计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200900155

基于 Attention-CNN 的加密流量应用类型识别

陈明豪　祝跃飞　芦　斌　翟　懿　李　玎
信息工程大学网络空间安全学院　郑州 450001
数学工程与先进计算国家重点实验室　郑州 450001
(1069304038@qq.com)

摘　要　随着流量加密技术的不断发展,加密流量已逐渐取代非加密流量成为当前网络环境的主流,其在保护用户隐私的同时,也常被各种恶意软件用来规避传统的基于端口或载荷关键字的入侵检测系统的防御,给网络安全带来了严重威胁。针对常规识别方法的局限性,研究人员尝试利用人工智能的方法来识别加密流量的应用类型,但现有研究对加密流量的特征信息的利用不够充分,导致相关方法在实际复杂的网络环境中表现不佳。为此,提出了一种基于 Attention-CNN 的加密流量识别方法,在加密流量数据初步特征提取的基础上,使用 BiLSTM+Attention 和 1D-CNN 模型对加密流量的时序和空间特征进行特征压缩和进一步提取,并利用基于全连接神经网络得到的混合特征进行最终的识别。文中采用通用的 ISCXVPN2016 开源数据集进行实验验证,结果表明所提方法的整体识别准确率达到了 0.987,且相比现有研究,对不同类别流量识别结果的 F1 评价指标有显著提升。

关键词:网络安全;加密流量;BiLSTM;Attention 机制;1D-CNN
中图法分类号　TP309

Classification of Application Type of Encrypted Traffic Based on Attention-CNN

CHEN Ming-hao,ZHU Yue-fei,LU Bin,ZHAI Yi and LI Ding
School of Cyberspace Security,Information Engineering University,Zhengzhou 450001,China
State Key Laboratory of Mathematical Engineering and Advanced Computing,Zhengzhou 450001,China

Abstract　With the development of traffic encryption technology,encrypted traffic has gradually replaced non-encrypted traffic and become the most important part of the current network environment. While protecting users' privacy,encrypted traffic is also used by malicious software to avoid the defense of traditional intrusion detection system based on the port or payload keywords of traffic,which brings serious threat to network security. In view of the limitations of conventional classification methods,researchers try to use artificial intelligence method to classify the application type of encrypted traffic,but the existing researches usually do not make full use of the characteristics of encrypted traffic,resulting in poor performance in the actual complex network environment. To solve the problems mentioned above,this paper proposes an encrypted traffic classification method based on Attention-CNN model. After the preliminary feature extraction of encrypted traffic,we use both BiLSTM+Attention and 1D-CNN model to compress and further extract the temporal and spatial features of encrypted traffic respectively. Finally,one fully connected neural network is used for the final classification based on the obtained mixed features. Experiments are carried out on the ISCXVPN2016 dataset which is the widely used open source dataset in encrypted traffic classification area. Experimental results show that the overall classification precision of the Attetnion-CNN could reach 98.7% and the F1 score is significantly improved compared with several existing studies.

Keywords　Cyber security,Encrypted traffic,BiLSTM,Attention mechanism,1D-CNN

1　引言

随着流量加密技术的普及和用户安全意识的提高,加密流量已经成为互联网环境中的重要成分。根据 Google 透明度报告"Chrome 中的 HTTPS 加密情况"[1] 显示,截至 2020 年 4 月,谷歌 Chrome 浏览器加载的网页中超过 95%启用了加密服务。2018 年思科公司对超过 40 万种恶意软件进行监控和分析后发现,其中已经有超过 70%的恶意软件在通信过

到稿日期:2020-09-21　修回日期:2020-11-02　本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:国家重点研发计划前沿科技创新专项基金(2019QY1300)
This work was supported by the Cutting-edge Science and Technology Innovation Project of the Key R&D Program of China(2019QY1300).
通信作者:祝跃飞(yfzhu17@sina.com)

程中使用了流量加密技术[2]。RadWare[3]报告显示,35%受调查的组织或者个人正面临采用 TLS/SSL 加密的攻击。流量加密技术是一把“双刃剑”,它在保护用户隐私的同时也为诸多恶意软件隐藏其攻击行为提供了便利。作为网络防御中的一项基础工程,识别加密流量所属的应用类型具有很高的研究价值。

加密流量采用的端口混淆和端口跳变技术导致传统的基于端口的流量识别方法的准确率大幅下降[3];而加密流量对传输载荷进行加密的特性也导致了基于载荷关键字的流量识别方法“水土不服”[4]。NSS 实验室对全球主要安全设备产商的下一代防火墙产品进行调研时发现,SSL 解密会导致平均 81%的性能损失。由此可见,对于加密流量,先解密再识别的模式已无法满足流量实时识别的需求,且对于加密流量无差别解密还存在容易侵犯用户个人隐私的问题。如何在不解密流量的条件下,对加密流量的应用类别进行高效准确的识别,成为了近年来学术界的一个研究热点。

随着深度学习在图像识别、文本翻译和自动驾驶等领域取得的巨大成功,越来越多的研究者尝试将深度学习的模型和方法应用到加密流量识别领域。从识别数据粒度的划分[5]来看,相关工作主要有 3 种:1)基于数据包层面的识别,对网络流量中每一个数据包的应用类别进行单独的识别;2)基于数据流层面的识别,对常见的“客户端—服务器”或者“浏览器—服务器”模型通信过程中的上行或者下行方向的流量进行识别;3)基于流量会话层面的识别,将客户端或者浏览器和服务器通信的上下行完整流量视为一个整体,以会话为基础识别单位进行识别。从使用数据的标注情况来看,主要包括有监督学习的识别和半监督学习的识别两个方向。有监督学习使用的所有数据都有对应的标签信息,而半监督学习一般采用大量容易获取的无标签流量数据和少部分人为标注的流量数据共同构成数据集。从采用的模型来看,主要有循环神经网络[6]、卷积神经网络[7]和自动编码机[8]等。

