Modeling and Simulation 建模与仿真, 2024, 13(3), 2347-2358
Published Online May 2024 in Hans. https://www.hanspub.org/journal/mos
https://doi.org/10.12677/mos.2024.133215

Hans汉斯

基于多阶段神经网络的加密流量分类

刘　亚[1]，王　晓[1]，曲　博[2*]

[1]上海理工大学光电信息与计算机工程学院，上海
[2]广东科技学院计算机学院，广东　东莞

收稿日期：2024年4月20日；录用日期：2024年5月14日；发布日期：2024年5月21日

摘　要

随着加密技术的普及，准确分类加密流量对于识别匿名网络应用程序和防止网络犯罪至关重要。现有方法局限于专家经验或局部数据包信息，无法理解数据包之间的依赖关系。为解决这个问题，提出了多阶段神经网络流量分类器(MSNTC)，使用卷积神经网络(CNN)将会话图像拆分为数据包序列，长短时记忆网络(LSTM)获取流量上下文嵌入，自我注意力机制获得多通道特征图，再利用多尺度卷积神经网络聚合全局信息。在ISCX-VPN和ISCX-Tor数据集上对MSNTC模型进行评估，并与其他方法对比。实验结果表明，MSNTC模型在网络流量分类任务中展现出更好性能，验证了其优越性和通用性。

关键词

加密流量分类，注意力机制，卷积神经网络，长短时记忆网络

Encrypted Traffic Classification Based on Multi-Stage Neural Networks

Ya Liu[1], Xiao Wang[1], Bo Qu[2*]

[1]School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai
[2]School of Computer Science, Guangdong University of Science and Technology, Dongwan Guangdong

Received: Apr. 20th, 2024; accepted: May. 14th, 2024; published: May. 21st, 2024

Abstract

With the proliferation of encryption technology, the accurate classification of encrypted traffic is of vital importance for the identification of anonymous network applications and the prevention

*通讯作者。

文章引用: 刘亚, 王晓, 曲博. 基于多阶段神经网络的加密流量分类[J]. 建模与仿真, 2024, 13(3): 2347-2358.
DOI: 10.12677/mos.2024.133215

of cybercrime. Existing methods are limited to expert experience or partial packet information, thereby failing to comprehend the interdependencies between packets. To address this issue, a novel multi-stage neural network traffic classifier (MSNTC) is proposed, wherein convolutional neural networks (CNN) are employed to decompose session images into packet sequences, long short-term memory networks (LSTM) capture traffic context embeddings, self-attention mechanisms obtain multi-channel feature maps, and multi-scale convolutional neural networks are utilized to aggregate global information. The MSNTC model is evaluated on the ISCX-VPN and ISCX-Tor datasets and compared with other deep learning methods. Experimental results demonstrate the superior performance of the MSNTC model in network traffic classification tasks, thereby corroborating its superiority and universality.

Keywords

Encrypted Traffic Classification, Self-Attention Mechanism, Convolutional Neural Network, Long-Short Term Memory Network

1. 引言

随着互联网和信息技术的快速发展，网络已经渗透到各个方面。然而，网络安全面临多种风险和威胁，如黑客攻击、数据泄露和网络间谍活动[1]。为了保护数据的机密性和完整性，在网络数据传输中采用流量加密技术成为一种重要手段。然而，加密技术也被恶意分子滥用，成为网络安全的新威胁[2]。为了及时识别网络安全风险并采取措施，网络管理员需要准确分类流量，尤其是判断是否存在恶意加密流量[3]。为了应对不断变化的网络安全威胁，需要持续研究新技术，例如基于机器学习和深度学习的加密流量分类方法，以及面向加密流量的安全防护方案[4]。这些技术的研究和应用有助于提升网络安全水平，确保网络正常运行。

流量分类技术现在已经有了显著的进展。最初的方法是使用端口号进行分类，通过检测网络数据包中的源和目标端口号来识别不同种类的网络流量。这种方法简单易用，但是，越来越多的应用程序使用动态或非标准端口，使得基于端口的方法变得越来越不可靠[5]。为了提高准确性，深度数据包检查方法则是尝试分析网络数据包的内容，以识别其类型。它比基于端口的方法更精确，但只适用于未加密的流量，并且计算开销较大[6]。相比于解析流量数据的有效载荷，机器学习方法通过捕捉流量的行为特征或统计特征来完成分类任务，例如数据包传输频率、包的大小以及流的持续时间等[7]。因此，它可以应对网络变化和应用更新，并且不受加密算法的影响，能够有效地对加密流量进行分类。这些方法通常依赖于网络流的统计特征，如 AppScanner [8]、BIND [9]等方法。然而，这些方法的性能取决于人工设计的特征，因此限制了它们的泛化能力。

深度学习技术具有自动提取特征的能力，避免了繁琐的特征工程步骤，并在大规模数据处理方面表现出色，可更准确、更快速地识别流量类型[10]。与传统方法相比，深度学习模型能够高效自动地提取原始流量数据的特征，并完成端到端的加密流量分类任务。Wang [11] [12]等人使用一维卷积神经网络(1D-CNN)和二维卷积神经网络(2D-CNN)将原始流量数据转换为灰度图像以实现流量识别，并成功地实现了加密流量的端到端分类。然而，该方法存在缺陷，仅取网络流前 768 个字节将其转换为图像，未考虑到以数据包为单位进行转换为图像的方法。因此，在卷积操作中会导致不同时间段的数据包图像信息错

