第 9 卷 第 2 期
2020 年 3 月
集 成 技 术
JOURNAL OF INTEGRATION TECHNOLOGY
Vol. 9 No. 2
Mar. 2020

引文格式：
黄璇丽, 李成明, 姜青山. 基于深度学习的网络流时空特征自动提取方法 [J]. 集成技术, 2020, 9(2): 60-69.
Huang XL, Li CM, Jiang QS. A deep learning-based spatio-temporal features extraction method for network flow [J]. Journal of Integration Technology, 2020, 9(2): 60-69.

基于深度学习的网络流时空特征自动提取方法

黄璇丽[1,2] 李成明[1] 姜青山[1]

[1](中国科学院深圳先进技术研究院 深圳 518055)
[2](中国科学院大学深圳先进技术学院 深圳 518055)

摘 要 流量异常检测是网络入侵检测的主要途径之一，也是网络安全领域的一个热门研究方向。通过对网络流量进行实时监控，可及时有效地对网络异常进行预警。目前，网络流量异常检测方法主要分为基于规则和基于特征工程的方法，但现有方法需针对网络流量特征的变化需重新人工收集规则或构造特征，工作量大且繁杂。为解决上述问题，该文提出一种基于卷积神经网络和循环神经网络的深度学习方法来自动提取网络流量的时空特征，可同时提取不同数据包之间的时序特征和同一数据包内字节流的空间特征，并减少了大量的人工工作。在 MAWILab 网络轨迹数据集上进行的验证分析结果表明，该文所提出的网络流时空特征提取方法优于已有的深度表示学习方法。

关键词 网络流量；网络入侵检测；卷积神经网络；循环神经网络；时空特征提取
中图分类号 TP 399 **文献标志码** A doi: 10.12146/j.issn.2095-3135.20191231002

A Deep Learning-Based Spatio-Temporal Features Extraction Method for Network Flow

HUANG Xuanli[1,2] LI Chengming[1] JIANG Qingshan[1]

[1](*Shenzhen Institutes of Advanced Technology*, *Chinese Academy of Sciences*, *Shenzhen* 518055, *China*)
[2](*Shenzhen College of Advanced Technology*, *University of Chinese Academy of Sciences*, *Shenzhen* 518055, *China*)

Abstract Network intrusion detection is one of the core research areas of cyber security. Network traffic anomaly detection is common in network intrusion detection systems. Through monitoring the network traffic, network intrusion detection systems can effectively track anomalous traffic and then give out alerts. This research area has developed for decades and the conventional methods for network intrusion detection systems include rule-based and feature engineering based methods. However, the changing features of network

收稿日期：2019-12-31 **修回日期**：2020-02-18
基金项目：广东省重点领域研发计划项目(2019B0101137002)；深圳市基础研究项目(JCYJ20180302145607677)
作者简介：黄璇丽，硕士研究生，研究方向为机器学习和信息安全；李成明(通讯作者)，副研究员，研究方向为未来中心网络，E-mail：cm.li@siat.ac.cn；姜青山，研究员，研究方向为数据挖掘。

traffic require the methods to continuously gather new rules and generate new features, which results in a labor-intensive workload and comparatively poor quality of features engineering. To solve this problem, a deep learning-based spatial-temporal features extraction method was proposed. It includes convolution neural networks and long short term memory neural networks to learn the spatial-temporal features of network raw traffic. This method is tested on the MAWILab network traces data to evaluate its effectiveness. Multi-layer perception, convolution neural networks alone and long short term memory are used for comparison with the proposed approach. The features generated by these methods are used to classify the traffic, which can assess the performance of the feature extraction process of each method. Experiments show that the proposed method outperforms other methods in its effectiveness of spatial-temporal features extraction.

Keywords　network traffic; network intrusion detection; convolution neural networks; recurrent neural networks; spatio-temporal features extraction

1　引　言

网络入侵检测系统(Network-based Intrusion Detection System，NIDS)相关技术在万物互联的时代是不可或缺的，也是网络信息安全中一个重要的研究领域。入侵检测系统包括主机入侵检测系统和网络入侵检测系统。其中，网络入侵检测系统通过对网络上的流量进行监控，并实时对异常流量发出预警，从而提高网络的安全性[1]。

网络流量分析是网络异常检测的重要方法，传统的检测方法包含基于规则和基于特征工程的方法[2]。其中，基于规则的检测方法需要网络安全专家针对已有入侵行为，生成对应的规则进行匹配检测，不具备检测新的入侵行为的能力。基于传统机器学习的方法需要特征工程等人为地构造特征，然后训练机器学习模型，模型的效果很大程度上取决于特征工程的质量[3]。目前，深度学习已在计算机视觉、自然语言处理、推荐系统、网络流量异常检测等领域广泛应用[4-5]。然而，已有的基于深度学习网络流量异常检测方法，只是针对网络流单一的时序特征或空间特征进行了提取，缺少对网络流量时空特征的综合表示。