在数据流或者流量会话识别层面,文献[9]记录流量会话前 150 个数据包的长度和到达时间间隔信息,搭配这 150 个数据包文本中 0x00 到 0xff 的分布情况,构造整个会话的流量指纹特征,进行 TLS 流量恶意和非恶意的判别。文献[10]以清华大学校园网的正常流量为白样本,以沙箱运行恶意流量样本产生的流量为黑样本,收集一个会话的前 50 个数据包的长度、传输时间间隔和传输方向信息,并用 LSTM 进行训练,实验结果表明,在实验环境下基于 LSTM 算法的加密流量识别模型准确率和误报率均优于基于决策树、支持向量机和随机森林等传统机器学习算法的识别模型。文献[11]选择流量会话的前 1521 字节作为特征向量,并将其转化为 39×39 的灰度图,通过二维卷积神经网络以图像识别的方式进行加密流量的识别,达到了 92.92%的平均准确率。文献[12]则选择了流量会话载荷部分的前 1 000 个字节作为特征向量,先采用 Skip-gram 对特征向量中的每一个字节信息进行维度为 300 的词嵌入展开,然后对新的特征向量使用一维卷积神经网络进行最终识别,识别的准确率为 91.03%。在数据包识别层面,文献[13]引入文本识别的经验,将数据包头部的文本内容视为定长的句子,其中的字节信息视为单词,将数据包识别的工作等价于对等长句子的识别,先将数据包头部每个字节的信息进行词嵌入操作以提升特征维度空间,然后使用长短时记忆神经网络模型来学习字节之间的时序关系,在 Mirai-RGU 数据集上对流量恶意性的判断准确率达到了 97.2%。针对流量数据集中存在的样本分布非均衡的问题,文献[14]使用了半监督学习的理论,先利用大量无标签的流量会话数据,以数据采样的时序特征预测整体会话的统计特征的方式对模型进行预训练,再利用少量的有标签的数据进行针对性的再训练,最终在 QUIC 数据集上取得了 84.53%的准确率。文献[15]则采用 AC-GAN 模型对数据集中的弱势样本进行填充,然后对填充先后的数据集均使用支持向量机、决策树和随机森林等算法进行流量识别,在填充后的数据集上学习的模型的整体识别准确率更高。

从上文列举的相关工作可以发现,学术界目前对加密流量识别和分类的研究主要集中于流量的时序特征[16]或者空间特征[17]。传统意义上的时序特征一般指流量会话中每个数据包所携带的可以构成序列的特征信息,如数据包的长度、传输时间间隔、方向等,其优点是特征规模小、模型训练速度快,且不受流量加密的影响,相关方法从非加密流量识别到加密流量识别的迁移性较好;但也会因为特征维度低导致特征代表性不足、易产生数据标签冲突等问题。而流量的空间特征主要指数据包所携带的文本信息,对常见的流量加密协议(如 SSL/TLS,IPsec 和 SSH)的流量数据进行分析可以发现,其通信流程一般分成两个阶段:握手阶段和加密数据传输阶段。其中握手阶段主要是通信双方在验证身份和协商加密时使用加密套件、压缩函数、支持的版本等参数的过程,该过程一般采用明文进行传输,且不同应用类型流量的握手流量存在较大差异。同时,即使是加密数据传输阶段,不同应用的密文传输的载荷部分在模式上也存在一定的差异,这些都可以作为加密流量应用类型识别的依据。空间特征的优点在于特征维度高、识别准确率相对较高,缺点是模型训练速度较慢,同时因为放弃了会话流量数据的时序特征,所以对文本内容信息相近但是时序特征不同的流量类型(如聊天流量和邮件流量)容易产生误判。上述相关工作提到的各种方法虽然在各自的实验环境下都取得了不错的实验结果,但通常只围绕加密流量的一个特征维度进行详细的研究,导致模型鲁棒性不足,在面对复杂网络流量时识别效果可能会出现严重下滑。

针对上述情况,本文提出了一种 Attention-CNN 模型,同时利用流量数据的时序特征和空间特征作为识别的依据。该模型分别使用 BiLSTM+Attention 模型和 1D-CNN 模型对流量的时序特征和空间特征进行特征压缩和进一步提取,并将处理后的时序特征和空间特征拼接在一起得到会话的混合特征向量,采用全连接神经网络进行最终的识别任务。模型的训练过程和学习过程不需要人工参与,且由于使用的特征信息更为全面,因此特征提取过程更为科学,模型相比现有方法在加密流量识别效果上有明显的提升。

本文的主要贡献如下:

(1)提高了加密流量识别结果的准确性和鲁棒性,同时使用加密流量的空间特征和时序特征作为模型判断的依据。

(2)解决了传统时序特征面对复杂网络环境存在的代表

性不足和易发生特征冲突的问题，在数据处理过程中使用匿名化处理后的数据包头特征对传统时序特征进行补充。

(3)实现了 Attention-CNN 的加密流量应用类型识别模型，充分利用不同深度学习模型各自的长处对流量不同维度的特征进行压缩和进一步的提取，解决了特征维度差异性较大的问题。同时采用离线训练、在线识别的模式，实现了对加密流量应用类型进行实时识别的目标。

