基于不定长卷积神经网络的恶意流量分类算法

杨 璇[1], 邬江兴[2], 赵 博[3]

[1] 东南大学网络空间安全学院 南京 中国 211189
[2] 国家数字交换系统工程技术研究中心 郑州 中国 450002
[3] 中国人民解放军战略支援部队信息工程大学 郑州 中国 450001

摘要 在当今信息爆炸、网络快速发展的时代, 网络攻击与网络威胁日益增多, 恶意流量识别在网络安全中发挥着非常重要的作用。深度学习在图像处理、自然语言处理上已经展现出优越的性能, 因此有诸多研究将深度学习应用于流量分类中。将深度学习应用于流量识别时, 部分研究对原始流量数据进行截断或者补零操作, 截断操作容易造成流量信息的部分丢失, 补零操作容易引入对模型训练无用的信息。针对这一问题, 本文提出了一种用于恶意流量分类的不定长输入卷积神经网络(Indefinite Length Convolutional Neural Network, ILCNN), 该网络模型基于不定长输入, 在输入时使用未截断未补零的原始流量数据, 利用池化操作将不定长特征向量转化为定长的特征向量, 最终达到对恶意流量分类的目的。基于 CICIDS-2017 数据集的实验结果表明, ILCNN 模型在 F1-Score 上的分类准确率能够达到 0.999208。相较于现有的恶意流量分类工作, 本文所提出的不定长输入卷积神经网络 ILCNN 在 F1-Score 和准确率上均有所提升。

关键词 恶意流量; 流量分类; 卷积神经网络; 不定长输入
中图法分类号 TP393.0; TP183 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2022.07.07

Malicious Traffic Classification Based on Indefinite Length Convolutional Neural Network

YANG Xuan[1], WU Jiangxing[2], ZHAO Bo[3]

1School of Cyber Science and Engineering, Southeast University, Nanjing 212289, China
2China National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China
3Information Engineering University, Zhengzhou 450001, China

Abstract In today's era of information explosion and rapid network development, network attacks and network threats are increasing, and malicious traffic identification plays a very important role in network security. And deep learning has shown superior performance in image processing and natural language processing, so there are many researches to apply deep learning to traffic classification. When applying convolutional neural networks to traffic classification, some studies truncate or zero-complement the original traffic data, which may cause partial loss of traffic information and zero-complement operation may introduce information that is not useful for model training, thus affecting the detection accuracy of the model. In this paper, we propose an Indefinite Length Convolutional Neural Network (ILCNN) for malicious traffic classification, which is based on indefinite length input, and uses the raw traffic data without truncation and zero filling in the input, and uses the pooling operation to transform the indefinite length. This network model is based on indeterminate length input, using untruncated and un-zeroed raw traffic data in the input, and using pooling operation to transform indeterminate length feature vectors into fixed length feature vectors for the purpose of classifying malicious traffic. Because ILCNN uses the original traffic data and retains all the information of the traffic data, it can better perform feature extraction in the training phase of the model, avoiding the impact of losing some traffic information and introducing useless information, and eliminating the need for manual feature extraction and the tedious process of feature extraction of the traffic data; multiple convolutional kernels of different sizes are used in the model, which can extract the traffic classification. The model uses multiple convolution kernels of different sizes to extract features from different fields of view of the traffic data, which is convenient for subsequent classification of malicious traffic. The experimental results based on the CICIDS-2017 dataset show that the classification accuracy of the ILCNN model on F1-Score can reach 0.999208. Compared with the existing work on malicious traffic classification, the proposed ILCNN with indefinite long input convolutional neural network improves on both F1-Score and accuracy.

Key words malicious traffic; traffic classification; convolutional neural network; variable length input

通讯作者: 杨璇, 硕士生, Email: 220194739@seu.edu.cn
收稿日期: 2021-04-30; 修改日期: 2021-08-01; 定稿日期: 2022-05-11

1 引言

在当今信息爆炸、网络快速发展的时代, 互联网用户日益增多, 用户数据成为虚拟资产广泛存在于互联网中, 随之而来的网络攻击与网络威胁日益增多, 且随着时间与技术的发展, 网络攻击与网络威胁呈现出复杂化与多形态的趋势, 网络攻击与网络威胁已经严重影响到了网络空间的稳定运行, 并且造成了不小的经济损失。承载着网络攻击与网络威胁的流量是非正常的恶意流量, 网络流量异常检测[1]作为一种有效的防护手段, 是网络安全防御中非常重要的一部分, 能够对流量进行检测, 识别出其中各类不正常的恶意流量, 实现对网络态势的感知。

流量分类技术主要可以分为基于端口的流量分类方法、基于深度包检测(Deep Packet Inspection, DPI)的流量分类方法、基于统计特征的流量分类方法以及基于深度学习的流量分类方法。基于端口的方法[2]是基于 TCP 或 UDP 协议中的端口号进行识别, 在网络中的流量都长期使用固定端口号的情况下, 才能发挥出较高的准确率, 但是, 随着恶意软件使用动态端口、随机端口[3]、端口伪装[4]等技术的发展, 基于端口的检测方法已经不足以适用; 基于 DPI 的方法[5]通过分析数据包的包头以及负载内容, 并使用模式匹配、签名技术等来对流量进行分类, 虽然该类方法较之基于端口的方法准确率更高, 但是在面对数据特征不固定的流量时, 此类方法就没有用武之地。

