研究与开发

融合Transformer和Inception的网络入侵检测研究

张万琪，王家兴，宋振峰
（中国人民公安大学信息网络安全学院，北京 100038）

摘　要：针对当前入侵检测模型流量特征提取信息能力不足且检测效率低的问题，提出一种结合特征预提取模块和残差注意力模块（feature pre-extraction module-residual attention module，FRAM）、Transformer-DSC-Inception-金字塔注意力机制（Transformer-DSC-Inception-pyramid squeeze attention，T-DIPSA）的入侵检测方法，即T-DIPSA-FRAM。该方法融合自适应过采样（adaptive synthetic sampling，ADASYN）、精简编辑最近邻（reduced edited nearest neighbors，RENN）和局部离群因子（local outlier factor，LOF）算法，提高模型在复杂网络流量环境下的检测性能。首先，采用自适应混合采样与离群点检测（AR-LOF）算法平衡数据集；其次，设计包含残差注意力模块的特征预提取模块，初步高效提取网络流量中的关键特征，改善高维特征的学习稳定性；最后，设计局部特征增强注意力模块，利用Transformer编码器结构捕捉长距离依赖关系的同时，融合DIPSA的前馈网络聚焦多尺度局部空间特征，增强模型对动态、非均匀分布流量的敏感性。实验结果表明，在UNSW-NB15数据集和ToN-IoT数据集的二分类和多分类检测任务中，T-DIPSA-FRAM取得的F1值分别为93.58%、95.35%，加权F1值分别为88.26%、91.03%。研究表明，T-DIPSA-FRAM方法能够有效提升网络入侵检测的可靠性。

关键词：网络入侵检测；Transformer；Inception；残差注意力模块；多尺度卷积
中图分类号：TP393
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2026009

Research on network intrusion detection based on fusion Transformer and Inception

Zhang Wanqi, Wang Jiaxing, Song Zhenfeng
College of Information Network Security, People’s Public Security University of China, Beijing 100038, China

Abstract: Aiming at the insufficient capability of current intrusion detection models in extracting information from traffic features and their low detection efficiency, an intrusion detection method named T-DIPSA-FRAM was proposed, which integrates a feature pre-extraction module-residual attention module (FRAM) and a Transformer-DSC-Inception-pyramid squeeze attention mechanism (T-DIPSA). This method combined adaptive synthetic sampling

收稿日期：2025−07−14；修回日期：2025−09−16
通信作者：宋振峰，songzhenfeng@ppsuc.edu.cn
基金项目：中央高校基本科研业务费专项资金资助项目（No.2023JKF01ZK07）
Foundation Item: The Central Universities Basic Research Business Special Funds Project of China (No.2023JKF01ZK07)

(ADASYN), reduced edited nearest neighbors (RENN), and local outlier factor (LOF) algorithms to enhance the detection performance of the model in complex network traffic environments. Firstly, an adaptive hybrid sampling and outlier detection with LOF (AR-LOF) algorithm was employed to balance the dataset. Then, a feature pre-extraction module incorporating a residual attention module was designed to efficiently extract key features from network traffic, improving the learning stability of high-dimensional features. Finally, a local feature-enhanced attention module was designed. While capturing long-range dependencies using the Transformer encoder structure, the feedforward network of DIPSA was integrated to focus on multi-scale local spatial features, enhancing the model's sensitivity to dynamic and non-uniformly distributed traffic. Experimental results demonstrate that on binary and multi-class classification tasks using the UNSW-NB15 and ToN-IoT datasets, T-DIPSA-FRAM achieved F1 scores of 93.58% and 95.35%, and weighted F1 scores of 88.26% and 91.03%, respectively. The study indicates that the T-DIPSA-FRAM method can effectively improve the reliability of network intrusion detection.

Key words: network intrusion detection, Transformer, Inception, residual attention module, multi-scale convolution

0 引言

随着云技术的大规模应用，网络设备规模持续扩大，在互联网中积累了大量宝贵的信息资源。然而，复杂多变的网络环境也使其更容易遭受各类攻击威胁。为有效检测未经授权的访问、异常活动或潜在的攻击行为，开发基于网络流量的数据监管工具，实现高效入侵检测，已成为一项迫切需求[1]。入侵检测系统（intrusion detection system，IDS）是一类用于监测网络流量或系统行为的安全工具[2]。IDS通过实时监测网络和系统活动，能够及时发现并告警潜在的恶意行为，防止数据泄露和系统受损。IDS还能识别多种网络攻击类型，如DDoS攻击、恶意软件、SQL注入等，从而增强整体防护系统的多层防御能力，提升网络环境的稳定性与安全性。

目前，国内外关于入侵检测技术的主要研究如下：文献[3]提出了一种基于树的堆叠集成技术，并通过计算网络流量数据集中特征的方差，选出波动较大特征，以达到特征降维、提高检测性能的目的。该方法能显著降低特征维度，但其使用的过滤式特征选择方法可能忽略了特征间的复杂关联。文献[4]融合基尼不纯度和随机森林方法来进行特征选择，该方法虽然能够自动、稳定地评估特征重要性，适应高维数据并减少过拟合，但也存在计算开销大、模型可解释性差且对冗余特征处理不足的问题。文献[5]使用一种改进的堆叠式自动编码器进行数据降维，虽通过引入Dropout来增强模型的鲁棒性并防止过拟合，但也可能导致训练过程不稳定，且在某些情况下会降低模型的收敛速度。文献[6]采用了基于深度卷积神经网络堆叠集成的方法，对恶意软件进行多尺度的特征提取，测试结果表明，该方法的分类效果显著。为提高模型对不同类型攻击的识别能力，现有针对入侵检测的研究大多采用公开的、类别不平衡的数据集，而文献[7]提出了一种基于生成对抗网络和深度神经网络相结合的入侵数据增强方法，以实现样本集的类别均衡。文献[8]提出一个结合去噪自动编码器和Wasserstein生成对抗网络（Wasserstein generative adversarial network，WGAN）的新型架构，有效解决数据不平衡问题，并增强基于异常的入侵检测能力。虽然WGAN通过Wasserstein距离改善了生成过程的稳定性，但生成的样本可能与真实数据存在差异。文献[9]针对数据类别不平衡的问题，采用SMOTE（synthetic minority over-sampling technique）对少数类别进行过采样，并重新生成数据。利用SMOTE均匀生成少数类样本的合成样本，可能导致对易分类样本的过度采样，从而忽视难分类样本的学习，进一步降低模型在复杂决

策边界上的表现。文献[10]则利用了 Borderline-SMOTE 平衡类别，避免了该问题。文献[11]通过对比 3 种数据不平衡处理技术（SMOTE、Borderline-SMOTE 和 ADASYN）结合卷积神经网络在内部威胁检测中的效果，发现自适应过采样（adaptive synthetic sampling，ADASYN）会优先生成难分类样本，帮助模型更好地学习决策边界，从而增强对少数类的识别能力，且其结合卷积神经网络在提高检测准确性方面表现最优，但未考虑原始数据中的噪声数据。

应用于入侵检测的深度学习模型的相关研究如下。文献[12]采用卷积神经网络结合长短期注意力网络的结构作为入侵检测模型，但长短期注意力网络对长序列依赖建模效果有限。文献[13]采用 Transformer 获取数据特征的长距离依赖，但未考虑聚焦流量的局部细粒度特征。文献[14]采用变分高斯模型和独热编码技术对流量数据进行预处理，再利用金字塔深度可分离卷积神经网络进行入侵检测，虽提高了检测效率，但增加了计算复杂度。文献[15]利用 Inception V4 网络，通过增加网络的深度与宽度，实现网络攻击的检测与分类。此外，文献[16]提出了将卷积神经网络和 Inception 网络结构结合在一起的入侵检测模型，借助多尺度特征提取方法获取流量特征，但由于多尺度特征提取过程中未对通道特征权重进行区分，检测效率较低。

综合上述分析，现有入侵检测技术存在以下问题：（1）传统特征降维方法多依赖统计特性，容易丢失重要信息，且难以充分捕捉特征间的非线性关联；（2）数据样本普遍存在类别不平衡问题，且不能有效剔除样本中的噪声点；（3）在多尺度特征提取过程中，通道权重分配往往缺乏动态调整机制，导致难以充分融合网络流量数据的全局上下文信息和局部细粒度特征，进而导致检测效率低下。针对上述问题，本文提出一种结合特征预提取模块和残差注意力模块（feature pre-extraction module-residual attention module，FRAM）、Transformer-DSC-Inception-金字塔注意力机制（Transformer-DSC-Inception-pyramid squeeze attention，T-DIPSA）的入侵检测方法——T-DIPSA-FRAM。T-DIPSA-FRAM 可深层次捕捉流量特征的局部和全局上下文信息，增强模型分类能力。同时，针对数据中存在的离群点和类别样本不平衡问题，引入自适应过采样（adaptive synthetic sampling，ADASYN）与局部离群因子（local outlier factor，LOF）算法。本文所提 T-DIPSA-FRAM 方法在实验中表现出对少数类别样本良好的检测效果。