2 算法设计

2.1 模型框架

本文提出的 Attention-CNN 模型在流量会话层面进行加密流量应用类型的识别，使用的数据格式为原始流量数据，整个模型的工作流程如图 1 所示。图 1 中，FC Network 指全连接神经网络，1D-CNN 指一维卷积神经网络，BiLSTM＋Attention 指注意力机制指导的双向长短时记忆网络。对于原始流量数据，首先需要按 2.2 节中介绍的方式进行数据处理和特征提取，然后采用 BiLSTM＋Attention 模型和 1D-CNN 模型对初步提取的时序特征和空间特征做进一步的特征压缩和提取，取两个模块对应全连接神经网络部分倒数第二层的输出作为新的时序和空间特征，将两者拼接在一起得到代表该会话流量的混合特征向量，并用一个新的全连接神经网络基于混合特征进行最终的流量分类和识别。

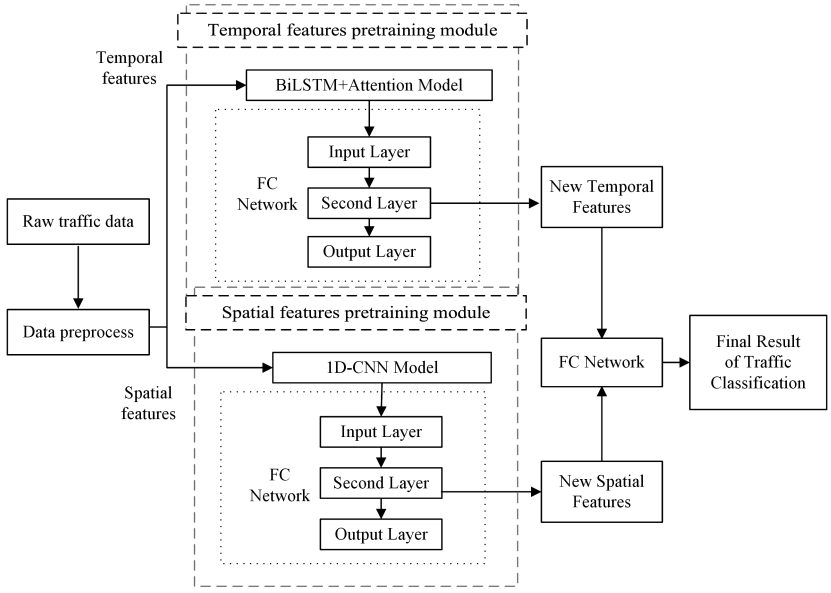

图 1 Attention-CNN 模型流程图

Fig. 1 Flowchart of Attention-CNN model

2.2 数据处理

整个数据处理流程如图 2 所示，首先使用 SplitCap 工具将 pcap 数据根据五元组信息(源 IP 地址、源端口、目的 IP 地址、目的端口、网络协议)以会话为单位进行划分。还需要对得到的会话流量进行过滤，主要是过滤掉以下 3 种流量。

(1)TCP 握手失败的会话。握手失败的数据包并没有携带任何有效的应用程序信息，对于实际的业务流量识别不能提供有效信息。该类型会话的过滤标准为传输层协议为 TCP 且整个会话握手信息不完整或者没有业务载荷。

(2)DNS 协议查询域名会话。由于 DNS 服务器的 IP 地址和端口一般与具体业务流量的目的 IP 地址和目的端口不同，因此即使一个流量业务需要使用 DNS 服务，其 DNS 查询的流量和具体业务的流量在进行会话划分时将会被划分到不同的会话中，单独的 DNS 流量会话对于流量类型的识别帮助较小。该类型会话流量过滤的标准为整个会话数据包均为 DNS 协议数据包。

(3)LLMNR 协议的会话。LLMNR 协议的应用情景与 DNS 协议类似，当 DNS 服务器不可用时，DNS 客户端计算机可以使用 LLMNR 协议来解析本地网段上的名称。LLMNR 会话最显著的特征是会话的第一个数据包必定是 UDP 协议，且目的 IP 是 224.0.0.252，目的端口是 5355，可以以此为规则进行过滤。

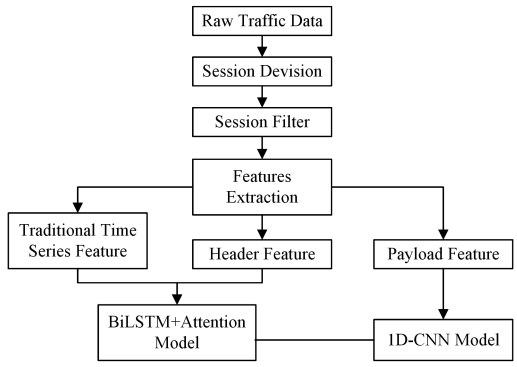

图 2 数据处理流程图

Fig. 2 Flowchart of data processing

上述 3 种会话在不同类型的流量中频繁出现，可以被视为共性的背景流量，但是其自身携带的有效信息较少，不利于模型对该类型流量具体业务特征的提取和学习，因此在数据处理流程中将其过滤掉。对于经过划分的会话和过滤后的会话流量，提取 3 种不同的特征作为用于后续模型的输入，3 种特征的名称和对应的维度情况如表 1 所列，其中 n_1 和 n_2 均为

正整数超参数，具体数值由第 3 节中的实验来确定。

表 1 特征提取维度情况

Table 1 Dimensions of extracted features

Feature Name	Feature Dimensions of a traffic session
Traditional Time Series Feature	$(n_1, 2)$
Header Feature	$(n_1, 40)$
Payload Feature	$(1, n_2)$

(1)传统时序特征。一个流量会话的传统时序特征选择的是该会话前n_1个数据包的长度和传输的时间间隔。将会话第一个数据包的方向设为该会话的正方向，对于后续的数据包，如果其方向与正方向相同，则方向特征设置为+1，否则设置为−1。将数据包的方向特征值与该数据包的长度值相乘，则可以通过数据包长度值的正负来代表数据包的方向。最终每个会话的时序特征的特征维度为$(n_1, 2)$，如果这个会话的实际数据包个数少于n_1，则在特征向量的对应位置用 0 来填充。