原始的网络流量是由按照网络协议规定格式的一串字节组成的。多个流量字节组合成一个数据包，通信双方的多个数据包则组成一个网络流[1]。其中，数据包以一个整体同时在网络上传输，故其内部的流量字节没有太多的时序关系，但字节间被认为存在着空间关系，可提取其相应的空间特征。而网络流中的每个数据包有不同的发送时间，被认为存在时序关系，可提取其相应的时序特征。因此，空间特征和时序特征是网络流量监测领域常用的两类流量特征[1-2]。

针对网络流所具有的时空特征，本文提出一种基于深度学习的网络字节流数据时空特征提取方法。其中，采用卷积神经网络结构提取网络流量的空间特征，采用循环神经网络结构提取网络流量的时序特征。本文研究的基本单元是网络流，对于原始网络轨迹流量，需将其切分为以网络流为单位的数据集。其中，每一条网络流包含一组双方通信的数据包，每个数据包包含一组字节(大小为 0～255)。实验使用的数据集是 MAWILab 网络轨迹数据集，并将原始的网络轨迹流量切分为以网络流为单位的数据集合，结合日志文件生成带标签的网络字节流数据。实验结果表明，本文所提出的网络流时空特征提取方法

优于已有的深度表示学习方法。

2 网络流量检测方法研究现状

国内外研究学者对网络流量分析的问题研究了近二十年，已有许多网络入侵检测相关的研究工作[2]。入侵检测系统应用广泛，常应用于工业系统、运输系统、医疗系统和建筑系统[4-5]。根据 Ahmed 等[6]的研究成果，网络流量异常检测方法可以分为四种：基于分类、基于统计、基于聚类和基于信息论。Chalapathy 等[5]研究重点为如何使用深度学习技术进行异常检测，其中包括将深度学习技术应用到主机入侵检测和网络入侵检测中。针对现有检测方法严重依赖人工制定规则、人工收集标签的问题，Nisioti 等[7]对无监督的网络流量异常检测进行了研究。此外，由于对抗机器学习的兴起，也有一些学者将其应用到入侵检测系统攻击相关的研究工作[8-9]。网络流量检测通常划分为基于规则、基于特征工程和基于特征学习 3 种方法，如图 1 所示。

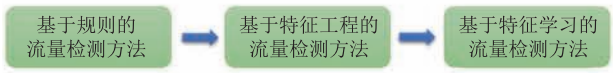

图 1 网络流量检测方法分类

Fig. 1 Categories of network traffic detection techniques

基于规则的网络流量检测方法主要包括：基于 DPI(Deep Packet Inspection)和基于端口[10]。对先前的网络流量人工制定规则，通过匹配规则来检测已出现的异常。该方法只能检测已出现过的网络异常，对没有出现过的异常难以检测；同时制定规则需要很多的人力资源，人力成本极高，无法适用于当前网络的急速发展。

基于特征工程的网络流量检测方法分两步：(1)使用特征工程进行特征组合、特征选择等构造特征集；(2)基于特征集使用传统机器学习和深度学习模型进行检测。其中，常用的网络流量特征有数据包个数、网络应用层协议和网络流量长度等统计特征和类别；常用的机器学习方法有线性回归、逻辑回归、决策树、随机森林、支持向量机和多层感知机等[11]。Aygun 和 Yavuz[12]使用自动编解码器进行流量检测。基于异常检测进行入侵检测的工作也有不少，如 An 和 Cho[13]使用变分自动编码器进行异常检测；Intrator 等[14]使用多个判别器的对抗生成网络(GAN)进行异常检测；Zhou 和 Paffenroth[15]使用鲁棒深度自动编码器去噪，再进行异常检测；Zhao 等[16]和 Lin 等[17]分别采用降维和聚类等无监督方法进行入侵检测。基于特征工程的方法常用的基准数据集有 KDD99 和 NSL-KDD 等，这类方法的检测效果依赖特征工程的质量，需要人工经验和特征工程技巧，在网络互联和大数据时代无法适用。

基于深度特征学习的网络流量检测方法，需要使用深度特征学习模型进行自动地提取特征。网络检测的基本单元是网络流，而网络流量底层归根到底是一串字节。依照网络协议规定，将流量字节组合成数据包，数据包再组成网络流，其中，数据包里的字节间存在着空间关系，而网络流中的每个数据包间又存在时序关系。针对原始的网络流量，构建合适的深度特征学习模型，即可学习网络流量的时间特征和空间特征。Wang 等[18-19]使用卷积神经网络学习空间特征分别进行加密流量分类和恶意流量分类相关的流量检测研究；Mirza 等[20]和 Shiravi 等[21]对原始流量数据 ISCX IDS 2012，建立循环神经网络进行特征学习以提高检测性能。

3 基于深度学习的网络字节流时空特征提取方法

基于深度特征学习的网络流量检测方法通过深度特征学习模型进行自动地提取特征，再使用提取的特征进行检测。基于深度学习的特征提取方法主要包括卷积神经网络结构和循环神经网络结构，