T-DIPSA-FRAM 中的特征预提取模块包含多层卷积和残差注意力模块（residual attention module，RAM），能够自动学习并提取输入流量数据的关键特征，降低数据特征维度。ADASYN-RENN-LOF 策略通过生成难分类样本和重采样，有效增强少数类样本的表示，解决类别不平衡问题，并提升数据集的平衡性。融合了 Transformer-DSC-Inception 和金字塔注意力机制的特征增强模块 T-DIPSA，借助 Transformer 的自注意力机制捕捉全局依赖，DSC-Inception 模块通过多尺度卷积提取多层次特征，金字塔注意力（pyramid squeeze attention，PSA）机制增强了对不同攻击模式的感知能力。此外，模型还增强了对位置信息的敏感性，使其能够关注序列中重要的时间点或特征，进一步提升对复杂入侵行为的区分能力。本文在 UNSW-NB15 和 TON_IoT 数据集上进行了实验验证，结果证明了所提方法的可靠性与有效性。

1　本文方法

T-DIPSA-FRAM 方法结构示意图如图 1 所示。

在数据预处理阶段，首先，进行缺失值填补，以保障数据的完整性，避免数据缺失引发模型训练不稳定或性能下降。接着，将非数值型数

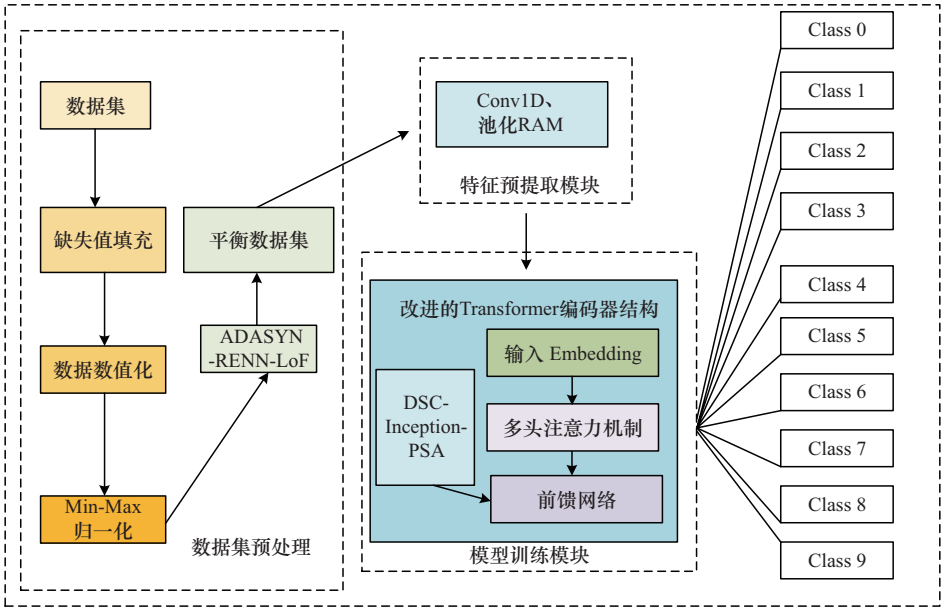

图1　T-DIPSA-FRAM方法结构示意图

据转换为数值形式数据，以满足模型的训练需求。然后，采用最小–最大归一化（Min-Max normalization）对数据进行归一化处理，消除不同特征之间的量纲差异，让每个特征处于相同的尺度，从而加快模型的收敛速度并提高模型的稳定性。最后，使用ADASYN-RENN-LOF处理数据的不平衡问题、噪声问题以及特征差异，帮助入侵检测模型更好地捕捉重要信息，提升其准确性和泛化能力。

特征预提取阶段，基于平衡后的数据集，通过逐步缩小的卷积来提取多尺度特征，聚焦于更重要的局部特征，有助于模型逐步精炼特征。模块中加入的残差注意力模块能够优化特征提取过程，提高模型对难分类样本的学习能力。最大池化层和Dropout的使用，有助于降低计算复杂度并防止过拟合，提升模型的泛化能力。此外，LeakyReLU和多个卷积层的结合增强了非线性表达能力，进一步提升了模型在复杂数据场景下的表现。

特征预提取为模型训练提供了关键特征并减少了冗余信息，改进的Transformer编码器借助自注意力机制捕捉长距离依赖，增强了时序信息建模能力。同时，添加了DSC-Inception-PSA模块的前馈网络通过多尺度卷积提取不同层次的特征，而金字塔注意力机制作为后处理模块能够优化这些提取出来的特征，将更多的注意力集中在最具辨识度的局部特征区域，进一步减少噪声干扰，增强了模型对少数类样本的识别能力。

此次方法设计涵盖了数据集平衡、特征预提取，并融合流行的Transformer编码器和多尺度卷积DSC-Inception-PSA模块，有效增强了模型的鲁棒性和可靠性，在入侵检测任务中具有显著优势。

2　模块技术

2.1　平衡数据集

ADASYN合成少数类样本时受邻域多数类样本干扰，易产生不同于原始少数类样本特征的噪声数据。因此，引入LOF算法，在数据采样处理后，针对每一少数类样本计算其离群因子$\mathrm{lof}_{(p_i)}$。设定$\mathrm{lof}_{(p_i)}>1$即噪声数据点，从而将其剔除。此次实验结合了ADASYN、RENN[17]和LOF算法，

以实现对数据集的平衡处理。具体处理过程如下。

（1）输入少数样本集A，多数样本集R。

（2）计算合成的类样本G。其中，A代表少数样本，R代表多数样本，$\beta \in (0,1)$。

$$G=(R-A)\beta \tag{1}$$

（3）对于每个少数类样本，计算其当前邻居中的多数类别的比例r_i，并通过欧几里得距离计算其邻居数k，K_1代表当前其邻居中的多数类样本。

$$r_i=\frac{K_1}{k} \tag{2}$$

（4）根据全局合成类样本G与每个少数样本的不平衡比率r_i动态计算少数类样本的合成数量g，实现数据分布自适应的过采样。每个少数类样本的合成数g的数学表达式如下：

$$g=G\times r_i \tag{3}$$

（5）每个少数类新合成的样本为z_i，x_i代表当前的少数类样本，x_{zi}代表x_i的k邻居中的随机少数类样本，$\lambda \in (0,1)$。

$$z_i=x_i+(x_{zi}-x_i)\times\lambda \tag{4}$$

（6）重复步骤（4）和（5），直至生成稳定的、新的少数类数据集new$A$。

（7）对于多数样本R中的每一个样本，计算其每个K_2近邻居数中的少数类数量，若大于$e=1$，则剔除该样本。

（8）重复步骤（7）直至生成稳定的、新的多数类样本集new$R$。

（9）对于new$A$，new$R$中的每个样本p，通过欧几里得距离计算近邻居数k'，计算样本p的k'近邻$N_{k'}(x_i)$。

$$N_{k'}(x_i)=\{x_1,x_2,x_3,\cdots,x_{k'}\} \tag{5}$$

（10）计算每个样本p_i的可达距离$d_(p_i,p_j)$，$k'_$代表p_j的第k'近邻距离，$d(p_i,p_j)$是样本p_i与样本p_j的实际距离。

$$d\_(p_i,p_j)=\max\{k'\_(p_j),d(p_i,p_j)\} \tag{6}$$

（11）对于每个样本点p_i，取其k'近邻$N_{k'}(p_i)$的可达距离$d_(p_i,p_j)$的平均数再取倒数，得到局部可达密度LRD(p_i)，为样本清洗提供指标。LRD(p_i)值越高，样本周围邻域越紧凑，属于密集区域；LRD(p_i)值越低，样本周围稀疏，可能是边界点或噪声。计算局部可达密度如式（7）。

$$\mathrm{LRD}(p_i)=\left(\frac{1}{\left|N_{k'}(p_i)\right|}\sum_{p_j\in N_{k'}(p_i)}d\_(p_i,p_j)\right)^{-1} \tag{7}$$

（12）计算每个样本p的$\mathrm{lof}_{(p_i)}$。$\mathrm{lof}_{(p_i)}$>1则说明该样本是离群点，可将其剔除。

$$\mathrm{lof}_{(p_i)}=\sum p_j\in N_{k'}(p_i)\frac{\mathrm{LRD}(p_j)}{\mathrm{LRD}(p_i)} \tag{8}$$

（13）重复步骤（9）至（12），直至剔除噪声。

（14）输出平衡后的数据集new$D$。

具体如算法1所示。

算法1　基于自适应混合采样与离群点检测算法

输入　原始数据集D中的少数样本集A，多数样本集R；

输出　平衡且净化后的少数样本集new$A$，平衡且净化后的多数样本集new$R$，新的数据集new$D$；

计算合成类样本$G=(R-A)\beta$，$\beta \in (0,1)$；

如果$G\leqslant 0$，则结束，返回原始数据集；否则，针对每个少数样本，计算其当前邻居中的多数类别的比例r_i和邻居数k；

对每个少数样本，计算合成样本数$g=G\times r_i$

初始化新少数样本集$A'=\varnothing$；

对每个少数样本$A_i\in A$：对$j=[1,g_i]$，随机选择邻居中的少数样本$A_j\in A$；生成合成样本$a'=A_i+\lambda\times(A_j-A_i),\lambda\in[0,1]$;

将a'加入A'结束循环;