(2)数据包头特征。一个流量会话的数据包头特征选择的是该会话前n_1个数据包头部所携带的文本特征。出于对结构统一性和特征稳定性的考虑，数据包头特征选择的是 TCP/IP 4 层结构中的网络层包头和传输层包头，同时为了避免模型将 IP 地址信息视为流量分类的关键特征，在特征提取过程中需要将传输层包头中的源 IP 地址和目的 IP 地址全部设置为 0.0.0.0 来实现匿名化。由于网络层 IP 包头长度一般为 20 字节，传输层 TCP 协议的包头一般为 20 字节，UDP 协议包头的长度一般为 8 字节，为了统一特征格式，对于传输层协议为 UDP 的数据包，将在该数据包的 UDP 包头结尾填充 12 个字节的 0x00 使其长度也为 20 字节。完成填充操作后，每个数据包从 IP 包头的第一个字节开始，提取长度为 40 个字节的信息，作为数据包头特征，对于每一个字节的信息，先将其从 16 进制转换成区间为[0,255]的 10 进制整数，然后除以 255 进行归一化。最终得到该会话的数据包头特征维度为$(n_1, 40)$，如果该会话的数据包个数不足n_1，则在特征向量的对应位置用 0 填充，整个数据包头特征结构如图 3 所示。

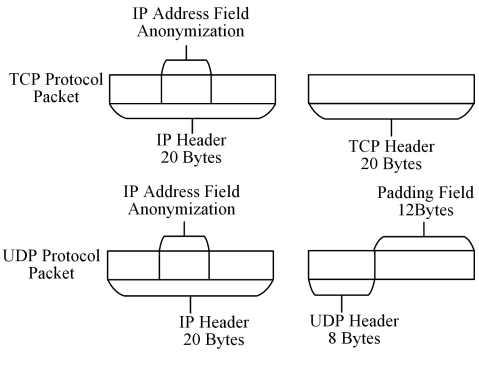

图 3 数据包头特征示意图

Fig. 3 Diagram of data packet header feature

(3)数据包载荷文本特征。一个会话的数据包载荷文本特征提取的是该会话传输层载荷部分前n_2个字节的信息。如果该会话所有数据包的传输层载荷总长度不足n_2个字节，则在记录该会话所有传输层载荷的基础上填充 0x00 直到长度为n_2字节；否则截取该会话前n_2字节的传输层载荷来构成特征向量。对于提取的数据包载荷文本特征向量，同样需要先将每个字节的信息从 16 进制转换成区间为[0,255]的 10 进制整数，然后除以 255 进行归一化，最终一个会话的传输层载荷文本特征维度为$(1, n_2)$。

对于从会话中提取的 3 种特征，将传统意义上的时序特征和数据包头特征拼接在一起生成维度为$(n_1, 42)$的时序特征，并将其作为 BiLSTM+Attention 模型的输入；将数据包载荷文本特征作为 1D-CNN 模型的输入。

2.3 基于 BiLSTM+Attention 的时序特征提取模型

循环神经网络(Recurrent Neural Network，RNN)是对具有较强序列相关性的数据进行特征提取和分类的一种行之有效的方法，其优点在于将历史输入视为背景信息，模型可以学习当前输入数据和背景信息之间存在的相关关系。但是，最初版本的 RNN 模型由于神经元内部结构的设计问题，在反向传播更新模型参数过程中容易出现梯度消失或者梯度爆炸的情况，导致随着输入层神经元个数的增多，后续的神经元实际上很难“记住”历史输入信息。LSTM 作为一种主流的 RNN 改进模型，通过在神经元内部添加 3 个阈值结构(遗忘门、输入门和输出门)的方式有效地提高了模型的长期记忆能力。考虑到会话流量中每一个时刻数据包的时序特征不仅受到历史数据包的影响，而且与未来的数据包息息相关，因此本实验采用了 LSTM 的变种模型——双向 LSTM(BiLSTM)对数据进行处理。同时，考虑到在会话流量中每一个数据包的重要性不同，为了突出这种差异性，以进一步提高模型的识别效果，本文在 BiLSTM 模型的基础上还采用了 Attention 机制对其最后一个时刻的隐藏层输出计算权重并进行加权求和。整个 BiLSTM+Attention[18] 模型的结构如图 4 所示。

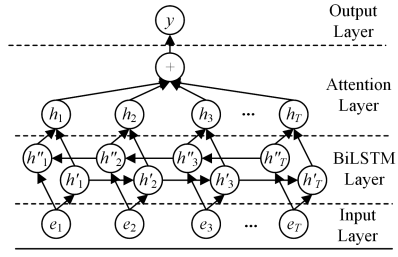

图 4 BiLSTM+Attention 结构示意图

Fig. 4 BiLSTM+Attention model structure

由图 4 可以看出，整个 BiLSTM+Attention 模型从功能上可以分为 4 层，为了更好地介绍在加密流量应用识别任务中模型的运行情况，取更为常见的文本分类任务进行类比。在本文中，一个流量会话可以类比为文本分类中的句子，会话中每一个数据包可以类比为句子中的单词。第一层是模型的输入层，图中的e_i表示每一个流量数据包的时序特征对应的输入神经元，根据 2.2 节中的介绍，本文使用的时序特征维度为$(n_1, 42)$，因此 i 的取值范围为$[1, n_1]$，同时e_i对应的输入维度为 42。第二层是 BiLSTM 层，将采用 BiLSTM 模型从会话流量的时序特征序列中提取更高级的特征，通过该层每个e_i的输入都会得到一个正向传递的隐藏层输出h_i'和逆向传递的隐藏层输出h_i''。第三层是 Attention 层，这一层首先将

