DOI:10.19651/j.cnki.emt.2209148

一种基于图卷积神经网络的加密流量分类方法*

王勤凡 翟江涛 陈 伟 孙浩翔
(南京信息工程大学电子与信息工程学院 南京 210044)

摘 要: 深度学习算法被广泛应用于网络流量分类领域并取得较好效果。然而对抗攻击的出现给其安全性带来了严重威胁,使得当前主流的基于卷积神经网络模型的分类算法的精度严重下降。针对此提出了一种抗流量分类中灰度图对抗攻击的加密流量分类方法。所提方法通过提取数据包负载长度、包序列、方向、簇等流量交互信息构建拓扑图,将加密流量分类问题转化为图分类问题。使用基于图卷积神经网络的分类方法进行特征的学习分类,图卷积神经网络模型可以自动从输入的拓扑图中提取特征,将特征映射到嵌入空间中的不同表示来区分不同的图结构。实验结果表明,所提方法不仅能够避免对抗攻击,且在公开数据集上的分类性能也较现有典型方法提高了 5%以上。

关键词: 网络流量分类;对抗攻击;图神经网络;深度学习
中图分类号: TP393.0 **文献标识码:** A **国家标准学科分类代码:** 413.20

An encrypted traffic classification method based on graph convolutional neural networks

Wang Qinfan Zhai Jiangtao Chen Wei Sun Haoxiang
(College of Electronical and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Deep learning algorithms are widely used in the field of network traffic classification and have achieved good results. However, the emergence of adversarial attacks has brought a serious threat to its security, and the accuracy of the current mainstream classification algorithms based on convolutional neural network models has been seriously reduced. In response to this, this paper proposes an encrypted traffic classification method that resists gray-scale adversarial attacks in traffic classification. The proposed method constructs a topology graph by extracting traffic interaction information such as packet load length, sending order, direction, and cluster, and transforms the encrypted traffic classification problem into a graph classification problem. Then, this paper uses the classification method based on graph convolutional neural network to learn and classify features. The graph convolutional neural network model can automatically extract features from the input topology and map features to different representations in the embedding space to distinguish different graph structures. The experimental results show that the proposed method can not only avoid adversarial attacks, but also improve the classification performance on public datasets by more than 5% compared with the existing typical methods.

Keywords: network traffic classification; adversarial attacks; graph neural network; deep learning

0 引 言

网络流量分类技术主要应用于网络行为分析、优化网络传输、异常检测等领域,以满足差异化 QoS 需要和保证网络的安全运行,是网络流量处理研究领域的热点问题之一。基于机器学习的流量分类技术在特征选取方面具有局限性,因此,基于深度学习[1-3]方法应运而生,基于深度学习的加密流量分类算法采用的是端到端的方式,可以通过神经网络对原始流量数据进行自主学习,避免了人工设计并提取特征的过程,很大程度上节省了人力成本。卷积神经网络(convolutional neural networks,CNN)模型作为当前最典型的深度学习模型之一,被研究者广泛应用于流量分类任务中。文献[4-6]均采用了将流量数据转换为灰度图的预处理方法,得到的灰度图数据输入 CNN 模型进行特征学习并最终分类,达到了较高的分类精度。

基于深度学习的加密流量分类方法在流量分类任务中

收稿日期:2022-03-03
* 基金项目:国家自然科学基金(61931004,62072250)、南京信息工程大学人才启动基金(2020r061)项目资助

表现较好，然而 Szegedy 等[7]在对计算机视觉领域中的对抗攻击研究时发现，现有的深度学习算法本身存在着很大的缺陷，仅仅在原始样本图片上添加一个很小的扰动，就可以达到欺骗深度学习模型将样本误分类的目的，使得分类精度严重下降。而这种扰动是人类的肉眼不可察觉的，这种现象被称之为对抗攻击，添加了扰动后的样本被称为对抗样本。目前已经有大量的对抗攻击方法被提出，主流的对攻击方法有 Goodfellow 等[8]提出的快速梯度符号法(FGSM)，Carlini 等[9]提出了基于优化的对抗攻击方法(C&W)，Madry 等[10]提出的投影梯度下降法(PGD)等。研究表明，在计算机视觉图像领域，这些对抗攻击方法都对图像分类有着巨大的干扰性，即便是最简单的单步攻击，都可以导致模型以高置信度输出错误的预测。当前主流的对抗防御方法之一就是通过训练对抗样本提高模型的鲁棒性，Zheng 等[11]在 MNIST 数据集上利用对抗训练使得对抗攻击后的分类精度达到 88.79%，这表明经过对抗训练后，模型对大部分的对抗样本有了判别能力。