更新少数样本集new$A=A\cup A'$;

初始化多数样本集new$R=R$;

重复执行：对每个多数样本$R_i\in$new$R$：计算邻居中少数样本数量n_i；如果n_i>阈值$\partial=0.1$，则从new$R$中剔除R_i；直到new$R$稳定不变；

对数据集new$D=$new$A\cup$new$R$，重复执行：对每个样本$D_i\in D$：根据式（5）计算近邻集合$N_k(D_i)$；根据式（6）计算每个邻居$D_j\in N_k(D_i)$可达距离$d_k(D_i,D_j)$；根据式（7）和式（8）计算局部可达密度$\mathrm{LRD}_k(D_i)$和$\mathrm{LOF}_k(D_i)$；剔除所有$\mathrm{LOF}_k(D_i)>1$的样本；直到new$D$稳定无离群样本；

输出数据集new$D$。

2.2 特征预提取模块

特征预处理模块包含残差注意力模块和多个不同卷积核大小的卷积层。通过堆叠不同卷积核大小的卷积层，能够提取多尺度数据特征。采用卷积核逐级递减的堆叠方式，有助于扩展后续卷积的感受野，充分捕捉输入的网络流量数据的上下文流动信息，从而利于提取初步的数据关键特征[18]。特征预提取模块设计如图2所示，从左至右的卷积核大小依次设置为7、5、5、3。ReLU可保持正输入的梯度消失，LeakyReLU允许在负输入上有一个固定的小斜率$\alpha=0.2$，避免神经元死亡问题，4层卷积的激活函数以ReLU、LeakyReLU交替设置，以利于流量数据特征预提取过程的收敛和稳健。

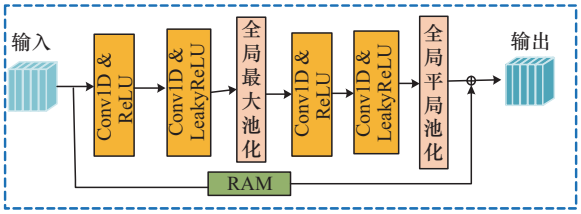

图2 特征预提取模块设计

注意力机制可增强特征信息，抑制冗余，且残差学习可缓解较深的网络在梯度反向传播更新参数时造成的梯度消失问题[19]。受深度收缩网络启发，本次设计的残差注意力模块先使用1×1、3×3、5×5并行提取数据特征，并将其输入、输出跳跃连接相加，以缓解梯度消失的问题。第二阶段卷积进一步提取并输出特征之后加入全局平均池化，生成每个特征图的平均值，记为a'，经Sigmod将其放缩为0和1之间的一个系数μ，取阈值为$\tau=\mu a'$。p为经过卷积层后的特征数据，经过稀疏处理l操作后与原始输入相加，达到去除特征冗余的效果。软阈值函数（曲线如图3所示）有去除噪声的作用，即将阈值在$[-\tau,\tau]$的特征置为0，并收缩距0较远的特征。这一过程用表达式表示如下：

$$l(p)=\begin{cases}p-\tau, & p>\tau\\ 0, & -\tau\leqslant p\leqslant\tau\\ p+\tau, & p<-\tau\end{cases}\tag{9}$$

其中，p为卷积层输出的特征值；$l(p)$为经过软阈值函数处理后的特征值。

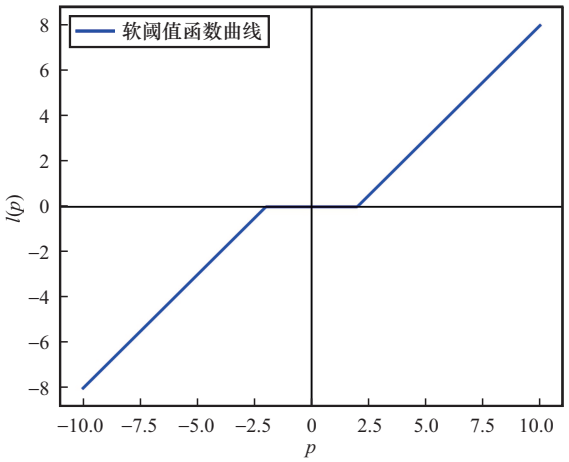

图3 软阈值函数曲线

该软阈值函数的核心作用机制在于对流量序列特征进行“剔除冗余”处理：首先，对于落在阈值区间$[-\tau,\tau]$内的特征值p，函数将其置为0，这相当于滤除较小幅度的冗余成分；然后，针对超出阈值区间的特征值p，则通过减去或加上阈值的方式进行收缩，在保留重要特征信息的同时减少冗余影响，防止出现潜在的无效特征被放大

导致的过拟合。

本次特征提取模块的创新性设计是将卷积层的激活函数以 ReLU 和 LeakyReLU 交替设置，适用于处理网络流量 1 维时间序列流。通过在堆叠逐级递减的卷积层之间加入 RAM，如图 4 所示，可自适应调整特征权重，并融合卷积层之间的关键特征，从而达到解决特征冗余问题的效果，这可为后续模型训练提供更有效的数据信息特征。

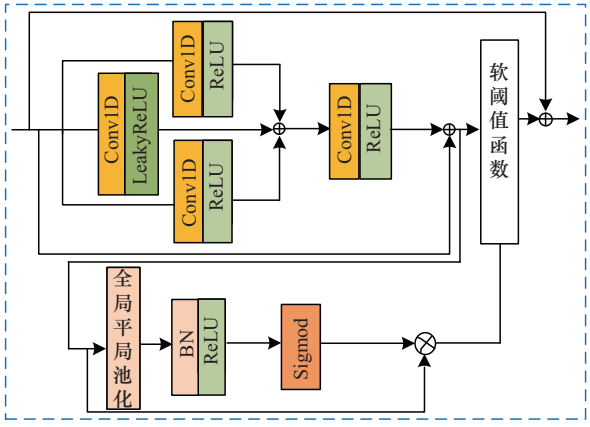

图 4　残差注意力模块

2.3　Transformer 编码器结构

Transformer 架构适用于处理序列数据，如自然语言处理（natural language processing，NLP）中的文本生成、机器翻译等任务[20]。Transformer 架构主要由编码器（Encoder）和解码器（Decoder）组成。其中，编码器负责将输入序列映射为一组包含上下文信息的特征表示，从而有效提取数据的全局特征；而解码器则利用编码器的输出，生成目标序列。在入侵检测任务中，核心目标是对网络流量或日志数据进行分类（如正常/恶意）或异常检测，这是一个典型的判别式任务，而非生成任务。因此，只采用 Transformer 编码器部分。

Transformer 编码器由多个相同的层堆叠而成，每层包含多头注意力机制（multi-head self-attention）和前馈网络（feed-forward network，FFN）两个子层。每个子层周围都采用残差连接和层归一化进行封装。在设计入侵检测模型时，首先通过特征预提取模块初步捕获网络流量数据的高维特征；随后将这些特征输入 Transformer 嵌入层，并转换为包含位置编码的数据表征；接着，通过多层自注意力机制堆叠的多头注意力模块计算加权，捕捉上下文信息；然后，经过残差和层归一化操作转为稳定的信息流；最后，到达 FFN，并进一步进行非线性变换和提取信息特征。

Transformer 编码器部分的多头注意力机制是其多个自注意力层的扩展，而每个自注意力模块通常使用矩阵乘法将输入序列转换为 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 3 种不同向量。其中，$\boldsymbol{Q}$ 代表查询向量，表示当前正在处理的位置或词的表示；$\boldsymbol{K}$ 代表键向量，表示当前信息的特征向量的重要程度；$\boldsymbol{V}$ 代表值向量，通常是输入序列的特征表示，最终与注意力权重结合，以决定输出的特征。过程实现通过如下表达式表示：

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \quad (10)$$

其中，d_k 为键向量的维度；$\sqrt{d_k}$ 为缩放因子，防止点积过大导致梯度消失或爆炸。

编码器中包含的多头注意力机制实际就是堆叠拼接多个自注意层的输出特征向量，可关注来自不同表示空间的多组注意力特征信息，具体可由式（11）实现。

$$\boldsymbol{Z}=\text{concat}(h_1,h_2,h_3,\cdots,h_h)w^{\circ} \quad (11)$$

其中，$(h_1,h_2,h_3,\cdots,h_n)$ 代表 n 个注意力头，$\boldsymbol{Z}$ 代表混合拼接后的输出，w° 代表一个权重矩阵。

通过多头注意力机制，Transformer 编码器能够学习输入数据在不同表示空间中的多样化特征，从而增强其特征捕捉能力。自注意力模块通过并行计算对输入序列中的所有元素同时执行点积注意力操作，捕获序列中任意位置之间的依赖