BiLSTM 层对应的正反向隐藏层结果h'_i和h_i''相加。

$$h_i = h_i' \oplus h_i'' \tag{1}$$

其中，$\oplus$运算代表的是向量的元素对位相加，通过该运算可以得到 Attention 层的输入向量 $\boldsymbol{H}$：$[h_1, h_2, \cdots, h_T]$，而 Attention 层的权重矩阵由以下公式获得：

$$M = tanh(H) \tag{2}$$

$$\alpha = softmax(w^{\mathrm{T}} M) \tag{3}$$

式(3)中，w^{T}是一个需要模型经过训练后学习到的权重矩阵。通过式(4)可以得到 Attention 层的最终输出h^*：

$$h^* = \tanh(H\alpha^{\mathrm{T}}) \tag{4}$$

最后一层输出层紧跟的是一个输入维度为h^*、隐藏层维度为 256、输出维度为流量分类的种类个数、激活函数为 Softmax 的全连接神经网络。在使用流量分类任务对整个模型进行预训练以确定 BiLSTM＋Attention 模型的内部参数后，输入模型的每一条会话流量可以提取全连接网络倒数第二层的输出，用于代表该会话的新时序特征，维度为(1,256)。

2.4 基于 1D-CNN 的空间特征提取模型

卷积神经网络(Convolutional Neural Network，CNN)由于善于提取数据局部特征和下采样特性，相比其他深度学习模型更适合处理特征维度较高的数据，如图像、文本和本文中使用的加密流量空间特征。本文中经过数据处理环节提取的初步空间特征为会话流量传输层载荷信息，如果采用 2D-CNN 进行识别，则首先需要将连续的文本转换成格式为 $n \times m$ 的二维矩阵(n,m 均为正整数)，但这会使得对于区间为(0，$n-1$] 的正整数 i，转换后的二维矩阵第 i 行开头部分的信息和第 $i-1$ 行末尾的信息被模型视为几乎完全分离的两部分，但实际上它们很可能是一个数据包传输层载荷中连续的部分。因此，2D-CNN 模型采用的数据转换方式会导致数据原始信息的一定损失，从而对加密流量的分类和识别产生不利的影响，而如果采用 1D-CNN 模型，则不存在类似的问题。因此，本文对于空间特征的压缩和进一步提取采用 1D-CNN 模型来代替更为常见的 2D-CNN 模型。

本文设计的 1D-CNN 模型流程如图 5 所示，从结构上来看可以分成两部分：1)由卷积层(Convolution Layer)、归一化层(Normalization Layer)、抛弃层(Dropout Layer)和池化层(Pooling Layer)组成的重复两轮的循环结构体部分；2)Flatten 层及其连接的全连接神经网络部分。在循环结构体中，卷积层的主要作用是对输入的空间特征向量进行特征提取，将全局的特征信息保存在多个局部特征矩阵中；归一化层则是在不影响数据真实分布的前提下，将数据尽量向原点靠拢，能够大大提高模型的训练速率；抛弃层的主要功能是避免过拟合，其在 1D-CNN 模型反向传播更新模型参数的过程中，以一定的概率将参数从模型中暂时“抛弃”——放弃本轮反向传播对于该参数的更新，保证不同迭代轮次训练得到的模型的差异性，能有效降低过拟合现象；池化层的主要作用是进行数据降维和特征压缩，本文的 1D-CNN 模型采用的是最大值池化(选取某一局部的最大值作为该局部数据的代表)。在全连接网络中，先通过一个 Flatten 层将前面循环结构体的输出转换成适合全连接网络输入的一维格式，然后用全连接神经网络进行加密流量的分类和识别。与 BiLSTM＋Attention 模型对加密流量的时序特征进行处理相似，在进行预训练确定模型参数后，对于输入模型的每一条会话流量，1D-CNN 模型提取全连接网络倒数第二层的输出代表会话流量的新空间特征，特征维度同样为(1,256)。

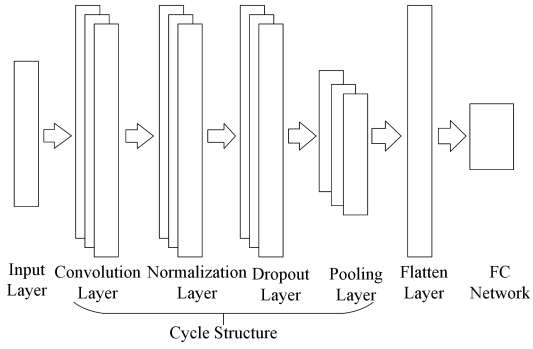

图 5 1D-CNN 模型

Fig. 5 1D-CNN model

3 实验

3.1 实验环境与数据

本文实验基于 Windows10 操作系统，以 Python3.7 为基础语言环境，使用 SplitCap 工具对 pcap 文件划分会话，数据读取、过滤和预处理操作涉及到了 python 下的 scapy 库和 dpkt 库，数据格式转换和矩阵运算使用 numpy 库，相关深度学习模型框架基于 pytorch 1.5.0 版本进行开发并采用 GPU 进行加速，GPU 加速硬件为 NVIDIAGeForce 1060 6GB。

实验采用的数据集为 ISCXVPN2016 数据集。该数据集由加拿大网络安全研究所整理和制作，是目前在加密流量识别领域中使用得最为广泛的原始流量数据集，其中包含了 Gmail，Skype 和 Facebook 等 17 种常见应用在正常环境和 VPN 环境下的流量，数据集的初始大小约为 28GB。在本实验中，将该数据集按照流量类型划分，不同应用程序的同一种应用类型流量被划分为同一个类别，如 aim_chat 流量、facebook_chat 流量和 icq_chat 流量均被划分到 Chat 类型下，从数据集中一共选择了 Chat，Email，Files，Stream，Torrent 和 Voip 这 6 种主要的应用类型及其对应的 VPN 流量，共计 12 种流量类别。