基于统计特征的方法将机器学习应用到流量分类当中, 通过提取数据流的总字节数、数据流的持续时间、数据包间隔时间等统计特征, 应用随机森林、支持向量机等机器学习算法对流量进行识别和分类。相比于基于端口以及基于 DPI 的方法, 基于统计特征的方法可以应用于加密流量, 但是该类方法的性能不仅依赖于机器学习算法的模型以及模型参数, 还高度依赖于对流量数据的特征提取, 不仅需要先验知识, 还耗时耗力; 基于深度学习的流量识别方法将深度神经网络应用于流量分类, 深度神经网络能够在训练过程中自动提取流量数据中的可区分特征, 最终对流量进行准确识别, 然而在使用深度学习进行流量分类时, 会对原始流量数据包进行截断或者补零操作, 确保输入到神经网络进行训练的数据分组保持固定长度, 但是截断以及补零操作会影响流量分类的精确度。

上述恶意流量分析方法面临着几个方面的问题, 一方面是需要耗时耗力对数据进行特征提取, 在数据的各类特征均被很好提取的情况下, 依旧会丢失部分信息; 另一方面是对原始流量的补零以及截断操作会造成部分信息的丢失或者无用信息的引入。针对上述问题, 本文提出了不定长输入的卷积神经网络模型, 使用原始数据分组进行训练, 通过实验结果证明, 相比于使用相同数据集的其他工作, 本文使用的模型具有更优的性能。具体来说, 本文贡献如下:

(1) 本文提出了不定长输入的卷积神经网络对恶意流量进行分类, 该方法能够保留全部的流量信息, 提取流量数据包中的空间特征, 完成对恶意攻击流量的精确分类;

(2) 本文在对数据进行预处理过程中无需进行数据的截断或者补零操作, 避免了丢失部分流量信息、引入无用信息所带来的影响;

(3) 文本使用的数据为原始数据包, 无需进行人工特征提取, 省略了对网络流量进行特征提取的繁琐过程;

(4) 本文使用 CICIDS-2017[6]数据集对模型进行训练以及测试, 并通过精确率、召回率、F1-score 等性能指标对其进行性能评估。

本文结构如下, 第二部分介绍了流量分类的相关工作; 第三部分详细介绍了本文使用的恶意流量分类模型; 第四部分详细展示了在 CICIDS-2017 数据集上进行测试的实验结果并进行了分析; 最后对全文进行了总结。

2 相关工作

在流量识别问题的研究上, 从基于端口的流量分类方法到基于深度学习的流量分类方法, 学者们都做出了非常多的尝试。在本节中, 将主要回顾基于深度学习流量识别方法的主要研究以及不定长卷积神经网络的研究。

基于深度学习的方法能够在训练过程中自动提取流量数据中的可区分特征, 无需进行人工特征选取, 只需对流量进行截断或者补零等预处理即可。部分研究会使用单个深度学习模型进行流量分类, 单个的深度学习模型复杂度相对较低。Go J H 等人[7]使用 ResNeXt[8]网络进行恶意软件分类, 将恶意软件的二进制执行文件转化成灰度图输入到 ResNeXt 网络进行训练, 取得了 0.9832 的准确率, 并且与 ResNet 网络[9]和 Inception V4 网络[10]进行比较, 在计算量相同的情况下, 能达到更高的精确度; Pascanu 等人[11]提出了基于动态分析的两层架构的恶意软件检测系统, 第一层 RNN 用于学习特征, 第二层逻辑

回归分类器使用学习的特征进行分类, 然而误报率较高, 达到了 10%; Wang 等人[12]采用基于 1D-CNN 的端到端的加密流量分类方法, 将流量数据固定为 784byte 的等长数据, 输入 CNN 进行分类, 实验证明采用 1D-CNN 效果好于 2D-CNN; Radford B J 等人[13]使用 CICIDS-2017 数据集进行恶意流量的分类研究, 该文目标是使用无监督学习进行未知恶意攻击的识别, 在文中使用 LSTM 模型进行建模, 在端口号已知的情况下, 最终准确率能达到 87%。

另外还有部分研究同时训练多个模型进行流量分类, 多个模型能够同时学习数据的空间特征以及时序特征, 但是模型复杂度会更高一些。吴迪等人[14]通过将 CNN 与 LSTM 结合, 自动学习数据流中的时间与空间特征, 进行恶意流量的分类, 实验结果证明, 该模型在 F1-Score 上的分类准确率能够达到 0.9976, 但是该类结合模型复杂, 运行时间较长, 在网络流量急剧增大的情况下, 不能够达到实时检测的效率; Zhang 等人[15]将改进的 LetNet-5 卷积神经网络与 LSTM 神经网络组合成分层网络, 其中卷积神经网络用来抓取流量的空间特征, LSTM 神经网络用来抓取流量的时间特征, 基于原始数据流进行恶意流量分类任务, 最终 F1-Score 能达到 0.999161; Dong Y 等人[16]在文中设计了入侵检测系统, 包括底层的数据抓取层、数据存储层、数据处理计算层以及应用层, 在数据处理计算层的流量识别模型中, 首先通过 SAE 对数据进行降维处理, 然后使用 ALEXNET 网络[17]进行分类, 最终获得了 0.9432 的准确率, 但是该模型在适应性方面不强。