可有效提取网络流量时空特征。

3.1　网络流

按照网络协议的规定，多个字节组成数据包，通信双方的数据包组成网络流，网络流携带着数据在不同的计算机之间传输，字节就是网络流量的原始形态。网络流(flow)、数据包(packet)和字节(byte)数据的层次关系如图 2 所示。

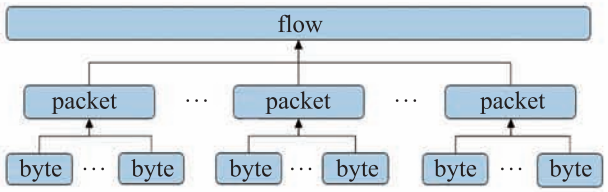

图 2　网络流的层次关系图

Fig. 2　The hierarchy of network traffic flow

原始的网络轨迹是由一串字节组成的，网络流是网络流量检测的基本单位。图 3 所示为某个网络流样例：该条网络流由 30 个数据包组成，一个数据包由多个字节组成。

一个网络流样本由网络流的字节向量和标签组成，其样本集合 D 的表达如公式(1)所示。

$$D=[(B_1,Y_1),(B_2,Y_2),\cdots,(B_k,Y_k)] \tag{1}$$

其中，Y_k 表示第 k 个网络流的标签；$\boldsymbol{B}_k=[b_1,b_2,\cdots,b_{m\times n}]^{\mathrm{T}}$为字节向量。$m$ 是一条网络流中数据包的数量，n 是一个数据包中包含字节的数量，因此一条网络流的字节数量为 $m\times n$，即字节向量的长度。

3.2　网络流时空特征提取方法

从网络流的层次结构可知，数据包之间存在着显著的时序特征，数据包内的字节被认为存在着空间特征。因此，本文设计了如图 4 所示的特征提取方法，其中输入为网络流字节向量(flow bytes vector)，包含卷积层(conv1 和 conv2)、最大池化层(maxpool1 和 maxpool2)、全连接层(Full Connection，FC)、长短期记忆(Long-Short Term Memory，LSTM)和 softmax 层，其主要的流程分为以下 3 步。

(1)网络流量空间特征学习过程：由于卷积神经网络可以学习每个数据包内的空间特征，故本文采用卷积神经网络来学习原始流量的空间特征。该网络结构使用两层卷积层和两层最大池化层，可学习到网络流量数据的局部特征，输出为空间特征表示 h_1，之后将其应用于网络的深层结构以学习更多的全局特征。