随着研究的不断进展，研究者发现对抗攻击不仅在视觉领域中对图像分类有着较大影响，也严重威胁了网络流量分类的分类效果。在网络安全领域，当前主流的流量分类方法是 CNN 模型进行流量特征学习并分类，该方法首先将流量数据经过一定的预处理转换成灰度图像，而对灰度图像的攻击会对基于 CNN 的网络流量分类造成极大的影响。普通样本被攻击后生成的对抗样本可以很轻易的骗过基于 CNN 的流量分类模型，使其将样本错误地分类为其它类别，对流量分类任务造成极大的干扰。胡永进等[12]为了应对流量分类攻击，从防御者的角度出发，提出了一种基于对抗样本的网络欺骗流量生成方法。羊洋等[13]验证了对抗攻击对基于 CNN 模型的网络流量分类同样有极大的影响，并提出了基于混合对抗训练的防御措施，将对抗攻击形成的对抗流量样本和原始流量样本混合训练以增强分类模型的鲁棒性，并取得到较好的分类效果。然而，相较于未被攻击前的分类准确率仍有一定的差距。如何进一步降低对抗攻击对于流量分类的影响依然是当前需要研究的重点问题。

文献[14-15]分别采用 GCN 和 GNN 模型对网络流量和恶意软件进行分类并取得不错的分类结果，证明了 GCN 模型在流量分类上的可行性。其中文献[14]采用的流量构图方式是基于主机行为的方法，它通过分析网络流量跟踪图中的主机行为或局部结构来对网络流量进行分类。该方法在处理流量分类时依靠的是主机在传输层的行为模式，虽然该方法不需要读取数据包负载信息，但是也存在一定的缺陷，对于传输层加密的流量无法进行分类。基于此，本文提出了一种利用流量会话流中数据包的交互信息的非欧几里得构图方式，将会话流表示为数据包交互拓扑图，并通过图神经网络模型进行图的学习和分类，在实现较高分类精度的同时也避免了对抗攻击的影响。

所提算法基于数据流中数据包的交互信息，即会话流中的数据包负载长度、方向、包序列、簇等特征构建流量交互拓扑图，然而再通过图卷积神经网络(graph convolutional networks，GCN)模型进行训练分类。GCN 模型可以自动从输入的拓扑图中提取特征，并通过将特征映射到嵌入空间中的不同表示来区分不同的图结构。本文主要贡献如下：

1)分析了对抗攻击对基于 CNN 的网络流量分类模型的影响，并通过对 UNB ISCX VPN-nonVPN 2016 加密流量数据集和 USTC-TFC2016 加密流量数据集进行了对抗攻击实验进行验证。

2)提出了将流量样本转化成拓扑图结构的预处理方法，简化预处理步骤，使用极少的特征就可以达到极高的分类效果，并且避免了基于 CNN 的分类方法中将流量转化成的灰度图会受到的对抗攻击威胁。

3)提出了基于 GCN 的分类算法，算法以 GCN 模型为主体，通过图卷积层自主地在输入的拓扑图中提取特征，将拓扑图映射到嵌入空间中的图表示并进行图学习。本文使用 UNB ISCX VPN-nonVPN 2016 加密流量数据集和 USTC-TFC2016 加密流量数据集对所提分类算法进行性能测试，实验结果表明了所提分类方法的有效性。

1　本文方案

现有对抗防御方法不能完全避免对抗攻击对基于 CNN 模型的流量分类算法的影响，导致当前典型的将流量数据转换为灰度图像的处理方法面临着诸多安全问题。因此，本文从流量数据处理方式的角度出发，提出了一种基于 GCN 模型的加密流量分类算法，所提算法首先根据五元组对数据集中的流量数据按照会话级别进行分流，然后利用会话流中的数据包负载长度、方向、包序列、簇信息作为特征构建流量交互拓扑图，再通过有监督图学习的方式，采用 GCN 模型从不同类别流量构成的拓扑图中学习图特征表示，最后通过分类器对拓扑图分类，以此实现对加密流量的分类。该过程的整体框架结构如图 1 所示，主要分为以下 3 个步骤：