乱，从而影响分类准确性。Kumano [13]等人在其研究中提出，可以通过保持分类准确性的前提下，减少所需的数据包数量，只需使用 10 个数据包即可保证准确性。Deep Packet [14]结合堆叠自动编码器和卷积神经网络作为分类模型，使用网络中的所有层进行分类。但网络和传输层的协议字段主要设计用于网络传输而非应用程序识别，应用层以下协议字段几乎不包含有用信息。因此，包含太多无关信息会增加数据复杂性，并可能导致模型过度依赖不重要的信息，从而降低其对细粒度流量分类的差异能力。TSCRNN [15]利用 CNN 从网络流量数据中提取空间特征，并使用 LSTM 从 CNN 提取的特征中提取时间特性以捕捉网络流的时空特征。然而，该方法无法考虑到网络流中的全局特征，这将导致其分类性能受到限制。

基于上述研究，提出了多阶段神经网络流量分类方法。在预处理阶段，将原始流量分割为双向流，即会话，仅对每个会话中数据包的有效载荷以字节为单位进行读取并转换为图像，最后，将数据包图像序列进行合并，构建出会话图像。以 1D-CNN 为基准模型，这种预处理方法比 Wang 的预处理方法性能提高了 7.5%。在模型设计部分，结合上下文嵌入，自注意力机制和卷积神经网络完成加密流量分类任务。具体来说，在初始嵌入中，会话图像会被划分为固定大小的图像序列，该序列中的每个图像表示会话中的数据包。用 2D-CNN 对网络会话中的数据包图像进行线性投影，接着利用 LSTM 根据数据包的时间关系获取上下文嵌入，使用多头自注意力机制来捕获出会话中数据包之间的依赖关系，提取网络会话的多通道特征图，最后用卷积神经网络对特征图进行进一步的特征提取，挖掘出数据包之间更深层的关系。

2. 多阶段流量分类方法

MSNTC 的架构图如图 1 所示。首先对原始流量按照会话流进行分割，将会话流中的每个数据包转为灰度图后进行拼接，得到会话流的灰度图。接着对于会话流图像，使用 2D-CNN 将会话拆分为数据包序列并进行线性嵌入，通过 LSTM 来获取会话流中每个数据包的时序关系，将 LSTM 的每个隐层的输入作为会话中数据包序列的最终嵌入。将会话流嵌入馈送给多头注意力层来获取每个会话的特征图，对该特征图使用 CNN 来对会话流的最终特征进行提取，最后使用全连接层来实现加密流量的识别。

2.1. 数据预处理

在预处理阶段，只关注应用层即数据包的有效载荷，而不考虑其他层，因为其他层主要包含控制信息而非有效信息。数据预处理过程如图 2 所示。网络流中的所有数据包具有相同的源和目的 IP 地址、协议、源和目的端口号，形成一个五元组。网络会话是双向流，包括具有可互换的源和目标 IP/端口对的两个流，因此它包含比单向流更多的交互信息。在本研究中，网络会话被用作模型的输入，根据五元组将原始流量划分为不同的会话，会话定义为：

$$S_e = \left[P_1, P_2, \cdots, P_{n-1}, P_n\right]$$

其中，P_i 表示会话的第 i 个位置中的数据包，n 为每个会话中数据包的个数。不同的数据包具有不同的长度，而 MSNTC 可接受的输入大小是固定的。需要对数据包的格式进行统一处理，由于网络最大传输单元(MTU)最大为 1500 字节，为方便处理，将数据包的大小固定为 1600 字节，这完全可以包含每个数据包的全部信息。对于长度不够 1600 字节的数据包，在其末尾填充 0 × 00 至 1600 字节。实验中对该网络流取前 t 个数据包，若会话中的数据包数量不足 t，则用 0 × 00 填充至 t 个数据包。以字节为单位读取会话中的每个数据包，并将其转为 40×40 的灰度图，生成数据包图像序列，对数据包图像序列进行合并操作得到最终的网络会话流图像 S 来构建网络流图像数据集。

