and BiLSTM[J]. Computer Engineering,2019,45(12):127-133.

[16] DENG X,LIU Z H,OUYANG Y,et al. Identification of encrypted Malicious traffic based on CNN CBAM-BiGRU Attention[J]. Computer Engineering,2023,49(11):178-186.

[17] PACHECO F,EXPOSITO E,GINESTE M,et al. Towards the Deployment of Machine Learning Solutions in Network Traffic Classification:A Systematic Survey[J]. IEEE Communications Surveys & Tutorials,Secondquarter 2019,21(2):1988-2014.

[18] ZHOU F Y,JIN L P,DONG J. Review of Convolutional neural network[J]. Chinese Journal of Computers,2017,40(6):1229-1251.

[19] HOCHREITER S,SCHMIDHUBER J. Long short-term memory[J]. arXiv:1412. 3555,2014.

[20] CHUNG J Y,GULCEHRE C,CHO K,et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling[J]arXiv:1412. 3555,2014.

[21] OYELAKIN A,AMEEN A O,OGUNDELE T S,et al. Overview and Exploratory Analyses of CICIDS 2017 Intrusion Detection Dataset[J/OL]. https://api. semanticscholar. org/CorpusID:262063000.

[22] MASEER Z K,YUSOF R,BAHAMAN N,et al. Benchmarking of Machine Learning for Anomaly Based Intrusion Detection Systems in the CICIDS2017 Dataset[J]. IEEE Access,2021(9):22351-22370.

YANG Yongping, born in 1980, master, lecturer. His main research interests include network security and machine learning.