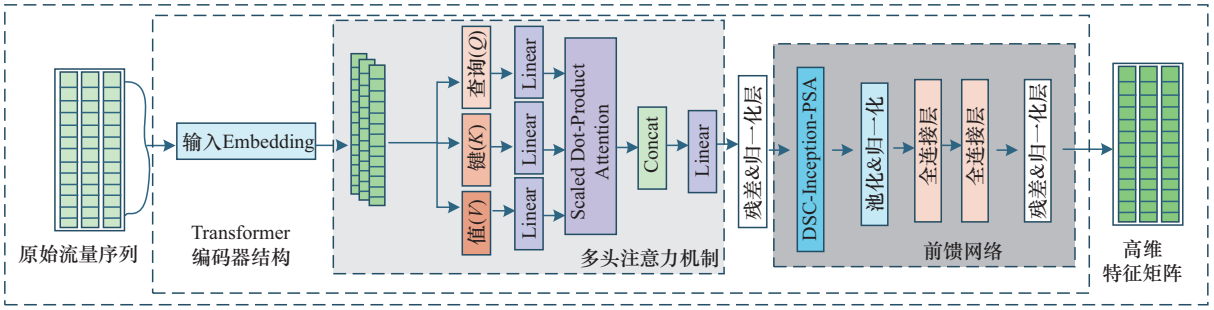

图5　改进后的Transformer编码器结构

关系，是实现长距离依赖的关键。为捕捉元素位置信息，Transformer中加入了位置编码器，可在快速捕捉信息内容的同时兼顾其位置信息。位置编码的实现如式（12）、式（13）所示。

$$PE_{(pos,2i)}=\sin\left(\frac{pos}{1000^{\frac{2i}{d}}}\right) \tag{12}$$

$$PE_{(pos,2i+1)}=\cos\left(\frac{pos}{1\,000^{\frac{2i}{d}}}\right) \tag{13}$$

其中，$PE_{(pos,i)}$代表位置编码操作，pos代表元素位置，d代表位置编码的维度。特征维度是偶数时，采用sin函数计算其位置编码；特征维度是奇数时，采用cos函数计算其位置编码。

FFN通常由两个线性层组成，中间有激活函数ReLU，无法有效提取重要特征[21]。本次设计将DSC-Inception-PSA结构插入FFN中，从而进一步提取序列关键信息特征。改进后的Transformer编码器结构如图5所示。通过多头注意力机制捕捉全局信息，利用融合DSC-Inception-PSA结构的FFN有效提取多尺度特征信息，并自适应地加权重要的局部细粒度空间信息。引入批归一化处理，以加速训练且稳定学习过程，再结合全局平均池化减少特征维度，有效提升了Transformer编码器结构对流量关键特征区域的建模能力。

2.4　DSC-Inception结构

受GoogLeNet团队提出的Inception网络结构启发，考虑到深度可分离卷积相比原始卷积更轻便[22]，设计DSC-Inception结构，如图6所示。该结构通过3个并行分支提取多尺度的特征矩阵，而后在拼接层获得深度特征表示，以提高模型泛化能力。DSC-Inception先利用1×1卷积降维，再扩大感受野，采用并行的卷积操作来提取输入流量数据的不同维度特征表征，最后聚合不同维度特征矩阵输出更高效的特征信息序列。

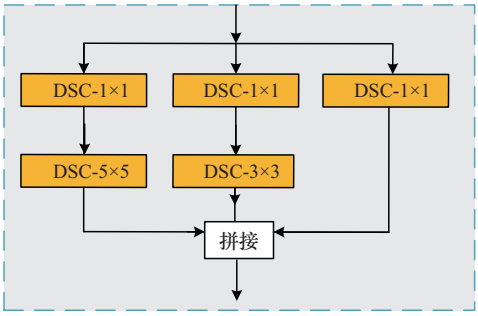

图6　DSC-Inception结构

2.5　金字塔注意力

为提高检测效率，引入图像领域的金字塔注意力[23]，并将其转换为适合处理流量数据的实现方式。该机制首先通过挤压拼接（squeeze-and-concatenation，SPC）模块将输入的特征序列划分为S部分，然后通过SEWeight计算各部分数据相应的权重并拼接，使用Softmax函数对通道数据特征权重再次校准，最后将校准后的多尺度特

征与原始输入的相应特征进行逐元素乘积运算，输出获得的多尺度融合特征图。整个流程包括分离输入、卷积、注意力加权和特征融合，旨在捕捉输入数据中的不同位置和尺度的细粒度特征，并通过 SPC 和 SE 分组动态加强通道权重，增强重要特征的表达。PSA 的实现过程如图 7 所示。

其中，SPC 模块是实现 PSA 权重的关键部分。首先，将输入的流量数据序列转换为张量 $m=(b,h,c)$，其中 b 代表批次大小，h 代表序列长度，c 代表每个序列的通道数。其次，借助不同卷积核挤压通道张量维度，获取多尺度关键信息表示，主要是将输入的数据序列张量均分为 s 个片段，每一片段的通道维数为 c/s，其中 c 代表原本的通道维数。然后，通过独立的卷积层对每个片段应用卷积操作，输出特征图 f_i，其中 f_i 代表第 i 个片段 $(i=1,2,3,\cdots,s)$。每一层卷积核大小 $k''=2\times(i+1)+1$，其中 i 代表第 i 层卷积（$i=1,2,3,\cdots,s$）。随着 i 的增加，卷积核会变大，导致计算量增加，故每一片段的分组大小需动态确定，具体的分组为 $G'=2^{\frac{k''-1}{2}}$，其中 G' 为分组大小，k'' 代表卷积核的大小。将所有的特征图堆叠输出新的张量 $\mathrm{spc}_{\mathrm{out}}=\mathrm{Cat}\left[\left(f_1,f_2,f_3,\cdots,f_i\right)\right]$，其中，$i=s$，Cat 代表拼接操作。较大 s 会引入大核卷积，增加卷积操作感受野，较大分组 G' 则强化 SEweight 注意力效果，故本次 PSA 中 s 和 G' 的设置见第 3.3 节。这种设计能够同时处理多个不同尺度的局部信息，并通过后续的注意力机制（SE 模块和 Softmax）来加权和融合这些信息。

2.6　DSC-Inception-PSA 模块

DSC-Inception-PSA 结构如图 8 所示。此次实验在模型部分堆叠了 2 个改进后的 Transformer 编码器，用于捕捉流量数据全局上下文时序特征信息。在 FFN 层中，先对输入数据进行归一化处理，再通过两个标准化卷积层对特征进行初步增强。随后，引入 DSC-Inception 结构，利用不同尺度的卷积核并行提取局部特征，并通过 PSA 机制对不同尺度特征加权。DSC-Inception-PSA 模块可借助输入流量的全局时序信息特征，自适应地

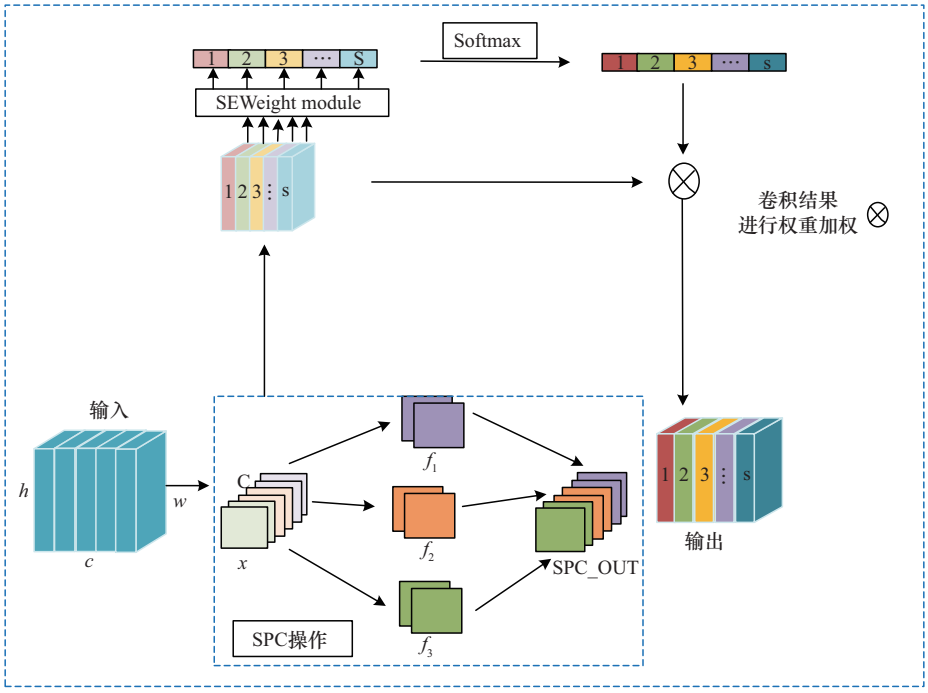

图 7　PSA 的实现过程

调整对局部流量特征的关注重点，从而提升对攻击流量行为的学习能力。具体而言，设FFN的输入特征矩阵为$X1$，经归一化和两层卷积处理后的输出数据为$X2$，再经DSC-Inception多尺度特征提取：$F_{X_k}=\mathrm{PW}(\mathrm{DW}(X2)), k\in\{1,3,5\}$。其中，PW代表逐点卷积，DW代表逐通道卷积。接着，经过PSA卷积加权使模型可对关键攻击特征有更强的响应能力：$F_{\mathrm{PSA}}=F_{\mathrm{concat}}\sigma(f(F_{\mathrm{concat}}))$。其中，$\sigma$代表Sigmoid函数，$f(\cdot)$代表卷积操作。最后，通过LayerNorm、Dropout与平均池化等操作进一步提升模型训练的稳定性与泛化性能。该结构有效增强了Transformer编码器结构对复杂流量的全局和局部特征的捕捉能力。