2.2. 嵌入层

在经过数据预处理后，使用 2D-CNN 将会话图像分解为多个数据包图像序列并对每个数据包完成初

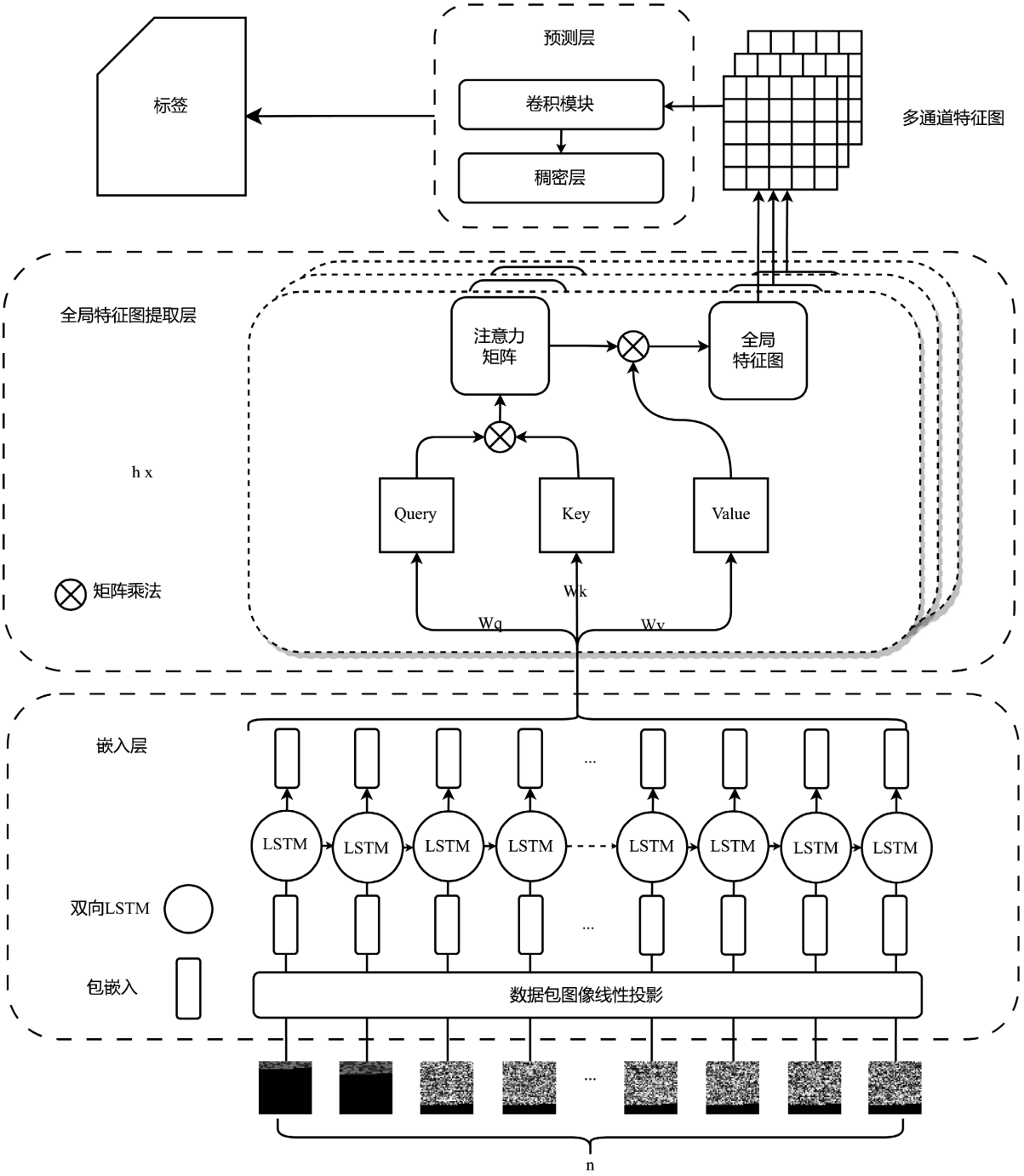

Figure 1. Overall of MSNTC
图 1. MSNTC 的整体架构

始嵌入。首先，将会话流图像分割成多个分组图像，将这些分组图像重新排列成序列。然后，对每个数据包图像执行线性嵌入操作。具体而言，通过在会话流图像上应用步长为 40、卷积核大小为 40、通道数量为嵌入维度的二维卷积运算，可以获得每个数据包的线性嵌入。

对于网络会话图像 S，这个会话图像被划分成了 t 个数据包图像 D_i，其中每个数据包是一个 40×40 的二维矩阵。对于初始嵌入 $E_i \in \mathbb{R}^{1\times1\times d_{embed}}$，计算公式如下：

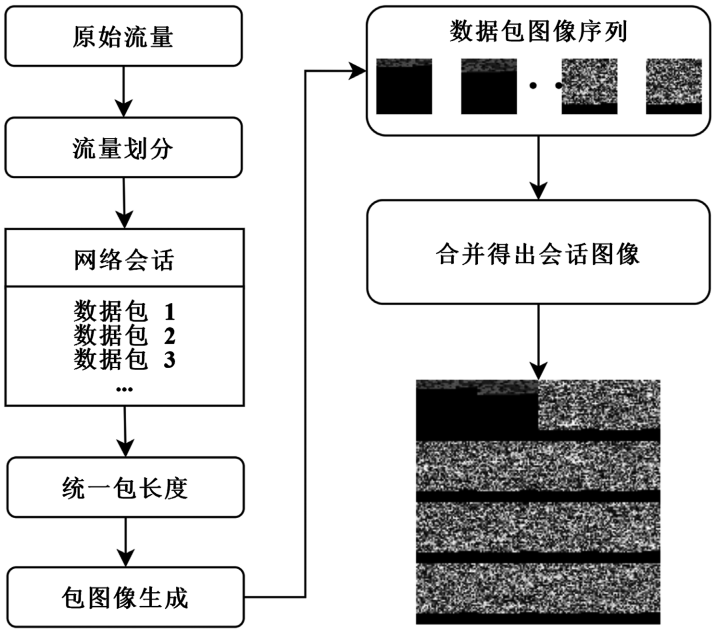

Figure 2. Data preprocessing of MSNTC
图 2. MSNTC 的数据预处理

$$E_i = W_{conv} * D_i + b_{conv}$$

其中，$W_{conv} \in \mathbb{R}^{40\times40\times c_{in}\times d_{embed}}$ 表示卷积核的权重矩阵，c_{in} 是输入数据包的通道数 $b_{conv} \in \mathbb{R}^{d_{embed}}$ 表示偏置向量。