1) 数据预处理：将原始流量按会话粒度进行划分，提取会话流中用于构建拓扑图的信息。

2) 图构建：利用会话流中数据包负载长度、方向、包序列、簇特征构建拓扑图，拓扑图中节点对应会话流中的数据包，以数据包负载长度作为对应节点的节点特征，节点之间的边对应数据包在传输顺序和传输方向上的交互。

3) 图分类：以包含节点属性和标签的图作为输入，通过图卷积网络模型进行图表示的学习，最后通过分类器进行分类。

1.1　拓扑图构建算法

所提拓扑图构建算法首先是对原始流量按照会话粒度进行分流，其中流的定义是具有相同五元组的一系列数据

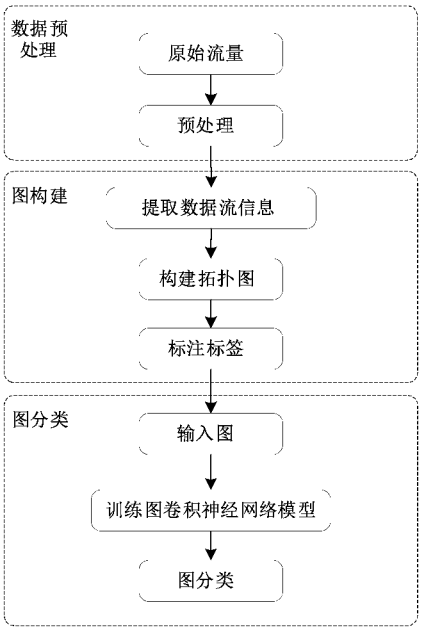
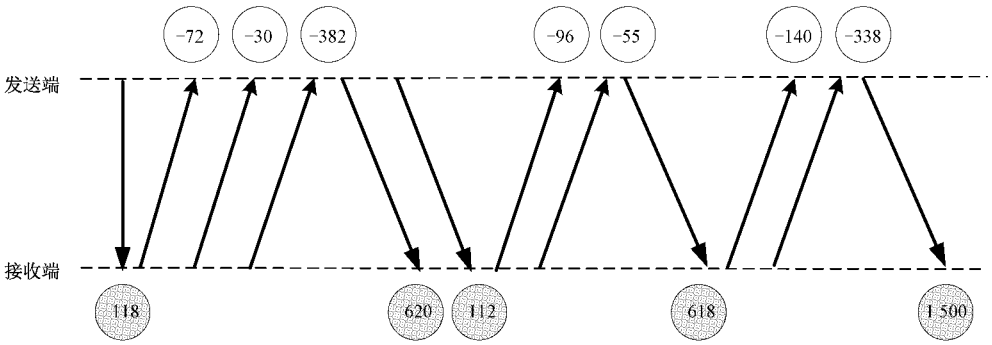


图 1 本文所提算法流程图

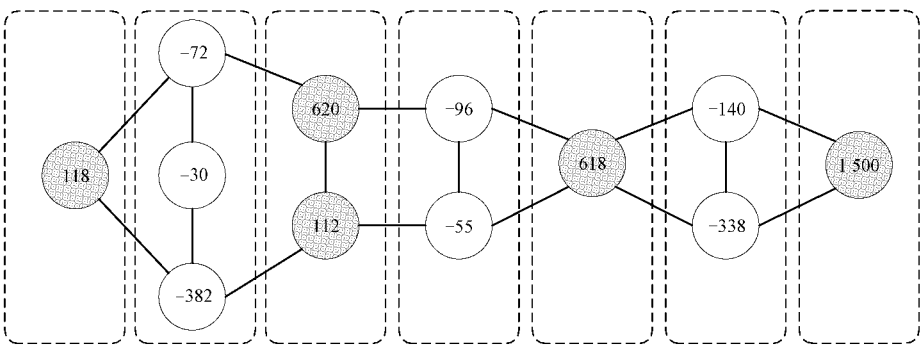
包构成的数据流,而会话流则是指由双向流组成的所有包,即五元组中的源端口和目的端口、源 IP 和目的 IP 可以互换。相较于单向流,研究者通常选择使用会话流进行流量分类,这是因为会话流中所包含信息比单向流更丰富。本文中我们将要利用到会话流中数据包交互的方向和次序信息。首先,我们提取会话流中各个数据包的负载长度,然后根据会话流中数据包的 IP 地址可以获取到数据包的方向信息。我们随机选择某条会话流为例,对构成拓扑图的几个要素进行说明。