(2)网络流量时序特征学习过程：对于步骤(1)中得到的空间特征表示 h_1，由网络流量的内部层次结构可知存在着时序特征。本文采用一个

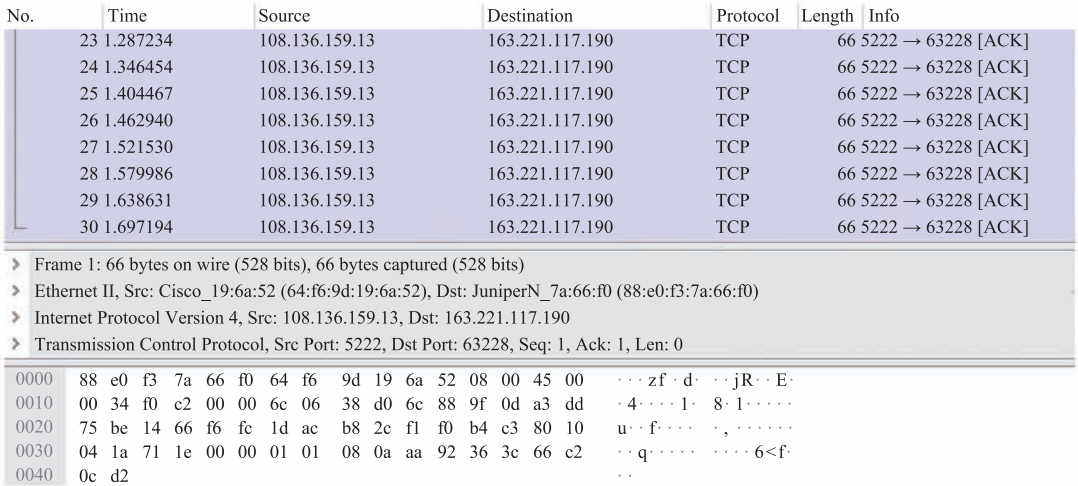

图 3　网络流样例图

Fig. 3　Network traffic flow sample

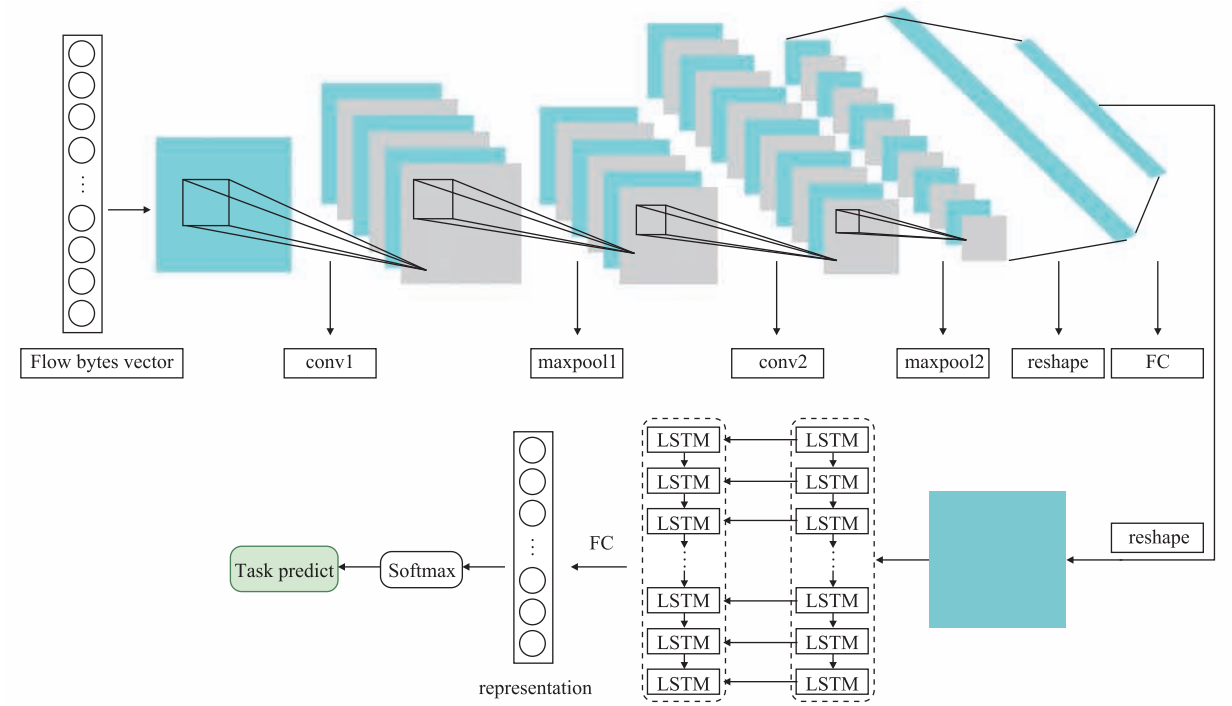

图 4 网络字节流数据时空特征提取方法

Fig.4 Spatio-temporal features extraction method for network traffic flow

LSTM 网络来进一步学习原始流量的时序特征，该网络结构使用两层隐藏层，输出为时空特征表示 h_2。

(3)网络流量检测过程：对于经过步骤(1)和(2)得到的网络流时空特征表示 h_2，设计一个全连接层和输出层，使用 softmax 分类器进行分类检测，输出网络流所属类别的预测概率。

网络流量的内部结构揭示其存在着的时序特征和空间特征。本文中，需先将原始网络流量的字节流向量变形成大小为 $n\times n\times channels$ 的张量，再作为时空特征提取网络的输入。其中，$channels$ 取 1。图 4 所示为本文方法的详细框架图。

3.2.1 空间特征提取网络结构

由于卷积神经网络具有空间特征学习的能力，常应用于计算机视觉等相关领域。现有的卷积神经网络通常包括输入层、卷积层、池化层、全连接层和输出层，其稀疏连接和参数共享的独特训练方式是该网络的最大优势[22-23]。其中，卷积层和池化层的设计极为关键，通常底层结构学习数据的局部特征，而网络的深层结构则能得学到全局特征。

本文中采用卷积神经网络结构来提取网络流量的局部空间特征。该网络结构含有两层卷积层，第一层卷积层使用 32 个卷积滤波器(大小为 5×5、步长为 1)作用于局部区域，学习局部特征。给定一个大小为 $d\times d\times channels$ 的张量作为输入，得到的输出张量大小为$(d-4)\times(d-4)\times(channels\times 32)$。第二层卷积层使用 64 个卷积滤波器，大小为 3×3，步长为 1。给定一个大小为 $d\times d\times channels$ 的张量作为输入，得到的输出张量大小为$(d-2)\times(d-2)\times(channels\times 64)$。

假设输入层是大小为 $d\times d\times channels$ 的张量，其卷积操作如公式(2)所示。

$$c_i=f\left[\boldsymbol{w}\cdot x_{i:(i+h-1)}\right] \tag{2}$$

其中，$x_{i:(i+h-1)}$为第 i 行到第 $i+h-1$ 行组成的

$h\times d$ 滑动窗口；$\boldsymbol{w}$ 为权重矩阵。

时空特征提取网络使用了两层最大池化层，即对区域内的特征点取最大值。两层最大池化层采用的池化滤波器大小为 2×2、步长为 2。给定一个大小为 $d\times d\times channels$ 的张量作为输入，得到的输出张量大小为 $\frac{d}{2}\times\frac{d}{2}\times channels$。