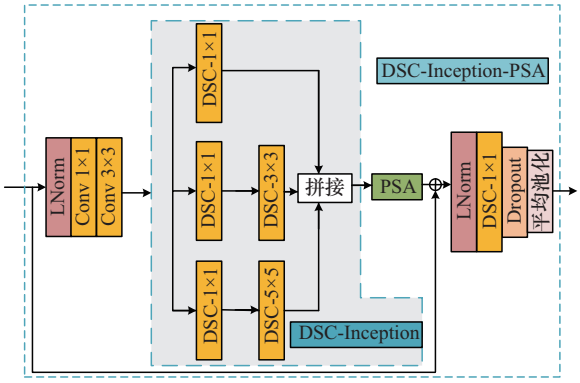

图8　DSC-Inception-PSA结构

3　实验结果与分析

3.1　数据集

UNSW-NB15数据集[24]由澳大利亚新南威尔士大学提供，是广泛用于网络入侵检测研究的公开数据集。该数据集模拟了真实网络环境中的流量状况，包含49个特征。其流量涵盖10种类别，包括9种攻击类别和1种正常流量。在多分类时，将数字标签0、1、2、3、4、5、6、7、8、9分别赋给Analysis、Backdoor、DoS、Exploits、Fuzzers、Generic、Normal、Reconnaissance、Shellcode、Worms。UNSW-NB15数据集中各类别的情况见表1。

表1　UNSW-NB15数据集

类别	训练集/个	测试集/个	描述
Worms（9）	130	44	计算机蠕虫
Shellcode（8）	1 133	378	漏洞利用载荷（常用于溢出等攻击）
Reconnaissance（7）	10 491	3 496	侦察攻击
Generic（5）	40 000	18 871	泛型攻击
Exploits（3）	33 393	11 132	漏洞利用攻击
Fuzzers（4）	18 184	6 062	模糊测试攻击
DoS（2）	12 264	409	拒绝服务攻击
Backdoor（1）	1 746	583	后门攻击
Analysis（0）	2 000	677	分析型攻击（如流量嗅探、信息收集）
Normal（6）	56 000	37 000	正常的流量

ToN-IoT数据集[25]由新南威尔士大学堪培拉网络工程与信息技术学院的物联网实验室创建，专为物联网网络安全研究而设计，用于评估物联网设备的入侵检测系统，见表2。本研究选择其Network数据集，包含43个特征。流量共分为10个类别（9种攻击类型和1种正常流量），在多分类任务中，将数字标签0、1、2、3、4、5、6、7、8、9分别赋给Password、Injection、Backdoor、Scanning、Mitm、Normal、Ransomware、DoS、DDoS、Xss。

表2　ToN-IoT数据集

类别	数量/个	描述
Ransomware（6）	20 000	勒索病毒
Password（0）	20 000	密码攻击
Xss（9）	200	跨站脚本攻击
Mitm（4）	1 043	中间人攻击
Injection（1）	20 000	注入攻击
DDoS（8）	20 000	分布式拒绝服务攻击
DoS（7）	20 000	拒绝服务攻击
Backdoor（2）	20 000	后门攻击
Scanning（3）	20 000	扫描攻击
Normal（5）	50 000	正常的流量

3.2　评价指标

本文采用准确率（Accuracy）、召回率（Recall）和F1值（F1-score）来衡量本次实验在UNSW-NB15数据集和ToN-IoT数据集上的有效性。这些

指标均可通过混淆矩阵获得。真阳性（true positive，TP）：模型预测和实际是正类的样本数量；真阴性（true negative，TN）：模型预测和实际是负类的样本数量；假阳性（false positive，FP）：模型预测为正类，但实际是负类的样本数量；假阴性（false negative，FN）：模型预测为负类，但实际是正类的样本数量。准确率（Accuracy）、召回率（Recall）和F1值（F1-score）的表达式如式（14）～式（17）所示。

$$\text{Accuracy}=\frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}\times 100\% \quad (14)$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}}\times 100\% \quad (15)$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}}\times 100\% \quad (16)$$

$$\text{F1-score}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\times 100\% \quad (17)$$

针对多分类实验，需要对F1值的计算方式进行调整，即使用求加权平均的F1值（weighted F1-score，F1_{wei}），根据式（17）计算每个类别的F1值，即为F1_i，再按各类别的样本数量进行加权平均得到F1_{wei}。得其F1值，求加权F1值的计算方式如下：

$$\text{F1-score}_{\text{wei}}=\frac{\sum_{i=0}^{n-1}(\text{F1}_i\times N_i)}{\sum_{i=0}^{n-1}N_i} \quad (18)$$

3.3　实验参数设置

本实验中Transformer编码器、PSA模块的参数设置见表3、表4。

表3　Transformer编码器参数设置

参数名称	参数值
多头注意力头数	4
特征隐藏维度	64/128
前馈中间层网络维度	512
编码器堆叠层数	2
层归一化率	0.000 01

表4　PSA模块参数设置

通道分组数（$s=4$）	分组大小	核大小
0	1	3
1	4	5
2	8	7
3	16	9

3.4　模型计算复杂度分析

为评估本文所构建模型的实际应用潜力，需对所构建模型的计算复杂度进行分析。神经网络模型的计算复杂度主要包括空间复杂度和时间复杂度，其大小直接反映模型在实际应用中的运行效率和资源损耗。本文所构建模型主要包括特征预提取模块、Transformer编码器和DSC-Inception-PSA模块，各模块在提取强化流量序列特征中协同作用，有效提升模型的整体检测能力。

针对单个序列长度为L的流量数据，特征预提取模块包含若干个卷积层和阈值函数，总体时间复杂度为$O\left(Ld_{\text{in}}d_{\max}\right), d_{\max}=64/128$。其中，$d_{\max}$代表二分类或多分类任务时的最大特征通道数，$d_{\text{in}}$代表输入的特征维度。空间复杂度为$O(d^2{}_{\text{out}})+O(Ld_{\text{out}})$。

Transformer编码器的计算复杂度则主要来自多头注意力机制和前馈神经网络。多头注意力机制中，输入序列经线性变换生成$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$矩阵，每个时间复杂度为$O(Ld_{\text{model}}^2)$，计算注意力矩阵分数所需点积操作的复杂度为$O(L^2d_{\text{model}})$。因此，多头注意力机制的总体时间复杂度为$O(L^2d_{\text{model}}+Ld_{\text{model}}^2)$。其中，$d_{\text{model}}$代表模型在二分类或多分类任务时的隐藏维度，本次设置$L=43/49$，$d_{\text{model}}=64/128$，故多头注意力时间复杂度取决于d_{model}。不过，注意力头数越多，线性变换矩阵$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$计算次数越多，这也会影响模型的实际运行效率，因此注意力头数取4。前馈神经网络部分采用两层全连接网络，特征维度从d_{model}变换为d_f，再回到d_{model}，从而这一过程的时间复杂度为$O(Ld_{\text{model}}d_f)$。其中，$d_f=512$，代表中间层维度。

综合可知，Transformer编码器的整体时间复杂度为$O(Ld_{\text{model}}d_f+Ld_{\text{model}}^2)$，空间复杂度为$O(Ld_{\text{model}}d_f+Ld_{\text{model}}+L^2)$。

DSC-Inception-PSA模块用于提取流量序列L的多尺度特征信息，其中，DSC-Inception结构的时间复杂度为$O(Ld_{\max}m)$，m为DSC-Inception各分支所有特征通道之和；PSA模块则将特征通道划分为s个片段，片段内采用不同核卷积构建信息细粒度特征，其时间复杂度为$O(Ld_{\max}S+d_{\max}^2/S)$；DSC-Inception-PSA模块采用分组卷积和轻量注意力机制控制了模块参数量，在提升特征表达能力的同时节约了计算开销。

综上分析，本文构建的网络流量入侵检测模型的主要参数如图9所示。结合Transformer与DSC-Inception-PSA模块，本文所提模型可同时实现时序特征建模与多尺度空间信息提取，提升捕获到的流量特征的区分度。输入维度为（43/49，1）的原始流量特征序列，经Transformer模块，二分类和多分类输出维度分别提升至64、128，增强了全局上下文信息建模能力；DSC-Inception-PSA模块在整合提取到的局部关键信息后，Flatten层的二分类和多分类输出特征维度达1 312、2 324。

为实现本文模型训练，实验环境设计为Windows 11系统，32 GB内存，搭配Intel Core i7-14650HX和GPU NVIDIA RTX 4060处理器，基于Python 3.12.0与TensorFlow 2.18.0，CUDA 11.8。模型计算复杂度见表5。当总训练轮数为80、批次大小为32时，本文所提模型每轮训练平均时间约为50 s、平均单个样本检测用时约8 ms，体现出本文模型在保证高性能检测的同时具备良好的训练效率与部署可行性，具有较强的实际应用价值。