虽然深度学习能够自动学习数据特征, 但是部分研究也会基于特征数据集进行流量分类, 特征数据集的研究方法需要首先提取数据的特征。Zeng Y 等人[18]分别使用了 1D-CNN、LSTM、SAE 3 种深度学习算法实现加密流量识别以及入侵检测, 并在两类公开数据集上进行了测试, ISCX VPN-nonVPN 数据集[19]以及 ISCX 2012 IDS 数据集[20], 前者用来测试算法模型对于加密流量分类的有效性, 后者用来测试算法对于入侵检测的有效性, 其中最好的模型在加密流量识别上能达到 99.85%的准确率, 在入侵检测上能达到 99.41%的准确率; Prasse 等人[21]研究开发了基于 LSTM 的恶意软件检测模型, 该模型仅仅使用 HTTPS 流量的握手阶段信息, 能够识别网络流中的大部分恶意软件, 包括以前未见过的恶意软件; Nayyar S 等人[22]提出了基于 LSTM 的检测模型对 DDoS 攻击进行检测, 在 CICIIDS-2017 数据集上能达到 96.25%的精确率以及 7 ms 预测时; Torroledo 等人[23]专门进行特征工程, 分析总结了 4 类 40 个数据特征, 用于识别恶意软件和钓鱼软件签名证书, 实验中采用 LSTM 以及 5 折交叉验证进行检测, 结果表明具有较高的精度。

另外在不定长卷积神经网络方面, 2015 年 He K 等人[24]提出了金字塔池化(Spatial Pyramid Pooling, SPP), SPP 可以不限制输入图片的大小, 将特征向量图划分为固定的区域数, 在每个区域进行最大池化操作, 最终生成固定长度的特征向量, 文献表示在图像分类上, SPP-net 相对于 no SPP-net 可以提高分类精确度; 而 Zhang Y 等人[25]也将不定长的卷积神经网络应用到了句子分类上, 并对模型中卷积核大小、池化操作等的选择进行了实验, 以此来探寻各类参数对模型性能的具体影响; 卓勤政[26]则使用上述文献中的模型对网络流量进行分类, 一种学习场景是对正常流量与恶意流量进行二分类, 另一种学习场景是对网络流量进行业务分类, 在两种学习场景下均可以达到 98%的准确率。

虽然现有的基于深度学习的分类模型均取得了不错的性能, 但是对数据分组进行截断或者补零, 在一定程度上会影响准确率。基于以上不足, 本课题提出了一种不定长卷积神经网络 ILCNN(Indefinite Length Convolutional Neural Network), 使用不定长的网络流量数据进行训练和流量分类, 避免了补零或截断操作对准确率带来的影响。

3 模型构建

因为全连接层需要确定单元数, 所以卷积神经网络一般会固定输入长度。在过去将卷积神经网络应用于恶意流量分类时, 会对流量数据分组进行补零或者截断, 将数据分组控制在固定的长度。但是补零或者截断都会对模型的训练产生一定的影响, 在对数据分组进行补零时, 补充的数据会参与到卷积计算中, 影响到特征的抓取等, 而对数据进行截断, 会导致数据分组中的部分信息丢失, 导致数据分组中的特征不能全部被学习到。因此, 本文的工作主要是为了避免对流量数据分组进行补零或者截断所带来的影响, 从而设计了不定长输入的卷积神经网络。

卷积神经网络的特点在于一般采用原始信号直接作为网络的输入, 避免了复杂的特征提取过程; 卷积阶段利用权值共享减少了权值的数量进而降低了网络模型的复杂度; 同时池化操作利用图像局部相关性的原理对特征图进行抽样, 在保留有用信息的同时有效的减少数据量。卷积神经网络当中卷积操作对输入的向量并没有固定长度的约束, 只是全

连接层对输入向量有固定长度的约束, 若在全连接层前通过池化操作将不定长的输入向量转化为定长向量, 那么不定长输入的卷积神经网络是可行的, 实验结果也表明了不定长输入的卷积神经网络在恶意流量分类问题上是存在优势的。

虽然本文提出的模型不需要对流量数据分组进行补零或者截断, 但是也需要对数据进行相应的预处理, 以保证数据分组以合适的格式输入到模型中, 保证模型训练以及使用模型进行预测的顺利进行。关于数据预处理以及模型架构的具体细节, 接下来将进行详细讲解。

3.1 数据预处理

本文工作选取的是 CICIDS-2017 数据集中的原始数据包, 原始数据包的形式为 PCAP 格式, 但是 PCAP 格式的原始数据包并不能直接输入进深度神经网络进行训练学习, 需要经过一系列的预处理, 将数据分组处理成 CSV 格式文件, 通过读取 CSV 文件将数据分组转化为 Tensor。接下来将对原始数据流的预处理过程进行详细讲解。数据的预处理流程如图 1 所示, 具体讲解如下:

(1) 输入 PCAP 原始数据包: 在 CICIDS-2017 数据集中, 每一天收集到的所有数据分组为一个 PCAP 文件, 根据 PCAP 文件格式, 去除掉 PCAP 文件的头部信息, 开始循环读取该文件流量样本中的数据分组, 直至读取到最后一个数据分组;

(2) 数据分组过滤: 在 PCAP 文件中, 不仅包含了本文当中需要用到的恶意流量数据分组, 还包括了一些其他数据分组, 比如 DHCP、APR 等数据包的数据分组, 需要将此类数据分组进行过滤, 方便于后期模型的训练以及预测;

(3) 数据分组信息处理: 经过过滤的数据分组均为恶意流量数据分组, 但是需要去除每个数据分组中以太网协议的头部信息以及 IP 协议中的源和目的 IP 地址, 以太网协议的头部信息对于数据分组的分类不能提供有用信息; 在数据分组标记中需要使用到 IP 协议中的源和目的 IP 地址, 为了避免神经网络模型直接使用源和目的 IP 地址进行分类, 删除掉源和目的 IP 地址避免影响;