3.2.2 时序特征提取网络结构

循环神经网络常用于时间序列相关的任务中，常用的变种有 LSTM 和 GRU（门控循环单元）。传统的神经网络结构是前馈的，即每一层节点之间没有联系，而循环神经网络每一层的参数是共享的，即当前层的输出不仅要考虑上一层的输出，而且还要考虑上一时刻隐藏层的输出[19,24]。因此，隐藏层更新公式如下所示：

$$h_t=f(\boldsymbol{W}h_{t-1}+\boldsymbol{U}x_t) \tag{3}$$

其中，$\boldsymbol{W}$ 为 $t-1$ 时刻到 t 时刻的参数矩阵；$\boldsymbol{U}$ 为输入层到当前层的参数矩阵，是共享的；f 为激活函数。但这种简单的循环神经网络存在梯度消失、梯度爆炸和难以训练的问题[20]。LSTM 是循环神经网络中成功的扩展之一，能有效解决以上的问题，其引入了输入门、遗忘门和输出门，还有记忆单元。LSTM 隐藏层更新公式如下所示：

$$i_t=f(\boldsymbol{W}^i h_{t-1}+\boldsymbol{U}^i x_t) \tag{4}$$

$$f_t=f(\boldsymbol{W}^f h_{t-1}+\boldsymbol{U}^f x_t) \tag{5}$$

$$o_t=f(\boldsymbol{W}^o h_{t-1}+\boldsymbol{U}^o x_t) \tag{6}$$

$$\tilde{c}_t=\tanh(\boldsymbol{W}^c h_{t-1}+\boldsymbol{U}^c x_t) \tag{7}$$

$$c_t=f_t\circ c_{t-1}+i_t\circ\tilde{c}_t \tag{8}$$

$$h_t=o_t\circ\tanh(c_t) \tag{9}$$

其中，i_t、f_t 和 o_t 分别是为输入门、遗忘门和输出门，它们都有循环神经网络中特有的两个参数矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{U}$；tanh 为激活函数，其公式为 $\tanh(x)=\frac{e^x-e^x}{e^x+e^x}$；$c_{t-1}$ 为上一时刻的记忆；$\circ$ 操作为向量对应的元素相乘；$\tilde{c}_t$ 为当前时刻的记忆；c_t 为经过遗忘门和输出门后的记忆；h_{t-1} 和 h_t 分别为上一个单元隐藏层和当前单元隐藏层的输出。

4　结果分析与评估

4.1 数据集与评价指标

实验验证采用的数据集来源于 MAWILab 开放数据[25]。MAWILab 数据集每天收集日本和美国两个服务商节点之间在 14：00—14：15 期间 15 min 的网络流量，并提供日志文件。流量的标签有 anomalous、suspicious、notice 和 benign 四种，分别代表异常流量、可疑流量、通知过的流量和正常流量。

有许多研究者使用 MAWILab 进行网络流量检测研究，如 Kwon 等[26]通过提取 MAWILab 中网络协议等相关字段，人工构造特征后使用卷积神经网络进行检测；Siffer 等[27]使用极值理论进行检测异常流量。本文提出的是网络字节流数据时空特征提取方法，对 MAWILab 原始网络流量数据处理的流程如图 5 所示。

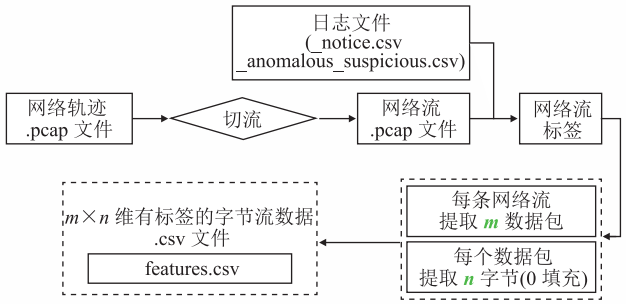

图 5　关于 MAWILab 网络轨迹数据的原始流量处理流程图

Fig. 5　Data processing of MAWILab network traces

本文选取了 2019 年 6 月 15 天的网络流量，每天采样约 1 万条网络流，其中各类别的网络流数量如表 1 所示。对于每一天的网络流量，MAWILab 提供了 .pcap 轨迹文件、_notice.csv 和 _anomalous_suspicious.csv 日志文件[24]。由于网络检测的基本单位是网络流，需先将 .pcap 文件中的原始网络轨迹切为网络流，其中每个网络流对应一个 .pcap 文件；然后，根据源 IP、端

IP、源端口、端端口四元组(srcIP, dstIP, srcPort, dstPort)，在日志文件中检索对应的标签。

表 1 MAWILab 数据集(2019 年 6 月)网络流数据描述

Table 1 Description of MAWILab network traffic flows (June, 2019)