ISCXVPN2016 数据集按第 2.2 节中的方式进行数据处理后，可以得到 147248 条会话流量数据。为了保证实验结果的有效性和可靠性，对于数据处理后的数据按照 8 ∶1 ∶1 的比例随机选取划分训练集、验证集和测试集，采用十折验证的方式降低训练过程中的偶然因素对最终结果的影响。

3.2 评价指标

本实验采用准确率(Accuracy)作为模型的整体评价指标，对于每一种流量子类别，在计算精确率(Precision)和召回率(Recall)的基础上计算 F1 评价指标，并将其作为模型对于该流量子类别识别结果的综合评价指标。上述评价指标的计算式如下：

$$Accuracy = \frac{T}{T+F} \tag{5}$$

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

$$Recall=\frac{TP}{TP+FP} \tag{7}$$

$$F_1=2\times\frac{precision\times recall}{precision+recall} \tag{8}$$

式(5)中，T 表示全部数据中模型识别结果正确的数目，F 表示全部数据中模型识别错误的数目。精确度、召回率和 F1 评价指标都是针对某一种具体的目标流量类型而言的，假设我们的目标流量类型为 A，那么式(6)和式(7)中的 TP 表示类别为 A 的数据被模型正确识别的数目，式(6)中的 FN 表示类别为 A 的流量被模型识别错误的数目，式(7)中的 FP 表示非 A 类型的流量被模型误识别为 A 的数目。将式(6)和式(7)的结果代入式(8)即可得到模型对于流量 A 识别的 F1 评价指标结果。

3.3 实验结果和分析

本文的实验分成两个阶段：阶段一的工作主要是确认第 2.2 节数据处理中涉及超参数的具体数值；阶段二则是在阶段一工作的基础上，使用 ISCXVPN2016 数据集将本文设计的 Attention-CNN 模型和前人的相关工作进行对比。

3.3.1 超参数对比实验

本实验要确认的超参数一共有两个，分别是用于时序特征提取的超参数 n_1 和用于空间特征提取的超参数 n_2。n_1 的主要功能是决定会话中提取时序特征的数据包数目，以最常见的加密协议 TLS 协议为例，因为一个完整的 TLS 通信过程中仅握手环节就至少需要 Client Hello，Server Hello，Server Certification，Client Key Exchange，Change Cipher Spec 和 Encrypted Handshake Message 这 6 种数据包，同时对数据集中的数据进行分析发现，98% 的会话数据包个数少于 50，因此本实验中 n_1 的取值范围为[10，50]，取值间隔为 5。n_2 的主要功能是决定提取会话空间特征时提取的传输层载荷长度，本实验中其取值区间为[500，4 500]，取值间隔为 500。相关超参数的选择不但要考虑模型识别的准确率，还需要兼顾运算开销和训练成本等其他因素。表 2 列出了不同超参数组合下本文提出的 Attention-CNN 方法在 ISCXVPN2016 数据集上的整体识别准确率。

表 2 超参数实验结果对比

Table 2 Experimental results comparison of hyperparameters

Temporal Feature Hyperparameter n_1	Spatial Feature Hyperparameter n_2								
	500	1 000	1 500	2 000	2 500	3 000	3 500	4 000	4 500
10	0.966	0.946	0.936	0.978	0.978	0.958	0.981	0.973	0.964
15	0.961	0.952	0.936	0.980	0.980	0.966	0.985	0.986	0.961
20	0.964	0.953	0.927	0.986	0.983	0.971	**0.987**	**0.987**	0.968
25	0.957	0.951	0.933	0.974	0.979	0.962	0.982	0.986	0.964
30	0.971	0.953	0.935	0.978	0.976	0.963	0.979	0.985	0.959
35	0.962	0.942	0.938	0.976	0.976	0.949	0.981	0.979	0.957
40	0.951	0.947	0.933	0.972	0.975	0.958	0.979	0.972	0.957
45	0.960	0.947	0.930	0.980	0.973	0.961	0.984	0.983	0.958
50	0.962	0.954	0.934	0.969	0.976	0.959	0.973	0.979	0.960

对超参数 n_1 进行分析，由表 2 可以看出，对于不同的超参数 n_2，识别结果最佳的 n_1 一般出现在区间[20，30]中，说明提取大约 20～30 个数据包的时序特征就可以较好地代表整个会话流量的时序特征。同时，实验结果表明模型的识别准确率与 n_1 的数值之间并不是简单的线性正相关关系，盲目增加提取时序特征的数据包个数可能会造成有效特征被干扰，从而降低整个模型对加密流量类别识别的结果。而对于超参数 n_2，实验结果显示，从整体趋势来看，n_2 的取值越大，模型的整体识别准确率就越高，但也不是绝对的，例如 n_2 的取值从 500～1 000 反而使得模型整体识别准确率普遍下降。在表 2 列举的所有情况中，模型识别的最佳准确率为 0.987，其对应的 (n_1, n_2) 值有两组，分别为(20，3 500)和(20，4 000)，考虑到降低特征提取工作量和模型训练开销的实际应用需求，n_1 和 n_2 的最终取值分别为 20 和 3500。

3.3.2 相关工作对比实验

为了进一步验证本文模型的有效性，本文在 ISCXVPN2016 数据集上选取 3 种相关工作的模型 LSTM[10]，2D-CNN[19] 和 1D-CNN[20] 与本文提出的 Attention-CNN 模型进行对比实验，相关工作所涉及的模型的结构和内部参数完全依照原文献中的标准进行复现。