(4) 标记数据分组: 在数据分组处理结束之后, 需要对每个数据分组进行攻击类型的标记。在 CICIDS 的官方文档中给出了每一类攻击的具体描述 (包括攻击时间、攻击方 IP 地址、受害方的 IP 地址等), 可以据此对数据分组进行标记;

(5) 样本类别不均衡处理: 在每天的 PCAP 文件经过处理后, 可以得到所有带标记的恶意流量数据分组。深度神经网络在训练时, 基于类别数据均衡的假设之下, 但是在实际的网络环境当中, 网络中的各类攻击流量分布并不均衡, 因此本文将最小数据量的类别数量作为标准, 从各类攻击流量中随机选取相同数量的数据分组作为数据集, 根据模型需要再分为训练数据集、验证数据集与测试数据集;

(6) 归一化: 网络流量数据分组中的每一个字节由 8bit 组成, 大小分布于 0~255, 非常类似于灰度图像中的一个像素点, 因此数据分组中的每个字节为一个像素点, 组成一个一维灰度图像, 并对每个像素点中的数据进行归一化, 方便后续模型的训练及计算。

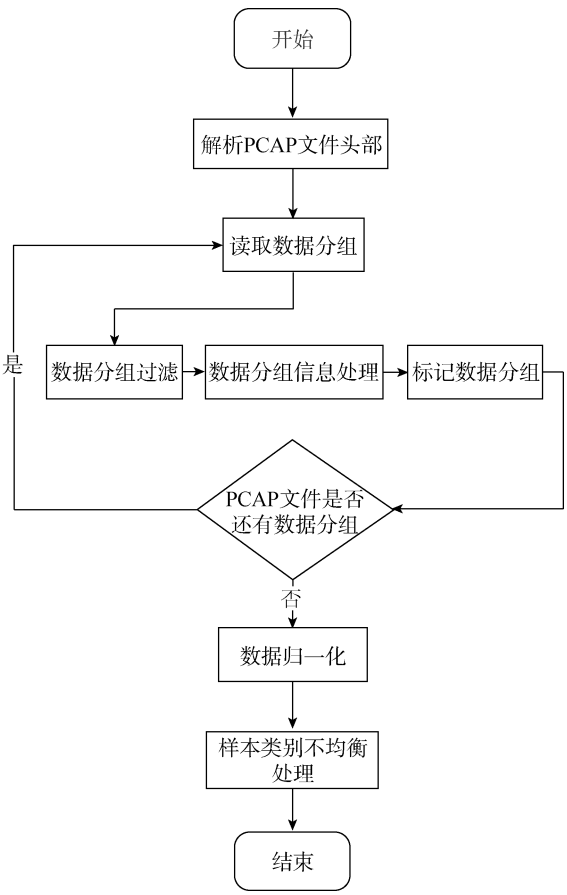


图 1 数据预处理流程图

Figure 1 Data preprocessing flow chart

3.2 模型结构

本文提出的不定长输入卷积神经网络 ILCNN 主要包括三个部分。第一个部分是多层卷积; 第二个部分是转化层, 将不定长的向量转化成定长向量, 包括一层卷积和一层池化, 得到的结果进行平铺输入到全连接层中; 第三部分是多层的全连接层, 最后接入 softmax 分类器, 完成对恶意流量的分类。网络模型结构如图 2 所示, 其中各种参数均为通过调参

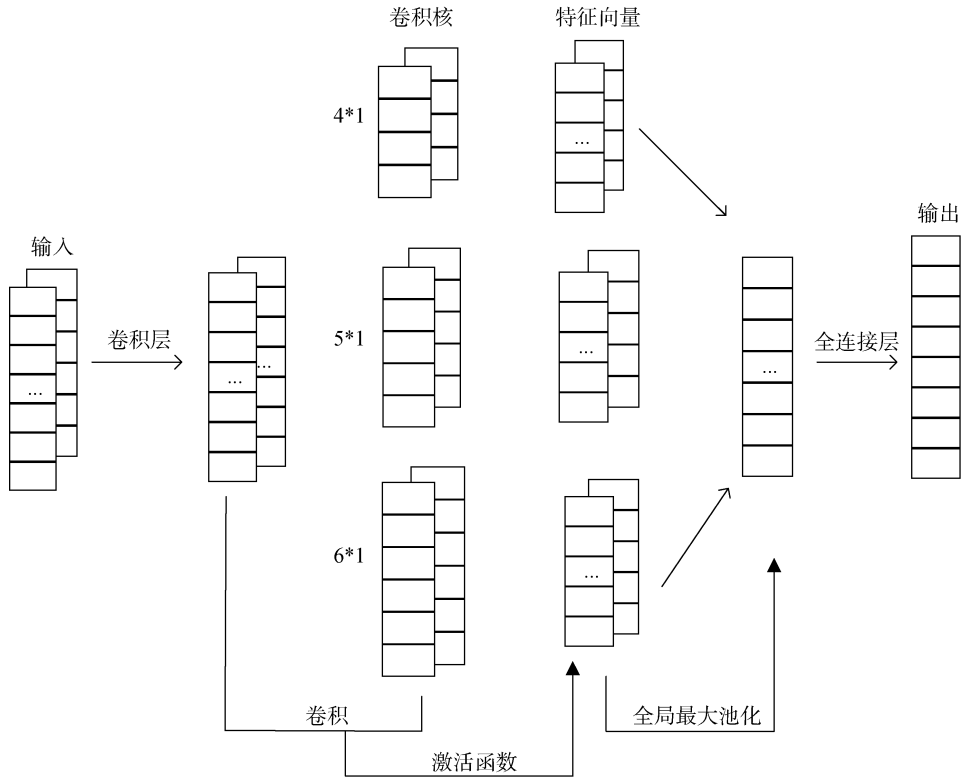


图 2　卷积神经网络结构图

Figure 2　Convolutional neural network structure diagram

得到, 接下来将进行详细讲解。

第一部分是多层卷积, 包括两层不带池化操作的卷积层, 第一层卷积层滤波器大小为 3, 数量为 64; 第二层卷积层滤波器大小为 4, 数量为 64, 步长均为 1。输入为经过数据预处理的流量数据分组, 经过卷积运算的向量再经过激活函数 Selu 进行运算, 输出为特征向量, Selu 运算定义为:

$$f(x)=\begin{cases} scale*x & if\ x>0 \\ scale*alpha*(\exp(x)-1) & if\ x<0 \end{cases} \tag{1}$$