类别	数量
Benign	97 072
Anomalous	31 446
Suspicious	8 611
Notice	11 171

由于神经网络的输入向量是固定长度，所以对于所有网络流，需要截取固定个数的数据包，同时每个数据包截取固定个数的字节，这样组成固定长度的网络流字节。如每条网络流截取 m 个数据包，每个数据包截取 n 个字节，此时多余的字节直接截断，而缺少的则选择零填充，这样得到每条网络流的字节数量都是 $m\times n$。数据处理中选择两种网络流截取：$m=8$、$n=98$ 和 $m=10$、$n=160$，对应的网络流的字节数量分别为 784 和 1 600。输入时空特征提取方法时，分别将其变形为 28×28×1 或 40×40×1 的张量。

本文所使用的评价指标是准确率(accuracy)、带权重的 $F1$(weighted_$F1$)、带权重的召回率(weighted_recall)、带权重的精准率(weighted_precision)。准确率为所有样本预测准确的数量占全部样本总数的比率。由于 MAWILab 网络流量检测为多分类任务，在计算每个类别的 TP(真阳性数)、TN(真阴性数)、FP(假阳性数)和 FN(假阴性数)时，将当前类别视为正样本，其他所有类别视为负样本。各评价指标的计算公式如下所示：

$$weighted\_precision=\sum_{i=1}^{4}\frac{TP_i}{TP_i+FP_i}support(i) \quad (10)$$

$$weighted\_recall=\sum_{i=1}^{4}\frac{TP_i}{TP_i+FN_i}support(i) \quad (11)$$

$$weighted\_F1=\sum_{i=1}^{4}\frac{2\times TP_i}{2\times TP_i+FP_i+FN_i}support(i) \quad (12)$$

其中，$support(i)$为支持度，即当前类别的样本数目占总样本数目的比重；TP_i 为 i 类且判别为 i 类的结果；TN_i 为非 i 类且判别为非 i 类的结果；FP_i 为非 i 类且判别为 i 类的结果；FN_i 为 i 类且判别为非 i 类的结果。

4.2 实验结果

本文使用表 1 的数据进行检测分析，实验数据的网络流长度(即特征维度)有 2 种：784 维和 1 600 维。评价指标为准确率、带权重的 $F1$ 值、带权重的召回率和带权重的精准率。训练网络字节流时空特征提取方法时，将数据集以 8∶2 的比例划分训练集和测试集，实验展示的结果为测试集上的检测结果。

图 6 和图 7 分别是训练 10 次的损失图和效果图。从图 6 可以看出，模型是在收敛的，且在 1 600 维数据上的收敛效果更好。从图 7 可以看出，随着迭代次数的增加，检测效果变好，且模型在 1 600 维数据上的表现更好，这应该是得益于 1 600 维保留的字节数比 784 维的多。

对比实验选择单一网络结构的特征提取方法，包括两层全连接的全连接网络(MLP)、卷积神经网络(CNN)和 LSTM，采用 softmax 层对得到的特征进行分类，分类结果如表 2 所示。

特征提取工作的质量无法直接进行衡量，需将提取到的特征用于相应的任务，通过任务的质量间接体现特征提取的质量。表 2 将不同特征提取的结果使用 softmax 分类器进行分类，以观察所提出方法在 784 维和 1 600 维数据上的准确率、带权重的 $F1$ 值、带权重的召回率、带权重的精准率。可以看出，不管是在 784 维的数据上还是 1 600 维的数据上，提取网络流量空间特征的卷积网络(CNN)方法的分类结果较差，提取网络流量时间特征的 LSTM 方法的分类结果有所提升，本文方法(CNN-LSTM)的分类结果最好。该实验结果表明不同特征对网络流量检测影响不同，其中时空特征比较好，时间特征次之，空间特征比较差。

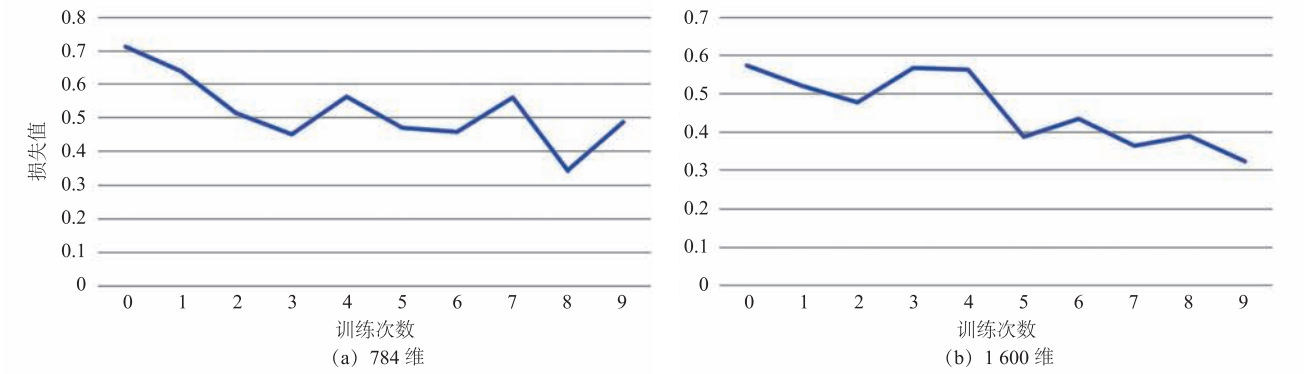

图 6　网络流检测实验损失图(10 次训练)