1) 节点:给定一个特定的会话流,由该会话流构成的拓扑图中各个节点对应会话流中的各个数据包,且以数据包负载长度作为节点特征。为了反映数据包传输方向,我们保留了数据包负载长度的符号。我们将会话流中第一个数据包的传输方向设为正向,即将此数据包以及后面与之方向相同的数据包负载长度设置为正值,与之方向相反的数据包负载长度设置为负值。本文将沿同一方向连续传输的数据包称为簇,即使符合条件的只有一个数据包,也将其称为簇,如图 2(a)中向发送端连续发送的三个数据包对应的−72,−30,−382 的 3 个节点即为一簇。

2) 边:在本文所提构图算法构成的拓扑图中边的类型有两种:簇内边和簇外边。簇内边依次连接每个簇中连续的节点,簇外边连接前后相邻的两个簇,即两个相邻的簇的第 1 个节点相连,最后 1 个的节点都与后 1 个簇相应的第 1 个节点和最后 1 个节点相连接,当 1 个簇中只有 1 个数据包时,此数据包即是第 1 个点也是最后 1 个点,相连接的节点之间最多只能添加一条边。由图 2(a)中特定流的信息构成的拓扑图如图 2(b)所示。



(a) 会话流数据包交互图



(b) 流量交互拓扑

图 2 基于拓扑图的数据流表示

本文方法利用不同类别流量之间的会话流数据包交互特征的差异对流量进行分类,其中所采用的数据包交互特征具体包括以下 4 种:会话流中数据包的负载长度、方向、包序列和由数据包组成的簇的特征。上述特征由数据

流转换的流量交互拓扑图展现，如图 2(b)所示。其中，拓扑图中节点上的数字表示数据包负载长度，数字的正负则代表数据包传输的不同方向。而会话流中连续向相同方向发送的一系列数据包所对应的拓扑图中的节点则体现了会话流的簇特征，如图 2(b)中的各个方框内一系列节点表示为图 2(a)中的一簇。拓扑图中节点的位置则体现了包序列特征，对于不同簇的数据包，拓扑图中节点的横向位置关系代表了包序列中数据包发送的先后顺序，左边节点代表的数据包发送顺序先于右边节点；对于同一簇中的数据包，包序列中数据包发送顺序则由拓扑图中节点的纵向位置关系所表示，上边节点代表的数据包发送顺序先于下边节点。

1.2　基于 GCN 的图分类算法

给定一组拓扑图 $\{G_1,\cdots,G_N\}\subseteq\mathcal{G}$ 以及标签 $\{y_1,\cdots,y_N\}\subseteq\mathcal{Y}$，训练 GCN 模型的目的在于通过学习能够预测每个拓扑图标签的表示向量 $\boldsymbol{H}_G$，即 $y_N=g(\boldsymbol{H}_G)$。基于 GCN 的分类模型由图卷积层和全连接层构成，图卷积层用于从拓扑图中提取特征进行训练，全连接层用于分类，模型使用交叉熵损失函数度量实际标签和预测标签之间的差异信息。基于 GCN 的分类模型的网络结构如图 3 所示。

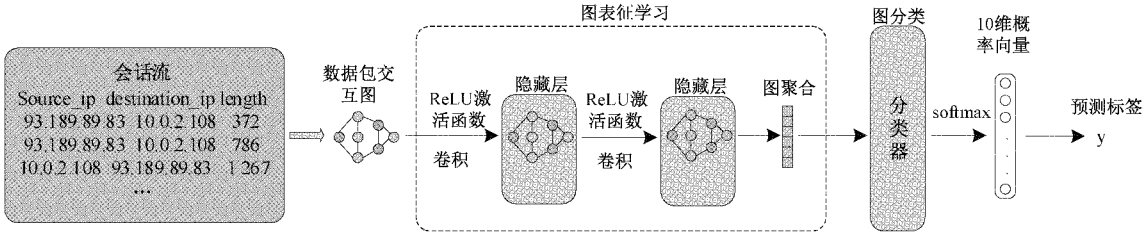


图 3　基于 GCN 的分类模型网络结构

1) 图卷积层：在图拓扑中，每个节点间相互联系，图卷积运算的方法是在更新节点状态信息时，聚合了邻居节点的状态信息，也就是将邻居节点的特征聚合到了中心节点上。如下：

$$\boldsymbol{H}_{G_i}^{(l-1)}=\sigma(\tilde{\boldsymbol{D}}-\frac{1}{2}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}-\frac{1}{2}\boldsymbol{H}_{G_i}^{(l)}\boldsymbol{W}^{(l)}) \tag{1}$$