对于第 i 个数据包的线性嵌入 E_i，使用一个双向 LSTM 层来获取其上下文信息，得到输出表示 $H_i \in \mathbb{R}^{1\times1\times2d_{lstm}}$。这个双向 LSTM 层由两个单向 LSTM 层组成，分别沿着时间方向和反方向处理输入序列，并在最后把它们连接起来。

$$\overrightarrow{h_i} = LSTM\left(\overrightarrow{h_{i-1}}, E_i\right)$$

$$\overleftarrow{h_i} = LSTM\left(\overleftarrow{h_{i+1}}, E_i\right)$$

$$H_i = \left[\overrightarrow{h_i}, \overleftarrow{h_i}\right]$$

其中，$\overrightarrow{h_i}$ 和 $\overleftarrow{h_i}$ 分别表示时刻 i 时从正向和反向处理输入序列所得到的隐状态向量，$[\cdot,\cdot]$表示将两个向量连接起来。最终，将每个数据包的输出表示 H_i 组装成会话的输出表示 $X \in \mathbb{R}^{t\times d}$，其中 t 是数据包的数量，d 为 $2d_{lstm}$，即数据包的嵌入维度。这种嵌入表示可以捕捉到数据包的上下文信息，同时，LSTM 的多层结构和双向结构可以提高数据包嵌入的准确性和丰富性。

2.3. 全局特征图提取层

网络流量的顺序性和应用层数据包之间的内在依赖性是流分类任务中不可忽视的关键因素。受 Transformer 的启发[16]，引入自注意力机制，该机制可以捕获每个数据包之间的交互信息，从而提取会话流的全局特征。

使用自注意力机制将所有数据包的信息集成在一起，并将它们以不同的权重融合在一起。如图 3 所示，不同类型的线表示不同数据包的注意力权重，通过自注意力机制学习数据包之间的交互关系，对数据包进行加权得到最终的全局特征图(Global Feature Map)。将网络会话的数据包上下文向量沿时间轴方向拼接成

一个矩阵作为整个会话的嵌入。使用自注意力机制来计算全局特征图 $M_{t\times m}$，其中 m 为全局特征的维度。

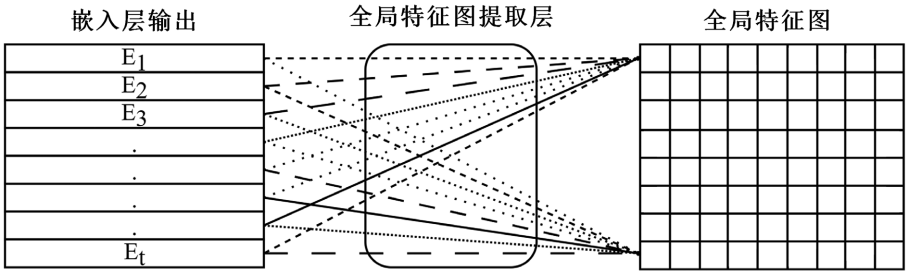

Figure 3. Global feature map extraction mechanism
图 3. 全局特征图提取机制

为了得到自注意力权重矩阵 $A\in\mathbb{R}^{t\times t}$，首先对输入矩阵 X 进行线性变换，得到查询矩阵 Q，键矩阵 K 和值矩阵 V：

$$Q=XW_q,K=XW_k,V=XW_v$$

其中，$W_q,W_k,W_v\in\mathbb{R}^{d\times m}$ 是需要学习的权重矩阵。然后，通过计算自注意力权重矩阵 A 来获得会话的全局特征：

$$A=softmax\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right)$$

其中，d_k 是查询矩阵 Q 和键矩阵 K 的维度，$\sqrt{d_k}$ 是为了缓解内积的数量级带来的影响。最后，使用注意力权重 A 对值矩阵 V 进行加权求和，得到全局特征图 M：

$$M=AV$$

在自注意力机制的基础上，进一步使用多头自注意力机制来获得更好的线性表示，这意味着全局特征图提取层可以获得多个特征图，每个注意力头可以强调不同的上下文子空间中的特征。在多头自注意力机制中，使用 h 个不同的投影矩阵来将查询、键和值矩阵分别投影到 h 个不同的子空间中，可以形式化为：

$$Q_i=XW_i^q,K_i=XW_i^k,V_i=XW_i^v$$

其中，$W_i^q\in\mathbb{R}^{d\times m},W_i^k\in\mathbb{R}^{d\times m},W_i^v\in\mathbb{R}^{d\times m}$。然后，分别对每个子空间进行自注意力计算，得到 h 个加权值矩阵 $Z_1,\cdots,Z_h$：

$$Z_i=softmax\left(\frac{Q_i\left(K_i\right)^{\mathrm{T}}}{\sqrt{d_k/h}}\right)V_i$$