其中 *scale* 和 *alpha* 为预先定义好的常数, 数值分别为 1.05070098 和 1.67326324。

第二部分是模型结构中最重要的部分, 如图 3 所示, 包括一层卷积和一层池化, 从第三节对卷积神经网络的相关介绍中可以知道, 卷积神经网络中的卷积操作是权值共享的, 需要进行修改的权值是卷积核中的参数, 对输入的向量并没有固定长度的约束, 但是全连接层对输入向量有固定长度的约束, 因此该部分使用全局最大池化操作将不定长向量转换成定长向量。该部分模型的输入为长度为 k 的特征向量, 通过与不同大小的卷积核(相关的卷积核参数如表 1 所示)进行卷积操作, 分别能够得到长度为 k–3+1、k–4+1、k–5+1 的特征向量, 相比于使用单个大小的卷积核, 使用多个大小的卷积核, 能够抓取流量数据分组中更多类型的特征。卷积运算得到的结果经过激活函数 Selu 进行运算, 输出到池化层, 池化层通过全局最大池化得到长度为 1 的特征向量, 所有特征向量进行连接, 能够得到长度为 32*3 的特征向量, 该特征向量接下来将会输入到全连接层中进行分类。

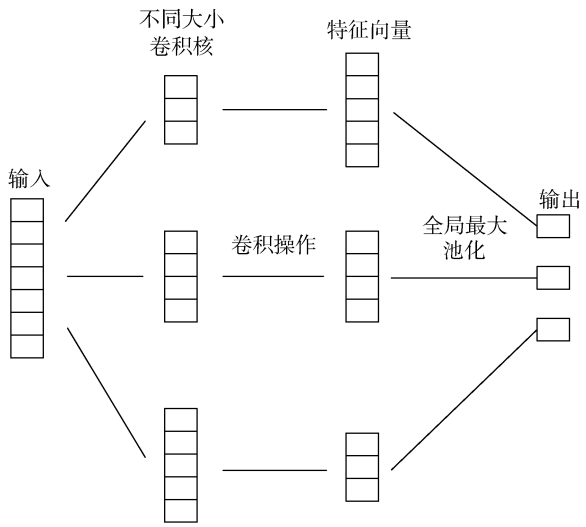


图 3　转化层网络结构图

Figure 3　Conversion layer network structure diagram

表 1 卷积核参数表

Table 1 Convolution kernel parameter table

	卷积核大小	卷积核数量	步长
转化层卷积核参数	4	32	1
	5	32	1
	6	32	1

第三部分是全连接层与 softmax 输出层, 全连接层神经单元数量为 32, softmax 输出层神经单元数量为 8, 使用的激活函数分别为 Selu 和 softmax, softmax 函数运算定义为:

$$y_j = \frac{\exp(x_j)}{\sum_{j=1}^{8}\exp(x_j)} \tag{2}$$

另外需要提到的是, 在网络模型的训练过程中使用的损失函数为交叉熵损失函数(Cross Entropy Loss function), 其定义如下:

$$J(w,b) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log\hat{y}^{(i)} + (1-y^{(i)})\log(1-\hat{y}^{(i)})\right] \tag{3}$$

4 实验结果与分析

4.1 实验准备

使用 keras 作为深度学习库, 并使用 Tensorflow 2.1 GPU 版本, 实现本文提出的卷积神经网络模型结构 ILCNN。硬件平台采用 NVIDIA Tesla V100 32GB 用于模型的训练和测试。为了验证网络模型的性能, 我们从 CICIDS-2017 数据集中的原始数据包提取出适应于深度学习的流量数据分组集, 对网络模型进行训练及评估。

在网络模型训练中, 本文将数据集进行随机分配, 大部分数据继续用于模型的训练, 分割出小部分数据用于对模型进行验证, 以得到具有更加稳定性能的网络模型。本文中将流量数据分组集划分为三个部分, 第一个部分占比数据集的 70%, 用于网络模型的训练, 第二个部分占比数据集的 15%, 用于网络模型的验证, 第三个部分占比数据集的 15%, 用于测试网络模型的性能。另外本文使用了 Early Stopping 策略, 该策略是指在网络模型的训练过程中, 一直监测验证集上损失函数的结果, 如果在连续几轮的训练过程中, 验证集的 loss 没有较明显的变化(变化值通过自定义设置), 训练过程就会提前停止, 其目的是为了避免网络模型出现过拟合的问题。

4.2 实验数据集

模型使用 CICIDS-2017 数据集, 该数据集是由加拿大网络安全研究所提供的公开数据集, 包含了良性的背景流量以及常见的攻击流量。加拿大网络安全研究所另外提供了 CICIDS-2018 的数据集, 该数据集与 2017 数据集所包含的攻击类型并无太大差异, 且均为通过模拟攻击对流量数据进行捕获, 但是在研究中更多学者使用的是 2017 数据集, 为了方便进行性能对比, 本文使用 2017 的数据集。数据集中捕获的原始数据包是 PCAP 文件, 同时该数据集也提供了经过特征提取、带有标记的 CSV 文件, 但是在本文中使用的是原始数据流, 即 PCAP 文件。