式中：$\tilde{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{I}_N$ 是引入自循环的无向图 $\mathcal{G}$ 的邻接矩阵，$\boldsymbol{I}_N$ 是单位矩阵，$\tilde{\boldsymbol{D}}$ 为 $\tilde{\boldsymbol{A}}$ 的度矩阵，$\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}-\frac{1}{2}$ 是对 $\tilde{\boldsymbol{A}}$ 进行对称归一化，$H_{G_i}^{(l)}$ 是上一个卷积层的输出，初始特征矩阵设定为 $\boldsymbol{H}_{G_i}^{(0)}=\boldsymbol{X}_i$，$\boldsymbol{X}_i$ 为节点的嵌入特征，$\boldsymbol{W}^{(l)}$ 是第 l 层的权重矩阵，$\sigma(\cdot)$ 是非线性激活函数，本文采用的是 $\mathrm{ReLU}(\cdot)=\max(0,\cdot)$。

2) 全连接层：GCN 在图卷积层之后使用了一个线性函数对卷积层的输出数据进行线性变换，并且为避免过拟合使用了 dropout 函数。将线性函数的输出作为每个拓扑图 G_i 的特征表示向量 $\boldsymbol{H}_{G_i}$，且为了便于接下来的预测过程，需要将 $\boldsymbol{H}_{G_i}$ 映射到一个新的潜在空间 $\boldsymbol{H}_{G_i}\in R^C$，$C$ 是 $\mathcal{Y}$ 中不同元素的数量，即要分类的流量类别数量。然后利用 softmax 函数得到预测的概率向量 $\hat{\boldsymbol{y}}_{ic}$，表示 G_i 属于每种流量类型的可能性。如下式所示：

$$\hat{\boldsymbol{y}}_{ic}=soft\max(\boldsymbol{H}_{G_i}) \tag{2}$$

3) 损失函数：分类模型中的 GCN 采用交叉熵函数作为损失函数，交叉熵主要用于度量两个概率分布之间的差异性信息，在深度学习中作为损失函数时用于衡量真实标签和训练后模型的预测标签之间的相似性，且使用交叉熵函数作为损失函数的优点还有：相对于均方差损失函数，交叉熵函数更具收敛性。交叉熵函数公式如下：

$$\mathcal{L}=-\frac{1}{|N|}\sum_{i=1}^{|N|}\sum_{c=1}^{C}y_{ic}\log(\hat{y}_{ic}) \tag{3}$$

其中，$|N|$ 是训练的样本数量，y_{ic} 是实际标签。

4) 优化器：分类模型中的 GCN 采用了 Adam 优化器，Adam 是一种可以替代传统随机梯度下降过程的一阶优化算法，它能够基于训练数据迭代地更新神经网络权重，实现简单高效且对只需要较小的内存。

2　实验与分析

2.1　数据集及预处理

本文选用“UNB ISCX VPN-nonVPN 2016”公开加密流量数据集和“USTC-TK2016”加密流量数据集对本文所提算法进行性能测试，上述两个数据集都是常见的用来测试模型分类效果的公开数据集。“UNB ISCX VPN-nonVPN 2016”数据集中按照服务类型分为 12 类加密流量，多种应用程序流量被包含在每种服务类型流量里。其中 VPN-Chat、VPN-Email、VPN-File、VPN-P2P、VPN-Streaming、VPN-VoIP 属于 VPN 加密流量，Chat、Email、File、P2P、Streaming、VoIP 属于非 VPN 加密流量，本文分别对 VPN 加密流量和非 VPN 加密流量进行六分类。每一类流量的样本数量如表 1 所示。USTC-TFC2016 数据集中包含恶意流量和良性流量，本文选取良性流量中的 BitTorrent、FTP、Facetime、MySQL、Gmail、Outlook、Skype、SMB、Weibo、WorldOfWarcraft 这 10 类流量进行分类，每一类流量的样本数量如表 2 所示。

本文首先按照会话粒度对数据集中的 pcap 格式的流量数据进行分流，通过实验了解到每个会话流中用于构造对应拓扑图的数据包数量超过一定范围时，对分类器的最终分类精度提升有限，并且会延长模型训练时间。最终我