通过对输入序列进行不同权重的加权平均，得到一组子空间特征矩阵 Z_i，这些子空间特征矩阵反映了输入序列在不同方面的关注度。由于注意力机制的引入，每个子空间对应一个不同的关注度，因此它们可以捕捉到数据包序列中不同方面的信息。将子空间特征矩阵拼接成一个三维张量 $Z\in\mathbb{R}^{t\times m\times h}$，可将 Z 看作一个多通道的会话特征图，h 表示注意力头数，即通道数。

2.4. 预测层

为了进一步提取有用的信息，使用卷积神经网络对 Z 进行特征提取，获得更高层次的抽象特征表示。这个过程充分利用了数据包序列中的局部和全局信息，能够提取出多层次、多尺度的特征。

首先，采用卷积神经网络对特征图进行特征提取。具体而言，对特征图进行卷积操作，输出新的特征图 $F \in \mathbb{R}^{t' \times m' \times c}$，其中 t' 和 m' 表示特征图的高和宽，c 表示输出通道数，即卷积核数量：

$$F = \sigma\left(W * Z + b\right)$$

其中*表示卷积操作，$W \in \mathbb{R}^{k \times k \times h \times c}$ 表示卷积核，k 表示卷积核的大小，σ 表示激活函数，通常使用 ReLU 函数。$b \in \mathbb{R}^c$ 表示偏置项。

经过多个卷积操作后，使用全局平均池化对特征图进行降维并提取出其中最显著的特征，从而简化模型的计算和参数量，输出尺寸为 c 的特征向量 $v \in \mathbb{R}^c$，表示对特征图每个通道上所有值所求得的平均值：

$$v_i = \frac{1}{t'm'} \sum_{j,k} F_{(j,k),i}$$

其中，i 表示特征向量 v 的第 i 个元素，也表示卷积操作后输出的第 i 个特征图，$\left(j,k\right)$ 表示特征图 F 的行列坐标。最后，将特征向量 v 输入到全连接层中进行分类。如果存在 r 个分类标签，则输出层的大小为 r。具体地，最终的分类结果 $\hat{y} \in \mathbb{R}^r$ 是一个长度为 r 的向量，每个元素表示会话对应分类标签的概率：

$$\hat{y} = softmax\left(W'v + b'\right)$$

其中 $W' \in \mathbb{R}^{r \times c}$，$b' \in \mathbb{R}^r$ 为可学习的参数。

3. 实验分析

3.1. 数据集

为了验证 MSNTC 模型的泛化性能，在本研究中采用了广泛应用的大型加密流量公开数据集，包括 ISCX VPN 2016 [17]和 ISCX Tor 2016 [18]。表 1 和表 2 分别展示了这两个数据集的详细信息。通过使用这些常用的加密流量数据集，我们可以充分验证 MSNTC 模型在不同加密场景下的性能和泛化能力。

Table 1. The description of ISCX VPN 2016 dataset
表 1. ISCX VPN 2016 数据集具体信息

类型	标签
NonVPN	Chat, Email, FTP, Streaming, VOIP
VPN	VPN-Chat, VPN-Email, VPN-FTP, VPN-Streaming, VPN-VOIP, VPN-P2P

Table 2. The description of ISCX Tor 2016 dataset
表 2. ISCX Tor 2016 数据集具体信息

类型	标签
NonTor	Chat, Email, FTP, Video, Browsing, Audio, VOIP, P2P
Tor	Tor-Chat, Tor-Email, Tor-FTP, Tor-Video, Tor-Browsing, Tor-Audio, Tor-VOIP, Tor-P2P

3.2. 评估指标

在多分类任务中，为了客观、准确地评估 MSNTC 的性能，选择了四个常用的测量指标，分别是准确率(Accuracy)、宏 F1 值(Macro-F1)、宏召回率(Macro-Recall)和宏精确率(Macro-Precision) [19]。

3.3. 比较方法

在本实验中，主要评估和对比了 MSNTC 和多种先进的流分类方法，其中包括 AppScanner [8]、BIND

[9]、Deep Packet [14]、TSCRNN [15]和 FlowPrint [20]。这些方法采用了不同的特征提取和分类技术。

Table 3. Details of experimental parameters
表 3. 实验参数设置

参数	设置
会话数据包数量	16
数据包有效载荷字节	1600
数据包图像尺寸	(40, 40)
会话图像尺寸	(160, 160)
初始嵌入维度	256
上下文嵌入维度	256
注意力机制层数(通道数)	8
卷积核大小	3
卷积核个数	512
激活函数	ReLU
密集层神经元数量	1024
损失函数	交叉熵
优化器	Adam
学习率	0.0005
预热学习率	0.1
批次大小	16
训练轮次	30
训练集、验证集、测试集数量之比	7:1:2