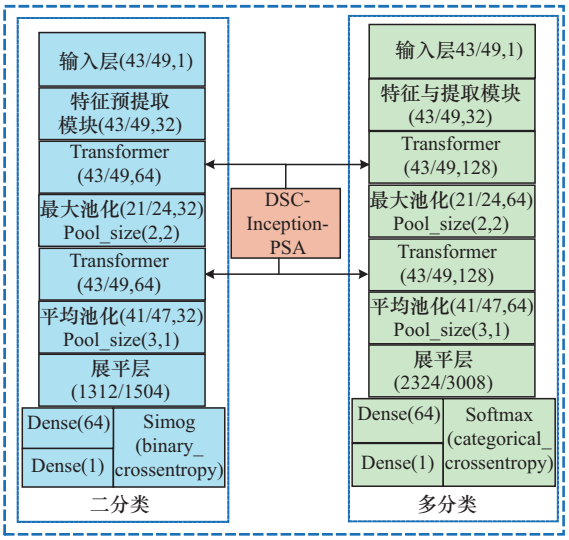

图9　模型的主要参数（二分类/多分类）

3.5　数据采样性能分析

使用AR-LOF对UNSW-NB15数据集进行采样，采样前后流量数据类别对比如图10所示。其中，少数类别如Analysis（0）、Backdoor（1）、Shellcode（8）、Worms（9）采样后分别上升了2.89%、1.72%、3.40%、1.69%，多数类别如Generic（5）、Normal（6）分别下降了5.97%、7.57%。由此可见，使用AR-LOF采样有效平衡了数据分布。

采用AR-LOF采样与ADASYN过采样方法合成少数类样本，以缓解类别不平衡问题。同时结合RENN方法适量削减多数类样本，并利用局部

表5　模型计算复杂度

模型	分类任务	数据集	空间复杂度/$\times10^6$	时间复杂度	
				每轮训练时间/s	平均检测时间/ms
本文模型	二分类	UNSW-NB15	0.82	51	7.3
		ToN-IoT	0.82	46	8.1
	多分类	UNSW-NB15	1.96	53	7.6
		ToN-IoT	1.96	49	8.3

离群值检测算法在混合采样过程中剔除噪声与异常样本。UNSW-NB15 数据集有无采样的评价指标对比如图 11 所示。由图 11 可知，该策略有效增强了对少数类别的辨识能力，使模型在多分类任务中的加权召回率提升 5.96%，加权 F1 值从 83.01% 提高至 88.26%。

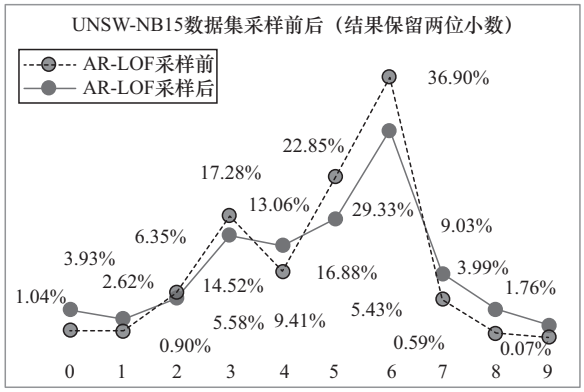

图 10　AR-LOF 采样前后 UNSW-NB15 数据集类别变化

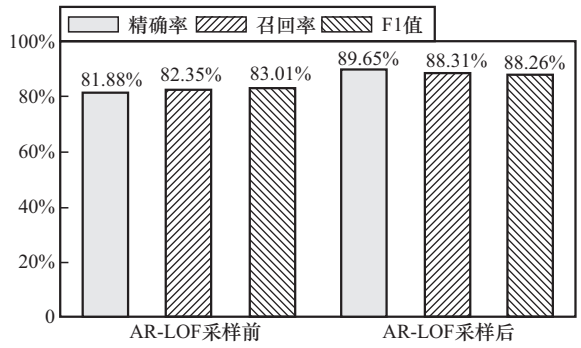

图 11　UNSW-NB15 数据集有无采样的评价指标对比

3.6　不同阈值对模型性能的影响

为进一步增强预提取流量序列的特征表示，本文设计的特征预提取模块包含逐级递减的卷积层和 RAM 模块。软阈值函数具有将不在特定阈值范围$[-\tau,\tau]$的特征值收缩为 0 的能力，将其引入 RAM 模块设计中可有效滤除提取的冗余特征信息，从而提升重要特征捕捉能力。由$\tau=\mu a'$知，阈值τ受放缩系数μ控制，μ的大小直接影响软阈值函数的特征收缩力度，从而影响模型检测能力。为探究阈值对本次模型性能的影响，设置μ范围为[0.0, 0.9]，分别在 UNSW-NB15 数据集和 ToN-IoT 数据集上进行对比实验。不同缩放系数在 UNSW-NB15 数据集和 ToN-IoT 数据集的表现分别见表 6、表 7。

表 6　不同缩放系数在 UNSW-NB15 数据集的表现

μ	F1 值	加权 F1 值
0.0	92.89%	87.62%
0.1	**93.66%**	87.75%
0.2	93.26%	88.24%
0.3	**93.58%**	**88.26%**
0.4	92.81%	**88.44%**
0.5	92.85%	87.22%
0.6	92.79%	86.91%
0.7	92.21%	86.88%
0.8	91.92%	86.5%
0.9	91.8%	85.5%

表 7　不同缩放系数在 ToN-IoT 数据集的表现

μ	F1 值	加权 F1 值
0.0	95.29%	90.62%
0.1	**95.35%**	**91.03%**
0.2	95.26%	**91.05%**
0.3	94.51%	90.22%
0.4	94.34%	89.38%
0.5	94.62%	89.66%
0.6	94.48%	88.71%
0.7	93.7%	87.9%
0.8	93.61%	88.13%
0.9	93.75%	88.59%

实验结果显示，在 UNSW-NB15 数据集和 ToN-IoT 数据集上，不同的阈值确实会显著影响模型的检测性能。当$\mu=0.0$时，RAM 模块丧失去除冗余特征的能力，仅表现为提取特征的普通卷积结构；随着μ的缓慢上升（临界值之前），模型性能有所增强。由表 6 可知，μ在 0.0 至 0.9 之间，模型在 UNSW-NB15 数据集的性能有所波动。其中，μ在 0.1 至 0.3 之间，F1 值和加权 F1 值逐渐上升，最终取得峰值为 93.58% 和 88.26%，表明较弱的特征放缩力度有助于去除冗余特征，提高模型对异常数据的敏感度；而当μ在 0.4 至 0.9 之间

时，特征放缩力度过强会导致潜在的有效信息丢失，进而使模型的F1值和加权F1值分别下降1.78%和2.76%。由表7可知，μ在0.1至0.2之间，模型性能稳定且有效，F1值和加权F1值保持在95%和91%以上，而当μ继续增大至0.9，模型性能则分别下降1.6%和2.44%。

综上分析可知，软阈值函数的引入直接影响模型检测结果。放缩系数μ在临界点（本文临界点为0.3和0.1）附近可显著剔除部分无价值的冗余特征，强化对关键数据特征的学习表达；而μ过高会导致大量真实有效的流量序列特征被收缩至0，削弱特征完整性，导致模型的F1值下降；在多分类任务中，类别样本间细粒度和特征复杂度更高，特征信息丢失负影响效果更为明显，加权F1值下降幅度更大，表明模型细分类别能力减弱。

3.7 消融实验

为探究不同特征融合模块对基于Transformer-DSC-Inception-PSA模型检测性能的影响，本文在多个模型结构上进行了对比实验，包括Transformer、T-DSC-Inception、T-DIPSA、T-DIPSA-F和T-DIPSA-FRAM。这些模型分别结合了DSC-Inception结构、PSA模块、特征预提取模块及RAM模块。

在UNSW-NB15数据集和ToN-IoT数据集的二分类和多分类检测任务中，对比模型的性能检测效果见表8。Transformer模型的表现最弱，加入了DSC-Inception-PSA和特征预提取模块的T-DIPSA-F模型，增强了对动态非均匀分布流量序列特征的捕获能力，能动态调整感受野范围并聚焦异常流量特征在空间维度上的分布不均匀特性，同时提炼深层流量序列特征。由表8可

表8 对比模型的性能检测效果

数据集	分类任务	模型	（加权）准确率	（加权）召回率	（加权）F1值
UNSW-NB15	二分类	Transformer	91.99%	91.84%	91.92%
		T-DSC-Inception	92.89%	92.84%	90.92%
		T-DIPSA	93.01%	92.72%	92.88%
		T-DIPSA-F	92.87%	93.35%	92.91%
		T-DIPSA-FRAM	**93.05%**	**93.78%**	**93.58%**
	多分类	Transformer	86.24%	85.17%	85.56%
		T-DSC-Inception	86.84%	86.5%	86.33%
		T-DIPSA	87.03%	87.03%	86.61%
		T-DIPSA-F	88.55%	87.45%	87.77%
		T-DIPSA-FRAM	**89.65%**	**88.31%**	**88.26%**
ToN-IoT	二分类	Transformer	93.49%	93.89%	94.82%
		T-DSC-Inception	92.80%	92.50%	93.58%
		T-DIPSA	93.81%	92.72%	94.88%
		T-DIPSA-F	94.87%	94.35%	94.91%
		T-DIPSA-FRAM	**95.17%**	**94.52%**	**95.35%**
	多分类	Transformer	88.64%	87.34%	87.99%
		T-DSC-Inception	88.51%	86.91%	87.54%
		T-DIPSA	88.03%	88.18%	88.76%
		T-DIPSA-F	90.12%	89.21%	90.12%
		T-DIPSA-FRAM	**91.96%**	**90.47%**	**91.03%**