表1 UNB ISCX VPN-nonVPN 2016 数据集中各类流量的样本数量

服务类别	样本数量	服务类别	样本数量
VPN-Chat	3 775	Chat	1 840
VPN-Email	93	Email	170
VPN-File	453	File	260
VPN-P2P	326	P2P	590
VPN-Streaming	311	Streaming	380
VPN-VoIP	3 232	VoIP	5 610

表2 USTC-TFC2016 数据集中各类流量的样本数量

服务类别	样本数量	服务类别	样本数量
BitTorrent	7 483	Skype	5 679
FTP	6 147	SMB	6 240
Gmail	8 552	Facetime	6 000
MySQL	6 942	Weibo	5 033
Outlook	7 524	WorldOfWarcraft	7 878

们将会话流中数据包个数的阈值设置为30个,将超过30个数据包的会话流进行截断处理,只选取会话流的前30个数据包,利用算法提取会话流中的数据包负载长度信息和方向信息形成数组序列,最后将数组序列通过图构建算法生成拓扑图。

2.2 性能评价指标

为了验证本文所提的基于GCN的分类模型的分类性能,本文将结合准确率、精确率、召回率以及F1度量3种性能评价指标对分类模型的分类性能进行全面的评估。上述性能评估指标公式如下:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

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

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{7}$$

其中,TP 为真阳,表示将正样本预测为正;TN 为真阴,表示将负样本预测为负。FP 为假阳,表示将负样本预测为正;FN 为假阴,表示将正样本预测为负。

2.3 攻击实验

1)实验设置

本文使用CNN模型作为攻击实验中的分类模型,采用流量预处理算法将流量数据转换为灰度图像,按照1∶9的比例将流量数据划分为训练样本和测试样本,学习率设置为0.001,迭代次数设置为50次。在攻击方面选择了白盒攻击进行攻击,采用的攻击方法为基于优化的CW攻击,CW攻击在范数下进行评估,实验在流量灰度图的不同扰动范围内对阈值进行测试,最终选取的扰动大小为100/255。对于基于L2范数的CW2攻击,将二分搜索步骤设置为9,最大迭代次数设置为100,初始常数设置为0.001,学习率设置为0.01,置信度设置为0,CW2攻击从测试集中生成1 000个对抗样本。

2)结果分析

本文首先采用CNN分类模型对原始流量数据集进行了分类性能测试,在UNB ISCX VPN-nonVPN 2016数据集上按照流量服务类型分别对VPN和nonVPN流量进行了六分类实验,在USTC-TFC2016数据集上对十类良性流量进行分类。结果如表3所示。

面对CW2攻击,普通CNN模型几乎没有任何防御效果,分类模型无法对CW2攻击后产生的对抗样本正确分类。本文在置信度为0的情况下,对CW2攻击后产生的对抗样本进行了分类测试,我们分别对UNB ISCX VPN-nonVPN 2016数据集的VPN和nonVPN流量以及USTC-TFC2016数据集的10类良性流量进行CW2攻击产生对抗样本,并进行六分类和十分类实验,得到的分类准确率均降到了0。实验结果表明,CW2在CNN分类模型上取得了很好的攻击效果,仅仅是对流量灰度图添加了轻微的扰动就对分类结果造成了极大的影响,足以证明对抗攻击对基于卷积神经网络的网络流量分类算法造成了巨大的威胁。

表3 基于卷积神经网络模型的分类结果

分类实验	Accuracy	Precision	Recall	F1
VPN六分类	0.971 4	0.945 0	0.917 1	0.928 9
nonVPN六分类	0.890 4	0.890 4	0.890 4	0.890 4
USTC十分类	0.996 8	0.996 8	0.996 8	0.996 8

2.4 对抗训练实验

1)实验设置

目前最好的防御对抗攻击的方法之一就是对抗训练,本文首先利用CW2攻击对流量灰度图进行攻击生成对抗样本,采用7∶3混合正常样本和对抗样本得到混合样本集,然后将混合样本放入普通CNN分类模型中进行训练,得到的模型称之为防御模型。防御模型通过对混合样本的学习将有效提高分类准确率,增强CNN模型的鲁棒性。