Fig. 6　Loss graph of network detection experiments (10 trainings)

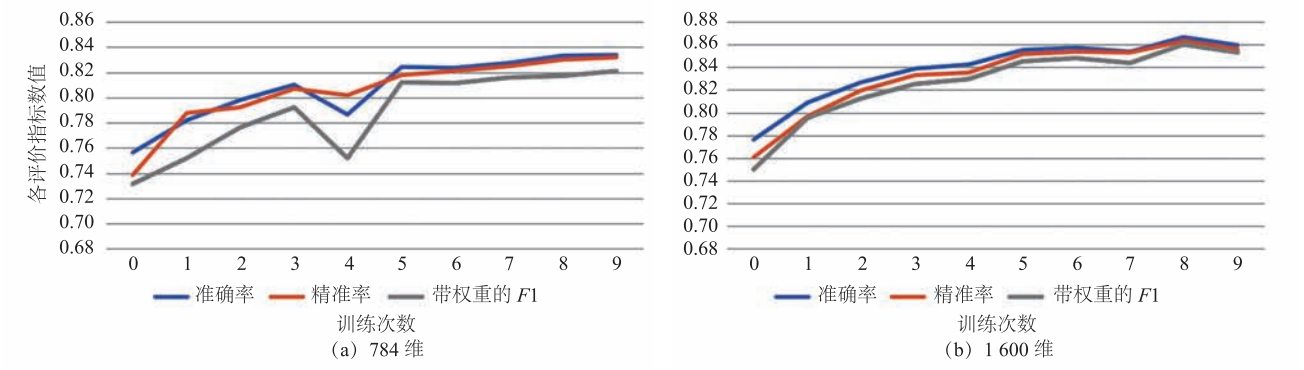

图 7　网络流检测实验效果图(10 次训练)

Fig. 7　Performance of network detection experiments (10 trainings)

表 2　对比实验结果

Table 2　Experimental results

方法	784 维				1 600 维			
	准确率	精准率	召回率	带权重的 $F1$	准确率	精准率	召回率	带权重的 $F1$
MLP	0.738 6	0.716 5	0.738 6	0.709 8	0.742 0	0.720 9	0.742 0	0.714 7
CNN	0.794 5	0.782 0	0.794 5	0.772 9	0.816 3	0.806 2	0.816 3	0.795 6
LSTM	0.816 9	0.803 4	0.816 9	0.801 9	0.823 4	0.811 8	0.823 4	0.807 9
CNN-LSTM（本文）	0.861 8	0.856 9	0.861 8	0.850 1	0.898 5	0.896 8	0.898 5	0.895 9

4.3　讨论与分析

网络流量检测方法可划分为基于规则、基于特征工程和基于特征学习三种。常见的网络流量检测研究先是人工提取相关的特征，再进行模型训练，因此检测效果依赖特征工程等技术的特征提取质量。Kwon 等[26]提取了 MAWILab 数据集中的 29 个特征，使用深度卷积神经网络进行检测，准确率能达 67.86%；Siffer 等[27]使用极值理论进行异常检测，对 MAWILab 数据集进行流量的异常检测时，取得了 86% 的真阳率和低于 4% 的假阳率。网络流量的内部结构揭示了其具有时间和空间的特性，这两类特征也常用于网络流量检测。针对特征依赖和人工经验依赖的问题，本文提出一种基于深度学习的网络字节流时空特征提取方法。与前面方法相比，本文方法节省了人为构造和提取特征的成本，其中在 1 600 维的数

据上进行实验时，准确率达到 89.85%，比 Kwon 等[26]方法的准确率有很大的提升。

5 结 论

针对网络流量特征的自动提取需求，现有的基于规则和特征工程的方法均需人工经验和特征工程技巧，比较繁杂且人工成本高。因此，本文提出一种更自动化的特征提取方法，通过使用深度学习提取网络字节流时空特征，及在网络原始流量数据集上进行特征提取，分类实验结果显示检测结果有较大的提升。这表明本文方法能有效地缓解网络流量检测任务中对人工提取特征的依赖并提高检测准确率。此外，在 MAWILab 数据集上进行的对比实验分类结果表明，本文所提出方法优于其他基于深度学习的网络流量检测方法(如全连接网络、卷积神经网络和长短记忆网络)，本文方法在 784 维和 1 600 维的数据上的准确率分别为 86.18% 和 89.85%。神经网络仍存在可解释性差的问题，未来工作可以提高神经网络在特征提取工作中的可解释性，挖掘更多有助于提高网络流量检测的特征类型。

参 考 文 献

[1] Mukherjee B, Heberlein LT, Levitt KN, et al. Network intrusion detection [J]. IEEE Network, 1994, 8(3): 26-41.