知，T-DIPSA-F 模型在 UNSW-NB15 数据集和 ToN-IoT 数据集的准确率、召回率、F1 值分别提升了 0.88%、1.51%、0.99% 和 1.38%、0.46%、0.09%，精确率、加权召回率、加权 F1 值分别提升了 2.31%、2.28%、2.21% 和 1.48%、1.87%、2.13%。实验结果表明，加入 DSC-Inception-PSA 结构和特征预提取模块能显著提升模型在多分类任务中的性能，其原因在于这些模块能够增强模型对类别间复杂决策边界的敏感性。DSC-Inception-PSA 通过动态调整感受野，优化了模型对空间特征分布不均匀的捕捉能力，而特征预提取模块则通过特征降维，强化了模型对异常流量数据细粒度特征的表达能力。嵌入 RAM 后，高维特征预提取过程中少数类特征得到多尺度增强，残差机制有效缓解了梯度消失和特征丢失问题，进一步提升模型的学习效率和鲁棒性。最终，F1 值在 UNSW-NB15 数据集的二分类和多分类中达到 93.58% 和 88.26%，在 ToN-IoT 数据集的二分类和多分类中达到 95.35% 和 91.03%，充分证明了 T-DIPSA-FRAM 模型在入侵检测任务中的优异性能。

3.8　特征可视化

为更直观地验证此次模型的有效性，利用 t-分布随机邻居嵌入（t-distributed stochastic neighbor embedding，T-SNE）对在 UNSW-NB15 测试集数据所学的高层次特征向量进行可视化表示。二分类前后的特征向量可视化如图 12、图 13 所示。图中，Class 0 代表正常样本，Class 1 代表异常样本分布。由图 12、图 13 可看出，在分类前，正常和异常样本之间无明显界限，分类后，类别特征明显聚集为 2 簇，且类别界限较为清晰。该结果展现了本文方法在区分正常样本和异常样本方面的有效性。

四分类前后特征向量可视化如图 14、图 15 所示。针对多分类任务，Analysis（Class0）、Backdoor（Class 1）、Shellcode（Class 8）、Worms（Class 9）4 个少数类样本在分类前的各部分特征向量分布混乱，Analysis 与 Backdoor 明显重叠，且特征分布几乎一致。分类后，Shellcode 和 Worms 形成聚集簇，而 Analysis 与 Backdoor 重叠部分减少，相比原始分布有所改善。

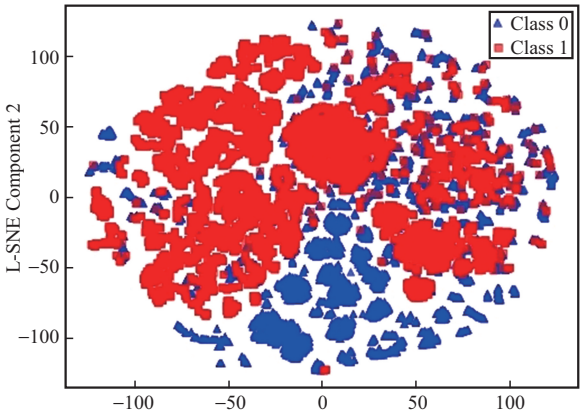

图 12　二分类前特征向量可视化

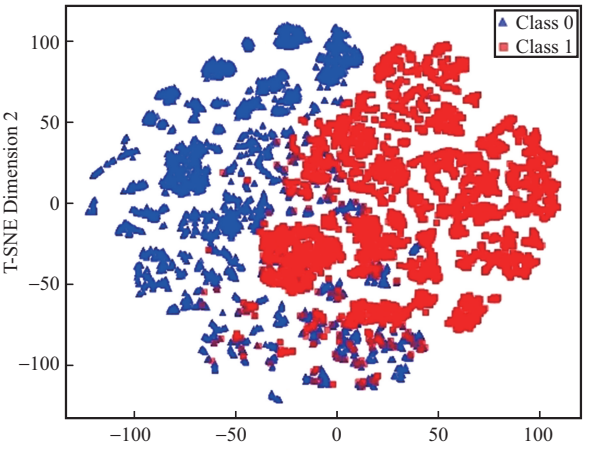

图 13　二分类后特征向量可视化

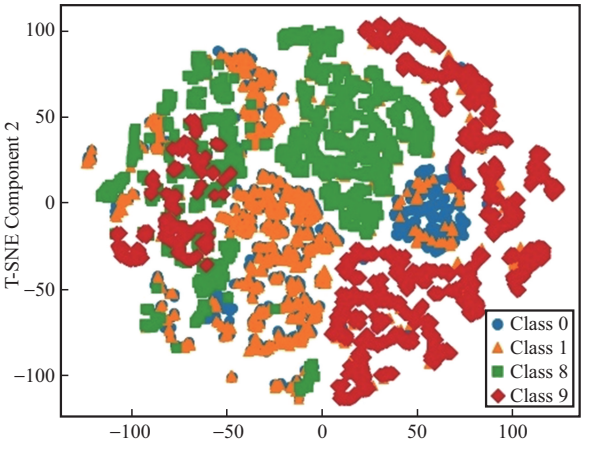

图 14　四分类前特征向量可视化

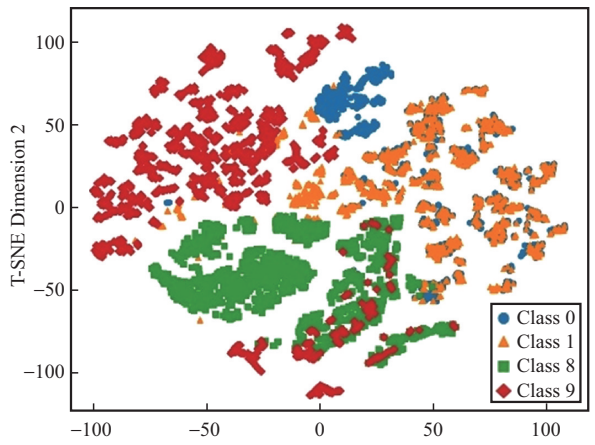

图15　四分类后特征向量可视化

3.9　对比实验

为验证本文所提方法的可靠性和有效性，选择深度神经网络（deep neural network，DNN）和现有的方法[26-32]进行对比实验。不同模型在UNSW-NB15数据集和ToN-Io T数据集的类别样本所得F1值如图16、图17所示。

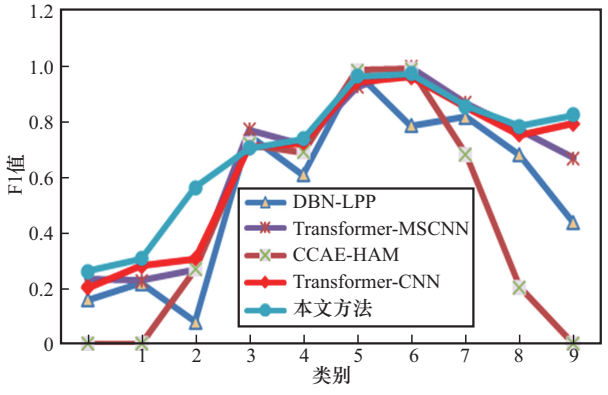

图16　各方法在UNSW-NB15数据集中的对比结果

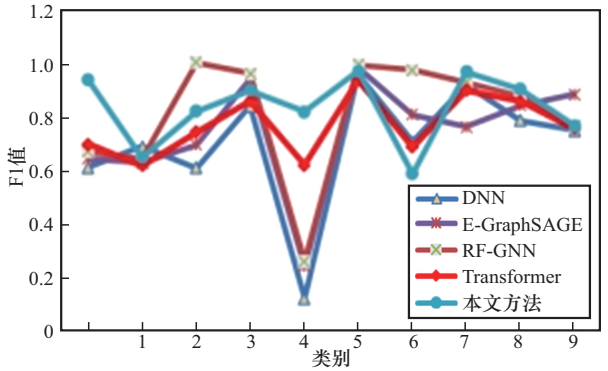

图17　各方法在ToN-IoT数据集中的对比结果

在UNSW-NB15数据集中，CCAE-HAM[26]通过卷积自编码器降低流量数据的冗余性，并采用头部注意力机制提取时序相关性信息，但难以捕捉细粒度特征信息，导致对Analysis（0）、Backdoor（1）、Reconnaissance（7）、Shellcode（8）、Worms（9）少数类样本区分度极差。DBN-LPP[27]通过深度置信网络（deep belief network，DBN）提取特征，并结合局部保持投影（locality preserving projection，LPP）去除冗余特征，但对时序特征建模能力有限。结合卷积网络的Transformer[28]提取少数类别细粒度特征能力不佳，区分效果差。Transformer-MSCNN[29]结合多尺度卷积与Transformer，并采用Borderline SMOTE过采样方法，其在少数类别Analysis（0）、Backdoor（1）、Shellcode（8）和Worms（9）的区分性有所提高。本文方法则是利用Transformer自注意力机制捕捉样本数据全局依赖，DSC-Inception模块提取多层次流量特征，融合残差注意力模块和金字塔注意力机制，极大增强了对流量特征的多尺度聚焦，在Analysis（0）、Backdoor（1）和Worms（9）达到更高的区分能力。