2)结果分析

对抗训练实验在UNB ISCX VPN-nonVPN 2016数据集及USTC-TFC2016数据集上进行。分类结果如表4所示。由表中数据可知,相较于普通CNN分类模型在对抗样本上0%的分类准确率,通过混合对抗训练得到的防御模型的在对抗样本上分类性能得到有效提高。

2.5 本文所提方法分类实验

1)实验设置

本文采用基于GCN的分类模型,通过使用本文所提

表 4　混合训练实验结果

分类实验	Accuracy	Precision	Recall	F1
VPN 六分类	0.685 8	0.654 4	0.625 6	0.639 7
nonVPN 六分类	0.592 6	0.585 4	0.566 2	0.575 6
USTC 十分类	0.625 4	0.588 5	0.577 6	0.583 0

的构图方式完成图的构造，然后将构造的图打上标签分批次输入模型，初始化权重参数，并对模型参数进行学习，经反向传播更新权重参数，选定模型学习率为 0.01，batch_size 为 32，最大迭代次数为 200 次。

2)结果分析

(1) 本文所提算法实验对比

为了进行对比，本文所提方法分类实验同样是在 UNB ISCX VPN-nonVPN 2016 数据集及 USTC-TFC2016 数据集上进行，最终分类结果如表 5 所示。

表 5　本文所提方法实验对比

分类实验	对比实验	Accuracy	Precision	Recall	F1
VPN 六分类	正常样本分类	0.971 4	0.945 0	0.917 1	0.928 9
	对抗攻击	0	0	0	0
	混合对抗训练	0.685 8	0.654 4	0.625 6	0.639 7
	本文方法	**0.980 5**	**0.980 0**	**0.960 0**	**0.970 0**
nonVPN 六分类	正常样本分类	0.890 4	0.890 4	0.890 4	0.890 4
	对抗攻击	0	0	0	0
	混合对抗训练	0.592 6	0.585 4	0.566 2	0.575 6
	本文方法	**0.944 6**	**0.944 6**	**0.944 2**	**0.944 4**
USTC 十分类	正常样本分类	0.996 8	0.996 8	0.996 8	0.996 8
	对抗攻击	0	0	0	0
	混合对抗训练	0.625 4	0.588 5	0.577 6	0.583 0
	本文方法	**0.998 8**	**1.000 0**	**1.000 0**	**1.000 0**

由表 5 中数据可以看出，共有 4 项对比实验，分别是普通 CNN 分类模型对未受到攻击的正常样本的分类实验、普通 CNN 分类模型对受到 CW2 攻击的对抗样本的分类实验、经过混合对抗训练得到的防御模型对受到 CW2 攻击的对抗样本的分类实验以及本文所提方法进行的分类实验。普通 CNN 分类模型在样本受到攻击后在 Accuracy、Precision、Recall 和 F1-Score 四项性能指标上的数据都降到了 0；混合对抗训练的方法可以有效提高模型的鲁棒性，得到的防御模型在对抗样本上的各项性能评估指标达到了 50%～60%，相比之下，本文所提分类算法在 VPN 六分类中各项性能评估指标达到了 98%，在 nonVPN 六分类中也达到了 94%，而在 USTC-TFC2016 十分类中更是高达 99.88%。因此，本文所提的基于 GCN 的分类算法采用了将流量转换为流量交互拓扑图的流量处理方式，不仅避免了当前典型的流量灰度图像容易受到的对抗攻击的影响，而且在分类性能上相比于当前主流的分类算法有明显提高。

(2) 与主流分类算法对比

为了体现本文所提分类算法相较于当前主流的基于 CNN 分类模型的加密流量分类算法在性能上的优势，本文对比了其他论文中应用到基于 CNN 分类模型的分类算法在相同数据集上的分类效果，对比结果如表 6 所示。表中所对比的算法都是在 UNB ISCX VPN-nonVPN 2016 数据集上进行流量分类实验的，但是分类类型有所不同，如文献[16]只对 VPN 流量中的六种服务类型进行了分类，文献[17]则针对流量中的 12 种服务类型进行分类测试。由于其他算法在论文中只使用了准确率来衡量算法的分类性能，本文也只在准确率上与其他算法相比。

表 6　当前主流分类算法准确率对比

方法	分类模型	分类类型	Accuracy
文献[16]	CNN	VPN 流量六分类	0.929 2
文献[17]	CNN+LSTM	服务类型十二分类	0.91
文献[18]	CNN	nonVPN 流量六分类	0.882
本文方法	GCN	nonVPN 流量六分类	**0.944 6**
		VPN 流量六分类	**0.980 5**