数据集的数据捕获从 2017 年 7 月 3 日周一 9:00 开始, 一直到 2017 年 7 月 7 日周五的 17:00 结束, 其中周一仅包括良性的背景流量, 其余的 4 天包含了各种攻击流量, 捕获的数据集共约 51.1G。该数据集捕获的攻击流量中包含 Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet 以及 DDoS, 涵盖了基于 2016 年 McAfee 报告的最常见攻击。

4.3 评价指标

最好的评估流量分类模型的方法是将算法应用到实际网络环境中, 但是因为所需的代价太大, 我们使用测试集上的网络性能表现来评估网络模型性能。在评估网络模型性能时我们使用了以下指标, 包括精确率(Precision), 召回率(Recall)以及 F1-score:

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1\text{-}score = \frac{2*precision*recall}{precision+recall} \tag{6}$$

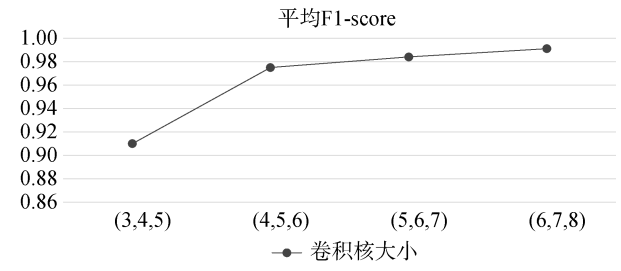
其中, True Positive(TP)表示检测模型将攻击类型正确识别为攻击类型的样本个数, False Positive(FP)表示检测模型将别的攻击类型错误识别为该类攻击类型的样本个数, False Negative(FN)表示检测模型将该类攻击类型错误识别为别的攻击类型的样本个数。

4.4 实验结果展示

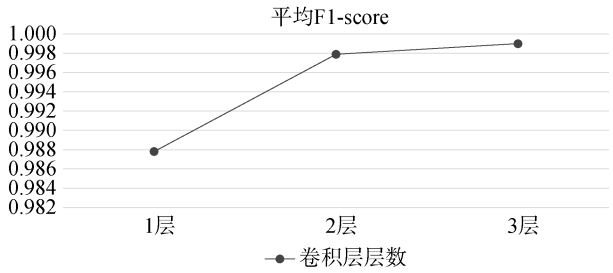
如第四节所述, 为了找到一个最佳的卷积神经网络结构, 本文依据实验结果, 对卷积层的数量以及卷积核的大小进行反复调整, 以求得合适的超参数。由于计算硬件的限制, 我们仅对超参数的部分可能值进行了实验, 以得到在测试集上每一类攻击类

型的 F1-score, 以及整体测试集上的准确度。由于最佳模型并没有一个明确的定义, 因此本文在对模型的选择上考虑了 F1-score、准确度以及网络复杂程度, 但是这并不代表本文最终选择的模型就一定是适合于恶意流量分类的最佳模型。

本文对网络模型结构中的滤波器大小以及卷积层数进行了实验, 使用在测试集上的平均 F1-score 作为评价指标, 图 4 对结果进行了展示。



(a) 卷积核大小对模型性能的影响
(a) The effect of convolution kernel size on model performance



(b) 卷积层层数对模型性能影响
(b) The impact of the number of convolutional layers on model performance

图 4 超参数对模型性能影响

Figure 4 The impact of hyperparameters on model performance

如图 4(a)所示是对将变长向量转换为定长向量的卷积加池化层的滤波器大小进行修改, 滤波器大小分别是(3、4、5)、(4、5、6)、(5、6、7)、(6、7、8), 随着滤波器大小的增加, 模型能够具有更好的性能, 滤波器大小从(3、4、5)增加到(4、5、6)时, 性能的提升比能够达到 7.5%, 但是再继续增加滤波器大小, 性能的提升比只有 0.9%左右, 为了使模型具有较好的准确率以及考虑到模型的复杂度, 滤波器大小为 4、5、6 是比较适合的参数; 如图 4(b)所示是对模型中第一部分卷积层层数的修改, 具体的卷积层参数如表 2 所示, 卷积层层数的增加, 使得网络结构能够抓取更加深层次的流量特征, 也使得网络结构的性能会更好, 但是当卷积层增加到一定层次后, 性能的增加几乎为零, 这一方面可能与卷积层中滤波器大小选择不当有关, 但是由于计算机硬件限制本文没有进行更多的实验, 另一方面可能与网络复杂程度增加有关。

当网络复杂程度增加时, 能够带来更好的性能, 但是当网络复杂程度持续增加时, 性能提升比会逐渐下降, 甚至可能出现性能下降的现象, 一方面是因为当网络复杂程度增加时, 深度神经网络中需要训练的参数会增多, 而在训练数据不变的情况下, 很可能现有的训练数据不足以支持更复杂的网络训练, 第二方面是更复杂的网络模型更有可能面临梯度消失等问题, 导致训练阶段拟合不足。

表 3 列出了本文中所得到的模型在测试集上进行恶意流量分类所获得的性能, 该模型在恶意流量分类任务上平均 F1-score 能够达到 0.9992, 这表明本文中所提出的 ILCNN 能够从恶意流量的数据分组中训练出可区分的特征, 并且能够成功的对恶意流量的数据分组进行分类。