[2] Hindy H, Brosset D, Bayne E, et al. A taxonomy and survey of intrusion detection system design techniques, network threats and datasets [Z/OL]. arXiv: 1806.03517v1, 2018.

[3] Buczak AL, Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection [J]. IEEE Communications Surveys and Tutorials, 2016, 18(2): 1153-1176.

[4] Tidjon LN, Frappier M, Mammar A, et al. Intrusion detection systems: a cross-domain overview [J]. IEEE Communications Surveys and Tutorials, 2019, 21(4): 3639-3681.

[5] Chalapathy R, Chawla S. Deep learning for anomaly detection: a survey [D]. Sydney: University of Sydney, 2019.

[6] Ahmed M, Mahmood AN, Hu J, et al. A survey of network anomaly detection techniques [J]. Journal of Network and Computer Applications, 2016: 19-31.

[7] Nisioti A, Mylonas A, Yoo PD, et al. From intrusion detection to attacker attribution: a comprehensive survey of unsupervised methods [J]. IEEE Communications Surveys and Tutorials, 2018, 20(4): 3369-3388.

[8] Rigaki M. Adversarial deep learning against intrusion detection classifiers [D]. Luleå: Luleå University of Technology, 2017.

[9] Lin Z, Shi Y, Xue Z, et al. IDSGAN: generative adversarial networks for attack generation against intrusion detection [Z/OL]. arXiv: 1809.02077, 2018.

[10] Dainotti A, Pescape A, Claffy KC, et al. Issues and future directions in traffic classification [J]. IEEE Network, 2012, 26(1): 35-40.

[11] Anderson B, McGrew D. Machine learning for encrypted malware traffic classification: accounting for noisy labels and non-stationarity [C] // Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017: 1723-1732.

[12] Aygun RC, Yavuz AG. Network anomaly detection with stochastically improved autoencoder based models [C] // 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing, 2017: 193-198.

[13] An J, Cho S. Variational autoencoder based anomaly detection using reconstruction probability [Z/OL]. Special Lecture on IE, 2015: 1-18.

[14] Intrator Y, Katz G, Shabtai A. Mdgan: boosting anomaly detection using multi-discriminator generative adversarial networks [Z/OL]. arXiv: 1810.05221, 2018.

[15] Zhou C, Paffenroth RC. Anomaly detection with robust deep autoencoders [C] // The 23rd ACM

SIGKDD International Conference, 2017: 665-674.

[16] Zhao S, Li W, Zia T, et al. A dimension reduction model and classifier for anomaly-based intrusion detection in internet of things [J]. IEEE Computer Society, 2017, 1: 836-843.

[17] Lin WC, Ke SW, Tsai CF. CANN: an intrusion detection system based on combining cluster centers and nearest neighbors [J]. Knowledge-Based Systems, 2015, 78: 13-21.

[18] Wang W, Zhu M, Wang JL, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks [C] // 2017 IEEE International Conference on Intelligence and Security Informatics, 2017.

[19] Wang W, Zhu M, Zeng XW, et al. Malware traffic classification using convolutional neural network for representation learning [C] // 2017 IEEE International Conference on Information Networking, 2017: 712-717.

[20] Mirza AH, Cosan S. Computer network intrusion detection using sequential LSTM neural networks autoencoders [C] // The 26th IEEE Signal Processing and Communications Applications Conference, 2018: 1-4.

[21] Shiravi A, Shiravi H, Tavallaee M, et al. Toward developing a systematic approach to generate benchmark datasets for intrusion detection [J]. Computers & Security, 2012, 31(3): 357-374.

[22] Lecun Y, Kavukcuoglu K, Farabet C, et al. Convolutional networks and applications in vision [C] // Proceedings of 2010 IEEE International Symposium on Circuits and Systems, 2010: 253-256.

[23] Yu Y, Long J, Cai Z, et al. Network intrusion detection through stacking dilated convolutional autoencoders [J]. Security and Communication Networks, 2017, 13(5): 1-10.

[24] Radford BJ, Apolonio LM, Trias AJ, et al. Network traffic anomaly detection using recurrent neural networks [Z/OL]. arXiv: 1803.10769v1, 2018.

[25] Fontugne R, Borgnat P, Abry P, et al. MAWILab: combining diverse anomaly detectors for automated anomaly labeling and performance benchmarking [C] // Proceedings of the 2010 ACM Conference on Emerging Network Experiments and Technology, 2010.

[26] Kwon D, Natarajan K, Suh SC, et al. An empirical study on network anomaly detection using convolutional neural networks [C] // The 38th IEEE International Conference on Distributed Computing Systems, 2018: 1595-1598.

[27] Siffer A, Fouque PA, Termier A, et al. Anomaly detection in streams with extreme value theory [C] // The 23rd ACM SIGKDD International Conference, 2017: 1067-1075.