由表 6 可知，相比于文献[16]中所提的分类模型，本文所提分类算法在准确率上有 5.13% 的提升，而相较于文献[18]，本文所提分类算法则提升了 6.26%。由此可以得出结论，与其他论文所提的分类算法相比，本文所提分类算法具有更优的分类性能。

3　结　　论

本文首先分析了对抗攻击对当前主流的基于 CNN 分类模型的网络流量分类方法带来的影响，并通过实验对该

问题进行了验证。针对该问题,本文提出了一种基于GCN的网络流量分类算法,所提方法采用会话流中数据包负载长度、方向、包序列以及簇特征进行流量交互拓扑图的构建,然后将拓扑图输入GCN模型进行特征学习并分类。以节点拓扑图代替灰度图作为模型的输入,避免了针对灰度图像的对抗攻击的影响。实验结果表明所提方法能够在流量数据集中实现很好的分类效果,在准确率、精确率、召回率以及F1-Score上的表现优于CNN模型。本文方法仍存在的不足之处在于未充分利用流量交互过程中的其它特征,未来考虑在流量交互拓扑图的构建中加入这些特征,进一步提升分类性能。

参考文献

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

[2] SALMAN O, ELHAJJ I H, KAYSSI A, et al. Data representation for CNN based internet traffic classification: A comparative study[J]. Multimedia Tools and Applications, 2021, 80(11): 16951-16977.

[3] 林鹏,翟江涛,许历隆,等.一种面向类不平衡加密流量的端到端分类模型[J].电子测量技术,2021,44(20):142-149.

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

[5] YONG W, HUIYI Z, HAO F, et al. Network traffic classification method basing on CNN[J]. Journal on Communications, 2018, 39(1): 14-23.

[6] HE Y, LI W. Image-based encrypted traffic classification with convolution neural networks[C]. 2020 IEEE Fifth International Conference on Data Science in Cyberspace(DSC), IEEE, 2020: 271-278.

[7] SZEGEDY C, ZAREMBA W, SUTSKEVER I, et al. Intriguing properties of neural networks[J]. ArXiv Preprint,2013, ArXiv:1312.6199.

[8] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples[J]. ArXiv Preprint,2014,ArXiv:1412.6572.

[9] CARLINI N, WAGNER D. Towards evaluating the robustness of neural networks [C]. 2017 Ieee Symposium on Security and Privacy (sp), IEEE, 2017: 39-57.

[10] MADRY A, MAKELOV A, SCHMIDT L, et al. Towards deep learning models resistant to adversarial attacks [J]. ArXiv Preprint, 2017, ArXiv: 1706.060837.

[11] ZHENG T, CHEN C, REN K. Distributionally adversarial attack [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 2253-2260.

[12] 胡永进,郭渊博,马骏,等.基于对抗样本的网络欺骗流量生成方法[J].通信学报,2020,41(9):59-70.

[13] 羊洋,陈伟,张丹懿,等.对抗攻击威胁基于卷积神经网络的网络流量分类[J].计算机科学,2021,48(7):55-61.

[14] JI X, MENG Q. Traffic classification based on graph convolutional network[C]. 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA), IEEE, 2020: 596-601.

[15] BUSCH J, KOCHETUROV A, TRESP V, et al. NF-GNN: Network flow graph neural networks for malware detection and classification [C]. 33rd International Conference on Scientific and Statistical Database Management, 2021: 121-132.

[16] 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.

[17] ZOU Z, GE J, ZHENG H, et al. Encrypted traffic classification with a convolutional long short-term memory neural network [C]. 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), IEEE, 2018: 329-334.

[18] SHAPIRA T, SHAVITT Y. Flowpic: Encrypted internet traffic classification is as easy as image recognition [C]. IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, 2019: 680-687.

作者简介

王勤凡,硕士研究生,主要从事多媒体与信息安全方面的研究。

E-mail:20201249328@nuist.edu.cn

翟江涛(通信作者),博士,副教授,主要从事多媒体与信息安全方面的研究。

E-mail: jiangtaozhai@nuist.edu.cn

陈伟,硕士研究生,主要从事多媒体与信息安全方面的研究。

E-mail: 1065155648@qq.com

孙浩翔,硕士研究生,主要从事多媒体与信息安全方面的研究。

E-mail: sun877903408@163.com