表 3 列出了 4 种模型在 ISCXVPN2016 数据集上的整体识别准确率。从表中可以发现，在本文的实验环境下，仅采用流量会话前 50 个数据包的传统时序特征作为输入的 LSTM 模型时识别准确率最低只有 64.1%。对于均采用会话流量空间特征作为输入的 1D-CNN 模型和 2D-CNN 模型，1D-CNN 模型的一个会话对应的输入空间特征维度是 784，而 2D-CNN 模型的一个会话对应的空间特征维度是 1 521，前者虽然使用的特征维度更低，但识别准确率却仍比后者提高了 8.3%，实验结果进一步证明了 1D-CNN 模型比 2D-CNN 模型更适合用于加密流量识别。本文提出的 Attention-CNN 模型同时利用了会话流量的空间特征和时序特征进行识别，准确率达到了 98.7%，相比 3 种相关工作中涉及到的模型 LSTM，1D-CNN 和 2D-CNN，其识别准确率分别提高了 34.6%，11.2% 和 19.5%。

表 3 所提模型与相关工作的识别准确率对比

Table 3 Comparison of classification accuracy between the proposed model and related works

Model Name	Classification Accuracy/%
LSTM	64.1
1D-CNN	87.5
2D-CNN	79.2
Attention-CNN	**98.7**

为了更好地说明 Attention-CNN 模型综合利用时序和空

间特征的优势，从表 2 中选择相关超参数最接近的实验结果与现有相关工作进行比较。当提取时序特征的数据包个数为 50 时，本文使用的 Attention-CNN 模型的最低识别准确率为 93.4%，相比 LSTM 模型提高了 29.3%。当提取的传输层载荷为 1 000 字节时（1 000 比 500 更接近 1D-CNN 模型的 784），Attention-CNN 模型的最低识别准确率为 94.2%，相比 1D-CNN 模型提高了 6.7%，而当提取传输层载荷为 1 500 时，Attention-CNN 模型的最低识别准确率为 92.7%，相比 2D-CNN 模型提高了 13.5%。上述 3 组对比结果更直观地说明了引入新的特征维度对于加密流量识别准确率提高的重要性，同时也为模型后续的改进提供了思路。

图 6 给出了 4 种模型对于不同类型流量识别的 F1 评价指标的对比情况，F1 评价指标是综合精确率和召回率的结果，可以更为全面地对比和分析 4 种模型对于不同类型加密流量的识别性能。从图中不难发现，LSTM，1D-CNN 和 2D-CNN 模型均有比较明显的识别结果不佳的流量：LSTM 模型对于 Voip 类型流量识别结果的 F1 评价指标为 0.48；1D-CNN 模型对于 Files 类型流量识别结果的 F1 评价指标为 0.786；2D-CNN 模型对于 Files 类型流量识别结果的 F1 评价指标仅为 0.281。而本文提出的 Attention-CNN 模型对于全部 12 种流量识别的最低 F1 得分，即对 Voip 类型流量的识别结果的 F1 评价指标为 0.936。从整体情况来看，LSTM 模型和 2D-CNN 模型的 F1 评价指标波动较大，1D-CNN 模型的波动稍小，但普遍低于本文提出的 Attention-CNN 模型。Attention-CNN 的 F1 评价指标波动极小且基本稳定在[0.95，0.99]区间内。上述情况充分说明了 Attention-CNN 模型对于不同类型流量的识别结果具有高度的稳定性和可靠性。

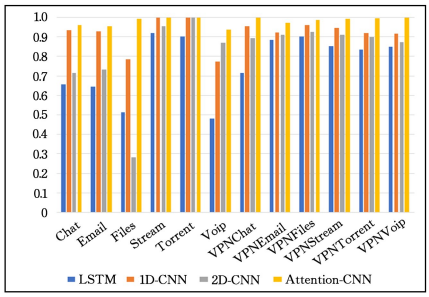

图 6 4 种模型 F1 评价指标对比图

Fig. 6 F1-score comparison of four models

图 7 给出了 Attention-CNN 模型在 ISCXVPN2016 数据集上识别结果的混淆矩阵，可以发现模型对不同类型的流量均取得了较高的识别水准，模型对实验数据中所有类型流量的识别准确率都超过了 0.9，且绝大部分类型流量的识别准确率约为 0.99。但从图 7 可以发现，现有版本的 Attention-CNN 模型在识别结果上还有两个不足有待改进：1）Chat 和 Email 两类流量存在一定程度上的相互误判，模型将 3.7%的 Chat 流量误判为 Email 流量，同时将 5.2%的 Email 流量误判为 Chat 流量；2）Voip 类型流量识别结果有待提高，模型将 7.6%的 Voip 流量误判为 Files 流量，这直接导致 Voip 类型流量识别的 F1 评价指标最低。在后续工作中可以通过引入新的流量特征维度、深入分析识别不佳的流量类型、有针对性地优化模型结构和调整参数等方式来继续改进模型。