4.5 模型分析

为了证明本文提出的不定长输入卷积神经网络的有效性, 对神经网络模型中输入到全连接层的特征向量进行提取, 使用 t-SNE 对提取出的高维特征向量进行降维处理, 降维到二维便于可视化, 从可视化图中观察特征向量的分布。从图 5 中可以看出, 每一类恶意流量经过卷积神经网络进行提取的特征向量在降维处理后都分布在一处, 不会过于分散; 且不同类型恶意流量的簇之间不会相互有交集, 因此可以认为本文提出的模型能够很好的学习恶意流量中的可区分特征, 学习到的可区分特征能对不同类型的恶意流量进行划分, 网络结构中后续的全连接层和 softmax 分类器可以根据这些良好的特征向量进行恶意流量的分类。

表 2 卷积层各层参数

Table 2 Parameters of each layer of the convolutional layer

卷积层数	卷积核大小	激活函数	卷积核大小	激活函数	卷积核大小	激活函数
1	6	selu				
2	3	selu	4	selu		
3	2	selu	3	selu	4	selu

表 3 恶意流量分类性能

Table 3 Malicious traffic classification performance

恶意流量	精确率	召回率	F1-score
Brute Force FTP	1.0	0.9996	0.9998
Brute Force SSH	1.0	0.9986	0.9993
DoS	0.9996	1.0	0.9998
Heartbleed	0.9993	1.0	0.9996
Web Attack	0.9996	0.9990	0.9993
Infiltration	1.0	0.9966	0.9983
Botnet	0.9956	1.0	0.9978
DDoS	0.9993	0.9996	0.9995
Average	0.9992	0.9992	0.9992

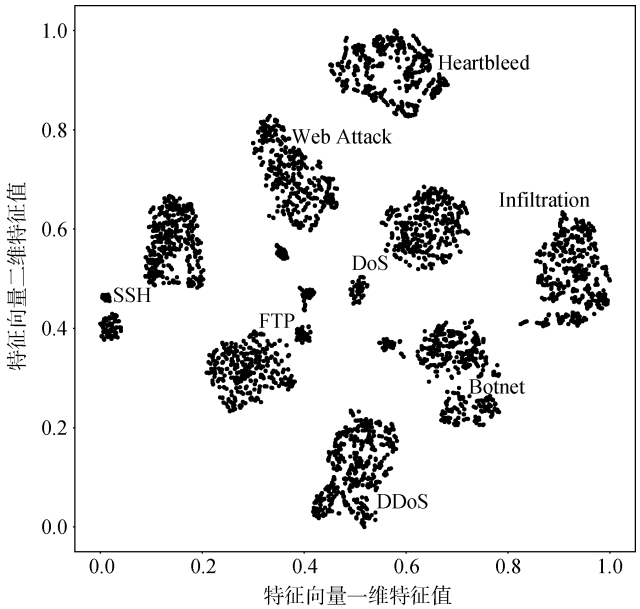


图 5 特征向量分布图

Figure 5 Feature vector distribution map

4.6 本文工作与其他工作的比较

另外使用 CICIIDS-2017 数据集的工作有很多, Zhang 等人[15]将改进的 LetNet-5 卷积神经网络与 LSTM 神经网络组合成分层网络, 其中卷积神经网络用来抓取流量的空间特征, LSTM 神经网络用来抓取流量的时间特征, 而 ILCNN 使用不定长输入的卷积神经网络, 其网络复杂度相对来说会更低; Zhang 提出的方法是基于原始数据流进行的恶意流量分类任务, 对每个数据包截取前 160 个字节, 每个数据流选取前 10 个数据包共 1600 个字节, 处理成 40*40 的二维矩阵进行实验, ILCNN 是基于原始数据包进行的恶意流量分类任务, 无需对数据包进行截断或者补零操作, 在数据预处理上会更节省步骤, 也不会对数据加入一些对网络训练无用的信息; 最终实验结果对比如表 4 所示, 结果表明 ILCNN 在性能上有细微的优势, 虽然召回率会略低于 Zhang 提出的方法, 但是准确率能够高出 1.14‰, F1-score 高出 0.05‰。

表 4 与其他工作的比较

Table 4 Comparison with other work

指标 / 模型	准确率	精确率	召回率	F1-score
ILCNN	0.999249	0.999211	0.999208	0.999208
Zhang 的模型[15]	0.998111	0.998475	0.999847	0.999161
Ferrag M A 的模型[27]	0.9981	—	—	—

Ferrag M A 等人[27]基于区块链以及深度学习提出了智能电网能源框架, 框架中使用循环神经网络实现了入侵检测部分, 使用的数据集包括 CICIDS-2017 数据集、power system 数据集以及 web robot (Bot)-Internet of Things (IoT)数据集, 接下来的对比中只会讨论 CICIDS-2017 数据集部分。Ferrag M A 的方法使用的循环神经网络更善于抓取流量的时间特征, 而本文中的卷积神经网络更善于抓取流量的空间特征; Ferrag M A 使用的是 CICIDS-2017 的特征数据集, 即对原始数据流经过了特征提取, 共包括 79 个流量特征, 而 ILCNN 使用的是原始数据包, 不需要经过特征抓取, 能够省去人工进行特征抓取的时间、精力消耗, 另外也可以避免特征提取对流量分类结果的影响; 最后在模型性能上, 结果对比如表 4 所示, 表中“—”标记表示相应的性能指标没有给出, Ferrag M A 的方法在循环神经网络中的隐藏节点达到 60 时, 能够达到 0.99811, ILCNN 能够达到 0.999249, 高出 Ferrag M A 的方法 1.14‰。

5 结论

本文提出了一种不定长输入的卷积神经网络 ILCNN。其处理对象是原始流量数据包, 不需要进行人工特征抓取造成人工消耗, 也不会造成流量信息损失从而影响分类准确率。实验结果表明, ILCNN 可以以 0.9992 的准确率处理恶意流量分类任务, 性能优于同样使用 CICIDS-2017 数据集的同类工作。

后续工作中, 将考虑提高网络模型的泛化能力, 能够检测更多的攻击类型, 以及能够通过小量样本的训练检测出低频攻击。

参考文献

[1] Ahmed M, Mahmood A N, Hu J. A survey of network anomaly detection techniques[J]. *Journal of Network and Computer Applications*, 2016, 60: 19-31.