3.4. 实验分析

首先，探究数据包个数对模型性能的影响，对 ISCX Tor 数据集中的每个类别各选择 600 个网络会话进行分析，选择网络会话的前 1、4、9、16、25、36 和 49 个数据包，并观察数据包数量对准确率的影响。图 4 展示了 MSNTC 模型在不同数据包取样数量下的准确率变化趋势。实验结果显示，在选择了 9 个数据包之后，准确率变化趋势不再明显。而当选择了 16 个数据包时，准确率达到最高点。本文将数据包个数设置为 16 个数据包，其他实验参数的配置见表 3。

为了验证本文提出的预处理方法的有效性，需要与传统预处理方法(这里采用网络流的前 1600 个字节)进行比较。由于传统方法不适用于输入到 MSNTC 模型中，本次实验选择了通用的 1D-CNN 模型对预处理方法进行性能比较。从表 4 中可以观察到，使用本文提出的预处理方法在加密流量分类任务中表现出更好的性能。这表明本文的预处理方法在捕捉加密流量的关键特征方面具有优势。因为以数据包为单位进行预处理，确保了不同时间点的数据包在预处理后保持整齐有序，并且本文将完整的数据包有效载荷保留下来。而传统方法以会话为单位简单地提取前 n 个字节会导致部分信息混乱或丢失，使得模型难以学习。

图 5 展示了不同模型在两种类型的测试集中的准确率对比，可以发现 MSNTC 模型在 ISCX VPN 数据集和 ISCX Tor 数据集上的准确率表现相对于其他模型是最好的。MSNTC 模型在 ISCX Tor 数据集上的

准确率提升幅度显著，较 AppScanner 模型提高了约 30 个百分点。在 ISCX VPN 数据集上，其准确率提高了约 25 个百分点。这说明深度学习的方法更加注重特征提取和拟合，能够更好地反映现实问题的本质特征，从而明显高于机器学习方法的分类效果。

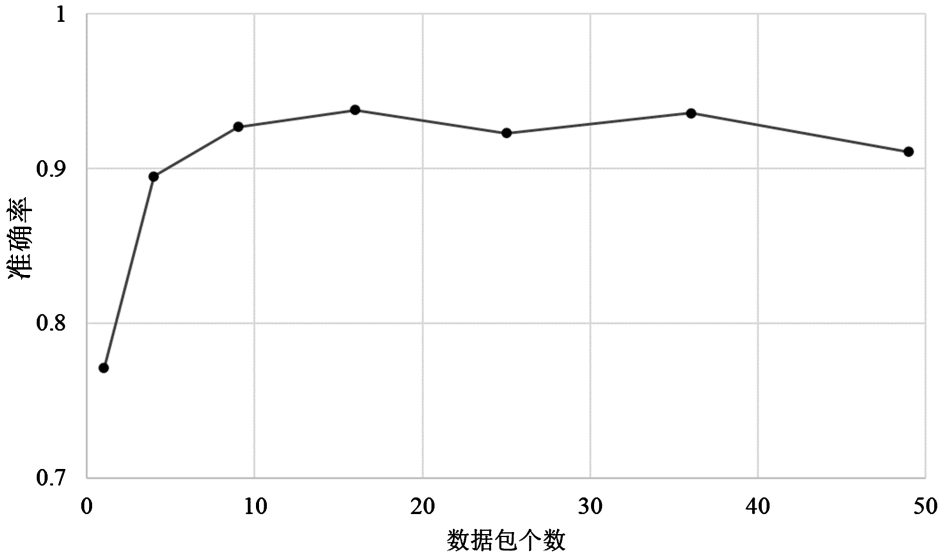

Figure 4. Comparison of accuracy of different packet counts in the ISCX Tor
图 4. 在 ISCX Tor 数据集种不同数据包个数准确率对比

Table 4. Performance comparison of preprocessing methods
表 4. 预处理方法性能比较

方法比较	准确率
本文预处理方法 + 1D-CNN	0.921
传统预处理方法 + 1D-CNN	0.857

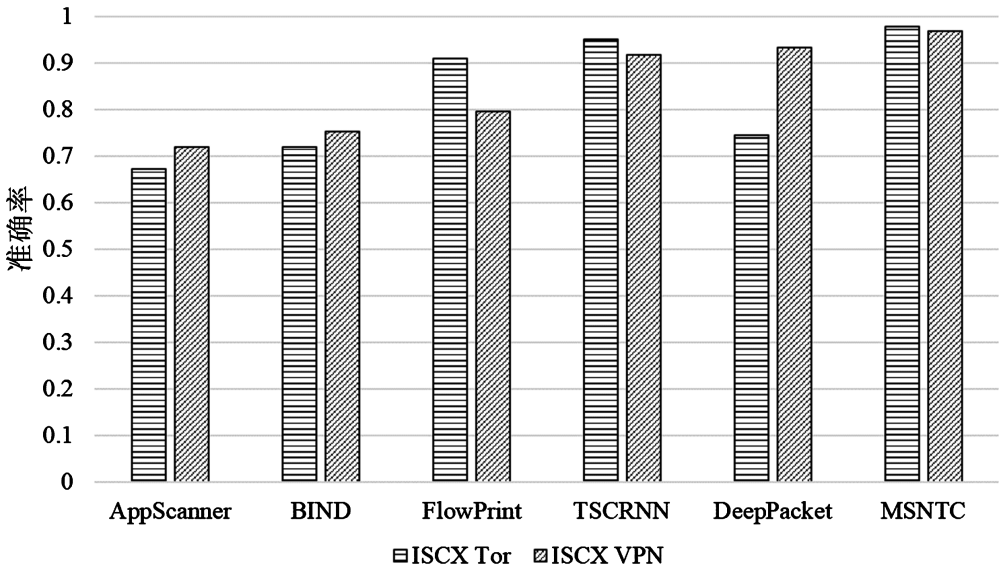

Figure 5. Accuracy of different models in Tor and VPN data sets
图 5. 不同模型在 Tor 和 VPN 数据集的准确率