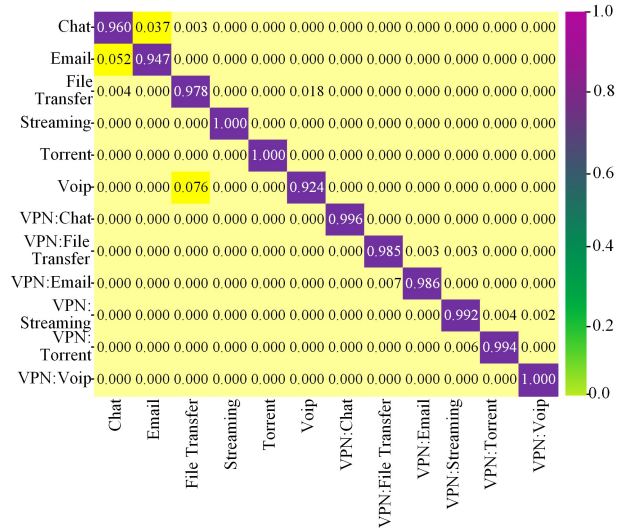

图 7 Attention-CNN 模型混淆矩阵

Fig. 7 Confusion matrix of Attention-CNN model

结束语 本文针对如何在不解密条件下对加密流量进行识别这一问题，提出了基于加密流量时空特征的 Attention-CNN 模型，用 BiLSTM＋Attention 模型来提取会话流量的时序特征，用 1D-CNN 模型提取会话流量的空间特征，并将提取出来的两种特征拼接起来作为最终识别的输入。在 ISCXVPN2016 数据集上进行实验，所提模型的整体识别准确率高达 0.987，同时在具体类别流量识别的 F1 评价指标上的表现均优于其他 3 种现有的相关工作，但其也存在两个小的不足需要继续改进。实验结果充分说明了本文模型的可靠性和有效性。

在未来的工作中，以下 3 个方向值得重点研究：

（1）为了满足模型输入维度一致的先决条件，目前学术界在流量识别领域利用流量空间特征进行识别的所有方法，对于不定长的流量文本信息，各自均在数据提取过程中首先进行截取或者填充操作，得到定长的流量文本，流量中被截取部分被直接丢弃，从而导致信息损失，可以结合 transformer 相关的研究，将不定长的会话流量通过一个编码器结构转化成定长的数据格式，然后进行后续的处理，将特征提取导致的信息损失降至最低。

（2）引入专业知识模块，从加密流量握手环节的明文信息中提取证书生命周期、证书是否自签名和加密套件选取详情等专业知识特征，结合现有工作搭建更加专业化的加密流量识别模型。

（3）引入深度学习模型可解释性方面的研究，从特征重要性排序、Attention 层参数代表的权重分析以及卷积神经网络特征图谱对应空间模式分析等角度入手，对模型的识别结果展开相关分析，进一步调整和优化模型结构。

参 考 文 献

[1] Google. Google Transparencyreport [R/OL]. (2020-07)[2020-07-01]. https://transparencyreport. google. com/https/overview.

[2] Cisco. Cisco Encrypted Traffic Analytics White Paper[R/OL]. (2019-07)[2019-07-20]. https://www.cisco.com/c/en/us/solutions/enterprisenetworks/enterprise-network-security/eta.html.

[3] Radware (2018). Global application and network security report [EB/OL]. https://www.datacomcz/userfifiles/radware_ert_report_2017_2018_fifinal.pdf.

[4] MADHUKAR A, WILLIAMSON C. A Longitudinal Study of P2P Traffic Classification[C]// modeling, analysis, and simulation on computer and telecommunication systems. 2006: 179-188.

[5] REZAEI S, LIU X. Deep Learning for Encrypted Traffic Classification: An Overview[J]. IEEE Communications Magazine, 2019, 57(5): 76-81.

[6] 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(99): 18042-18050.

[7] CHEN Z, HE K, LI J, et al. Seq2Img: A sequence-to-image based approach towards IP traffic classification using convolutional neural networks[C]// International Conference on Big Data. 2017: 1271-1276.

[8] HOCHST J, BAUMGARTNER L, HOLLICK M, et al. Unsupervised Traffic Flow Classification Using a Neural Autoencoder[C]// Local Computer Networks. 2017: 523-526.

[9] HU B, ZHOU Z H, LIAO L H, et al. TLS malicious traffic detection based on combined features of packet payload and stream fingerprints[J]. Computer Engineering, 2020, 46(520): 163-169.

[10] ZOU Y, ZHANG J, JIANG B. Detection of malicious encrypted traffic based on LSTM recurrent neural network[J]. Computer Applications and Software, 2020, 37(2): 308-312.

[11] GUO L, WU Q, LIU S, et al. Deep learning-based real-time VPN encrypted traffic identification methods[J]. Journal of Real-Time Image Processing, 2020, 17(1): 103-114.

[12] CHENG H, XIE J X, CHEN L H. CNN-based Encrypted C&C Communication Traffic Identification Method[J]. Computer Engineering, 2019, 45(8): 31-34, 41.

[13] HWANG R H, PENG M C, NGUYEN V L, et al. An LSTM-Based Deep Learning Approach for Classifying Malicious Traffic at the Packet Level[J]. Applied Sciences, 2019, 9(16): 3414.

[14] REZAEI S, LIU X. How to Achieve High Classification Accuracy with Just a Few Labels: A Semi-supervised Approach Using Sampled Packets[J]. arXiv: 1812.09761, 2020.

[15] VU L, BUI C T, NGUYEN Q U, et al. A Deep Learning Based Method for Handling Imbalanced Problem in Network Traffic Classification[C]// International Symposium on Information and Communication Technology. 2017: 333-339.

[16] LASHKARI A H, DRAPER-GIL G, MAMUN M S I, et al. Characterization of Encrypted and VPN Traffic Using Time-Related Features[C]// Proceedings of the 2nd International Conference on Information Systems Security and Privacy (ICISSP 2016). 2016: 407-414.

[17] LOTFOLLAHI M, SIAVOSHANI M J, ZADE R S, et al. Deep Packet: A Novel Approach For Encrypted Traffic Classification Using Deep Learning[J]. Soft Computing, 2020, 24(3): 1999-2012.

[18] ZHOU P, SHI W, TIAN J, et al. Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2016.

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

[20] WANG W, ZHU M, WANG J, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]// 2017 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, 2017.

CHEN Ming-hao, born in 1996, master. His main research interests include cyber security and encrypted traffic classification.

ZHU Yue-fei, born in 1962, professor, Ph.D, supervisor. His main research interests include intrusion detection, cryptography and information security.