[2] Yoon S H, Park J W, Park J S, et al. Internet application traffic classification using fixed IP-port[C]. *Asia-Pacific Network Operations and Management Symposium. Springer, Berlin, Heidelberg*, 2009: 21-30.

[3] Madhukar A, Williamson C. A longitudinal study of P2P traffic classification[C]. *14th IEEE International Symposium on Modeling, Analysis, and Simulation. IEEE*, 2006: 179-188.

[4] Thay C, Visoottiviseth V, Mongkolluksamee S. P2P traffic classification for residential network[C]. *2015 International Computer Science and Engineering Conference* (*ICSEC*). *IEEE*, 2015: 1-6.

[5] Lin P C, Lin Y D, Lai Y C, et al. Using string matching for deep packet inspection[J]. *Computer*, 2008, 41(4): 23-28.

[6] Sharafaldin I, Lashkari A H, Ghorbani A A. Toward generating a new intrusion detection dataset and intrusion traffic characterization[C]. *ICISSp*. 2018: 108-116.

[7] Go J H, Jan T, Mohanty M, et al. Visualization Approach for Malware Classification with ResNeXt[C]. *2020 IEEE Congress on Evolutionary Computation (CEC)*. *IEEE*, 2020: 1-7.

[8] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 1492-1500.

[9] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.

[10] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2017, 31(1).

[11] Pascanu R, Stokes J W, Sanossian H, et al. Malware classification with recurrent networks[C]. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. *IEEE*, 2015: 1916-1920.

[12] 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: 43-48.

[13] Radford B J, Richardson B D, Davis S E. Sequence aggregation rules for anomaly detection in computer network traffic[J]. *ArXiv Preprint ArXiv*: 1805.03735, 2018.

[14] Wu Di, Fang Binxing, Cui Xiang, Liu Qixu.BotCatcher: A Botnet Detection System Based on Deep Learning[J].*Journal of Communications*, 2018, 39(08): 18-28.
(吴迪, 方滨兴, 崔翔, 刘奇旭.BotCatcher: 基于深度学习的僵尸网络检测系统[J].*通信学报*, 2018, 39(08): 18-28.)

[15] Zhang Y, Chen X, Jin L, et al. Network intrusion detection: Based on deep hierarchical network and original flow data[J]. *IEEE Access*, 2019, 7: 37004-37016.

[16] Dong Y, Wang R, He J. Real-Time Network Intrusion Detection System Based on Deep Learning[C]. *2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS)*. IEEE, 2019: 1-4.

[17] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in neural information processing systems*, 2012, 25: 1097-1105.

[18] Zeng Y, Gu H, Wei W, et al. $ Deep-full-range $: A deep learning based network encrypted traffic classification and intrusion detection framework[J]. *IEEE Access*, 2019, 7: 45182-45190.

[19] Draper-Gil G, Lashkari A H, Mamun M S I, et al. Characterization of encrypted and vpn traffic using time-related[C]. *Proceedings of the 2nd international conference on information systems security and privacy (ICISSP)*. 2016: 407-414.

[20] 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.

[21] Prasse P, Machlica L, Pevný T, et al. Malware detection by analysing network traffic with neural networks[C]. *2017 IEEE Security and Privacy Workshops (SPW)*. IEEE, 2017: 205-210.

[22] Nayyar S, Arora S, Singh M. Recurrent Neural Network Based Intrusion Detection System[C]. *2020 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2020: 0136-0140.

[23] Torroledo I, Camacho L D, Bahnsen A C. Hunting malicious TLS certificates with deep neural networks[C]. *Proceedings of the 11th ACM workshop on Artificial Intelligence and Security*. 2018: 64-73.

[24] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2015, 37(9): 1904-1916.

[25] Zhang Y, Wallace B. A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification[J]. ArXiv Preprint ArXiv: 1510.03820, 2015.

[26] Zhuo Qinzheng. Research on network traffic analysis based on deep learning[D]. Nanjing University of Science and Technology, 2018.
(卓勤政. 基于深度学习的网络流量分析研究[D].南京理工大学, 2018.)

[27] Ferrag M A, Maglaras L. DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids[J]. *IEEE Transactions on Engineering Management*, 2019, 67(4): 1285-1297.

杨璇 于 2019 年在南京理工大学计算机科学与工程专业获得学士学位。现在东南大学网络空间安全专业攻读硕士学位。研究领域为流量分类。Email: 220194739@seu.edu.cn

赵博 于 2012 年在信息工程大学信息与通信工程获博士学位。现为信息工程大学信息技术研究所副研究员。研究方向: 拟态防御架构, 芯片设计。Email: lieut.zhao@126.com

邬江兴 国家数字交换系统工程技术研究中心主任, 教授, 中国工程院院士, 主要研究方向为拟态计算、内生安全、多模态网络等。Email: ndscwjx@126.com