相较于深度学习方法，MSNTC 模型在 ISCX Tor 数据集上的精确率、召回率和 F1 值分别比 TSCRNN 提高了 3.4%、3.6%和 3.5%；比 Deep Packet 提高了 30.0%、32.7%和 31.3%。在 ISCX VPN 数据集中，相比 TSCRNN，MSNTC 模型的精确率、召回率和 F1 值分别提高了 4.2%、6.4%和 5.3%；相比 Deep Packet，该模型提高了 3.0%、5.8%和 4.6%。这表明 MSNTC 模型不仅在加密流量检测方面表现优秀，而且相较于传统深度学习方法精度更高。由此可见，引入注意力机制的 MSNTC 模型具有更强的网络数据包特征提取能力，可以捕获数据包的交互关系，提取出更稳定的全局特征，表现结果比循环神经网络更佳。

为了全面评估 MSNTC 模型的效果，表 5 展示了对不同模型进行评估时使用的精确率(PR)、召回率(RC)和 F1 值三个指标。结果显示，MSNTC 模型在所有指标上均优于其他模型，同时也验证了 MSNTC 模型在不同加密场景下(包括 VPN 和 Tor)具有良好的泛化能力。

Table 5. Evaluation indicators of models in the ISCX Tor dataset
表 5. 模型在 ISCX Tor 数据集的评估指标

模型	ISCX Tor 数据集			ISCXVPN 数据集		
	PR	RC	F1	PR	RC	F1
AppScanner	0.376	0.442	0.391	0.739	0.722	0.719
BIND	0.460	0.452	0.451	0.758	0.748	0.742
FlowPrint	0.382	0.366	0.365	0.804	0.781	0.782
TSCRNN	0.949	0.948	0.948	0.927	0.926	0.926
Deep Packet	0.755	0.740	0.747	0.937	0.930	0.932
MSNTC	0.980	0.979	0.978	0.965	0.985	0.975

表 6 展示了不同深度学习模型的参数个数。与 Deep Packet 模型相比，MSNTC 模型的参数量显著减少，但略高于 TSCRNN 模型。这是因为 MSNTC 模型采用嵌入层进行了初步的特征提取和维度压缩，从而降低了全局特征图提取层中自注意力机制的权重计算量，并在卷积层使用了全局池化来减少模型的参数量。通过在保持准确率的同时优化参数量，MSNTC 模型在准确性和模型复杂度之间取得了平衡。

Table 6. Parameters in deep learning model
表 6. 深度学习模型参数

模型	参数量(百万)
Deep Packet	24.72
TSCRNN	2.90
MSNTC	4.69

4. 结束语

MSNTC 与其他端到端模型相比，引入了自注意力机制来提取网络流的全局信息，同时使用 LSTM 对数据包进行嵌入来弥补注意力机制无法考虑数据包时间特性的缺点。使用多种深度特征学习和自注意力机制的这种结构设计使得 MSNTC 模型能够更好地捕捉数据包之间的关系和时序信息，提取更稳定、更全面的特征表示，从而提高分类的准确性。该方法在不同的数据集上均取得最好的分类结果，能够有效地识别加密流量中的各种类型。相比于传统的基于规则的方法和机器学习的方法，MSNTC 方法具有非常显著的优势，对于其他深度学习的方法，MSNTC 展现出了更高的准确率和更强的鲁棒性。未来的

研究方向有多个方面。首先，进一步探究如何提高 MSNTC 方法分类的速度和效率，以满足实时监测和防御的需求。其次，可以考虑增加网络流的统计特征，以提高 MSNTC 方法的识别能力。最后，对 MSNTC 模型进行可解释性研究，研究如何解释和理解 MSNTC 模型的分类结果，以便更好地理解和信任其分类。

基金项目

国家自然科学基金资助项目(62002184)。

参考文献

[1] Xiang, J., Fulton, N. and Chong, S. (2021) Relational Analysis of Sensor Attacks on Cyber-Physical Systems. 2021 *IEEE* 34*th Computer Security Foundations Symposium* (*CSF*), Dubrovnik, 21-25 June 2021, 1-16. https://doi.org/10.1109/CSF51468.2021.00035

[2] Rezaei, S. and Liu, X. (2019) Deep Learning for Encrypted Traffic Classification: An Overview. *IEEE Communications Magazine*, **57**, 76-81. https://doi.org/10.1109/MCOM.2019.1800819

[3] Chen, L., Gao, S., Liu, B., Lu, Z. and Jiang, Z. (2020) THS-IDPC: A Three-Stage Hierarchical Sampling Method Based on Improved Density Peaks Clustering Algorithm for Encrypted Malicious Traffic Detection. *The Journal of Supercomputing*, **76**, 7489-7518. https://doi.org/10.1007/s11227-020-03372-1

[4] Anderson, B. and McGrew, D. (2016) Identifying Encrypted Malware Traffic with Contextual Flow Data. *Proceedings of the* 2016 *ACM Workshop on Artificial Intelligence and Security*, Vienna, 28 October 2016, 35-46. https://doi.org/10.1145/2996758.2996768

[5] Erman, J., Mahanti, A., Arlitt, M.F., *et al*. (2007) Identifying and Discriminating between Web and Peer-to-Peer Traffic in the Network Core. *Proceedings of the* 16*th international conference on World Wide Web*, Banff, 8-12 May 2007, 883-892. https://doi.org/10.1145/1242572.1242692

[6] 陈子涵, 程光, 徐子恒, 等. 互联网加密流量检测、分类与识别研究综述[J]. 计算机学报, 2023, 46(5): 1060-1085.

[7] Fan, Z. and Liu, R. (2017) Investigation of Machine Learning Based Network Traffic Classification. 2017 *International Symposium on Wireless Communication Systems* (*ISWCS*), Bologna, 28-31 August 2017, 1-6. https://doi.org/10.1109/ISWCS.2017.8108090

[8] Taylor, V.F., Spolaor, R., Conti, M. and Martinovic, I. (2018) Robust Smartphone App Identification via Encrypted Network Traffic Analysis. *IEEE Transactions on Information Forensics and Security*, **13**, 63-78. https://doi.org/10.1109/TIFS.2017.2737970

[9] Al-Naami, K., Chandra, S., Mustafa, A., *et al*. (2016) Adaptive Encrypted Traffic Fingerprinting with Bi-Directional Dependence. *Proceedings of the* 32*nd Annual Conference on Computer Security Applications*, Los Angeles, 5-8 December 2016, 177-188. https://doi.org/10.1145/2991079.2991123

[10] 冷涛. 基于深度学习的加密流量分类研究综述[J]. 计算机与现代化, 2021(8): 112-120.

[11] Wei, W., Ming, Z., Zeng, X., *et al*. (2017) Malware Traffic Classification Using Convolutional Neural Network for Representation Learning. 2017 *International Conference on Information Networking* (*ICOIN*), Da Nang, 11-13 January 2017, 712-717.

[12] Wang, W., Zhu, M., Wang, J., Zeng, X., *et al*. (2017) End-to-End Encrypted Traffic Classification with One-Dimensional Convolution Neural Networks. 2017 *IEEE International Conference on Intelligence and Security Informatics* (*ISI*), Beijing, 22-24 July 2017, 43-48.

[13] Kumano, Y., Ata, S., Nakamura, N., *et al*. (2014) Towards Real-Time Processing for Application Identification of Encrypted Traffic. 2014 *International Conference on Computing, Networking and Communications* (*ICNC*), Honolulu, 3-6 February 2014, 136-140. https://doi.org/10.1109/ICCNC.2014.6785319

[14] Lotfollahi, M., Siavoshani, J.M., Zade, R.S.H., *et al*. (2020) Deep Packet: A Novel Approach for Encrypted Traffic Classification Using Deep Learning. *Soft Computing*, **24**, 1999-2012. https://doi.org/10.1007/s00500-019-04030-2

[15] Lin, K., Xu, X. and Gao, H. (2021) TSCRNN: A Novel Classification Scheme of Encrypted Traffic Based on Flow Spatiotemporal Features for Efficient Management of IIoT. *Computer Networks*, **190**, Article 107974. https://doi.org/10.1016/j.comnet.2021.107974

[16] Vaswani, A., Shazeer, N., Parmar, N., *et al*. (2017) Attention Is All You Need. *Proceedings of the* 31*st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.

[17] Lashkari, A.H., Draper-Gil, G., Mamun, M.S.I., *et al*. (2016) Characterization of Encrypted and VPN Traffic Using Time-Related Features. *Proceedings of the* 2*nd International Conference on Information Systems Security and Privacy ICISSP*, Rome, 19-21 February 2016, 407-414. https://doi.org/10.5220/0005740704070414

[18] Lashkari, A.H., Gil, G.D., Mamun, M.S.I., *et al*. (2017) Characterization of Tor Traffic Using Time Based Features. *Proceedings of the* 3*rd International Conference on Information System Security and Privacy*, Porto, 19-21 February 2017, 253-262.

[19] Menzies, S., Greenwald, J. and Frank, A. (2007) An Analysis of Evaluation Metrics for Machine Learning Based Software Fault Prediction Models. *Journal of Systems and Software*, **80**, 1910-1923.

[20] Van Ede, T., Bortolameotti, R., Continella, A., *et al*. (2020) FlowPrint: Semi-Supervised Mobile-App Fingerprinting on Encrypted Network Traffic. *Proceedings of the* 27*th Annual Network and Distributed System Security Symposium*, San Diego, 23-26 February 2020, 1-18. https://doi.org/10.14722/ndss.2020.24412