在ToN-IoT数据集中，可以看出DNN方法的表现最差。E-GraphSAGE[30]和RF-GNN[31]构建流量拓扑图，并有效聚合节点与边信息特征，能提高各类别的区分能力；Transformer[32]未多尺度提取类间细微特征，区分能力差；本文方法则借助ADASYN-RENN-LOF算法增强少数类样本，如Xss（3）、Mitm（4），利用残差注意力模块在特征预提取阶段有力收缩少数类别流量细微特征，Transformer模块融合DSC-Inception、金字塔注意力机制，强化构建不同类别流量序列高维特征，从而增强模型在类别间的泛化能力且区分性能较为平稳。

4　结束语

为提升网络入侵检测系统对复杂网络流量的

分类性能，本文提出一种结合特征预提取模块、残差注意力模块与 Transformer-DSC-Inception-PSA 的入侵检测模型 T-DIPSA-FRAM。特征预提取模块高效提取流量数据中的关键特征，残差注意力模块则有效增强了少数类特征的学习能力，Transformer-DSC-Inception 架构结合自注意力机制和 DSC-Inception 结构的优势，捕捉长距离依赖关系和多尺度特征，并借助 PSA 加强对流量数据空间特征的聚焦。最终 T-DIPSA-FRAM 在二分类和多分类任务中均取得了明显的性能提升。尽管该模型在流量特征学习、异常检测方面效果显著，但仍无法有效区分个别异常流量特征，且模型实际决策过程难以解释。未来研究可围绕提高模型可解释性、优化少数类检测，以及增强模型的自适应能力展开，以进一步提升模型的综合应用能力。

参考文献：

[1] 张昊, 张小雨, 张振友, 等. 基于深度学习的入侵检测模型综述[J]. 计算机工程与应用, 2022, 58(6): 17-28.
Zhang H, Zhang X Y, Zhang Z Y, et al. Summary of intrusion detection models based on deep learning[J]. Computer Engineering and Applications, 2022, 58(6): 17-28.

[2] He K, Kim D D, Asghar M R. Adversarial machine learning for network intrusion detection systems: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2023, 25(1): 538-566.

[3] Rashid M, Kamruzzaman J, Imam T, et al. A tree-based stacking ensemble technique with feature selection for network intrusion detection[J]. Applied Intelligence, 2022, 52(9): 9768-9781.

[4] Disha R A, Waheed S. Performance analysis of machine learning models for intrusion detection system using Gini impurity-based weighted random forest (GIWRF) feature selection technique[J]. Cybersecurity, 2022, 5: 1.

[5] Fu Y F, Du Y S, Cao Z J, et al. A deep learning model for network intrusion detection with imbalanced data[J]. Electronics, 2022, 11(6): 898.

[6] Naeem H, Cheng X C, Ullah F, et al. A deep convolutional neural network stacked ensemble for malware threat classification in Internet of things[J]. Journal of Circuits, Systems and Computers, 2022, 31(17): 2250302.

[7] 孙佳佳, 李承礼, 常德显, 等. 基于生成对抗网络的入侵检测类别不平衡问题数据增强方法[J]. 科学技术与工程, 2022, 22(18): 7965-7971.
Sun J J, Li C L, Chang D X, et al. Data augmentation method for intrusion detection imbalance problem using generative adversarial networks[J]. Science Technology and Engineering, 2022, 22(18): 7965-7971.

[8] Arafah M, Phillips I, Adnane A, et al. Anomaly-based network intrusion detection using denoising autoencoder and Wasserstein GAN synthetic attacks[J]. Applied Soft Computing, 2025, 168: 112455.

[9] Khan M A, Iqbal N, Jamil H, et al. An optimized ensemble prediction model using AutoML based on soft voting classifier for network intrusion detection[J]. Journal of Network and Computer Applications, 2023, 212: 103560.

[10] Sun Y, Que H K, Cai Q Q, et al. Borderline SMOTE algorithm and feature selection-based network anomalies detection strategy[J]. Energies, 2022, 15(13): 4751.

[11] Al-Shehari T, Kadrie M, Al-Mhiqani M N, et al. Comparative evaluation of data imbalance addressing techniques for CNN-based insider threat detection[J]. Scientific Reports, 2024, 14: 24715.

[12] Halbouni A, Gunawan T S, Habaebi M H, et al. CNN-LSTM: hybrid deep neural network for network intrusion detection system[J]. IEEE Access, 2022, 10: 99837-99849.

[13] Wu Z H, Zhang H, Wang P H, et al. RTIDS: a robust transformer-based approach for intrusion detection system[J]. IEEE Access, 2022, 10: 64375-64387.

[14] He J X, Wang X D, Song Y F, et al. A multiscale intrusion detection system based on pyramid depthwise separable convolution neural network[J]. Neurocomputing, 2023, 530: 48-59.

[15] Karthick Raghunath K M, Kumar V V, Venkatesan M, et al. XGBoost regression classifier (XRC) model for cyber attack detection and classification using inception V4[J]. Journal of Web Engineering, 2022, 21(4): 1295-1322.

[16] 谢金鑫. 基于深度学习模型的网络入侵检测研究[D]. 天津: 天津理工大学, 2022.
Xie J X. Research on network intrusion detection based on deep learning model[D]. Tianjin: Tianjin University of Technology, 2022.

[17] Yang K, Wang J M, Li M J. An improved intrusion detection method for IIoT using attention mechanisms, BiGRU, and Inception-CNN[J]. Scientific Reports, 2024, 14: 19339.

[18] Ding W P, Abdel-Basset M, Mohamed R. DeepAK-IoT: an effective deep learning model for cyberattack detection in IoT networks[J]. Information Sciences, 2023, 634: 157-171.

[19] Li B, Li Z, Yang Y L. Residual attention graph convolutional network for web services classification[J]. Neurocomputing, 2021, 440: 45-57.

[20] Islam S, Elmekki H, Elsebai A, et al. A comprehensive survey on applications of transformers for deep learning tasks[J]. Expert Systems with Applications, 2024, 241: 122666.

[21] 向思羽, 刘才铭. 结合混合特征选择和Transformer的网络数据流异常检测[J]. 电子科技大学学报, 2025, 54(3): 442-454.
Xiang S Y, Liu C M. Network data anomaly detection combined with hybrid feature selection and Transformer[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(3): 442-454.

[22] 李聪聪, 袁子龙, 滕桂法. 基于深度学习的时空特征融合网络入侵检测模型研究[J]. 信息安全研究, 2025, 11(2): 122-129.
Li C C, Yuan Z L, Teng G F. Research on deep learning-based spatio-temporal feature fusion network intrusion detection model[J]. Journal of Information Security Research, 2025, 11 (2): 122-129.

[23] Zhang H, Zu K K, Lu J, et al. EPSANet: an efficient pyramid squeeze attention block on convolutional neural network[C]//Proceeding of the Computer Vision-ACCV 2022: 16th Asian Conference on Computer Vision. Heidelberg: Springer, 2022: 541-557.

[24] Zoghi Z, Serpen G. UNSW-NB15 computer security dataset: Analysis through visualization[J]. Security and Privacy, 2024, 7 (1): e331.

[25] Tareq I, Elbagoury B M, El-Regaily S, et al. Analysis of ToN-IoT, UNW-NB15, and edge-IIoT datasets using DL in cybersecurity for IoT[J]. Applied Sciences, 2022, 12(19): 9572.

[26] 金志刚, 刘凯, 武晓栋. 堆叠循环卷积和头部注意力的工业互联网入侵检测[J]. 哈尔滨工业大学学报, 2024: 1-11.
Jin Z G, Liu K, Wu X D. Industrial Internet intrusion detection based on stacking circular convolution and head attention[J]. Journal of Harbin Institute of Technology, 2024: 1-11.

[27] 武玉坤, 李伟, 陈沅涛. 深度置信网络融合局部保持投影的入侵检测模型[J]. 计算机应用与软件, 2024, 41(6): 62-71.
Wu Y K, Li W, Chen Y T. Intrusion detection model based on deep belief network fusing locality preserving projection[J]. Computer Applications and Software, 2024, 41(6): 62-71.

[28] Kamal H, Mashaly M. Advanced hybrid transformer-CNN deep learning model for effective intrusion detection systems with class imbalance mitigation using resampling techniques[J]. Future Internet, 2024, 16(12): 481.

[29] 李井龙, 刘胜全, 马宇航, 等. 融合Transformer和MSCNN双分支架构的工控网络入侵检测研究[J]. 东北师大学报(自然科学版), 2024, 56(3): 70-78.
Li J L, Liu S Q, Ma Y H, et al. Integrating Transformer and MSCNN dual-branch architecture research on intrusion detection in industrial control networks[J]. Journal of Northeast Normal University (Natural Science Edition), 2024, 56(3): 70-78.

[30] Lo W W, Layeghy S, Sarhan M, et al. E-GraphSAGE: a graph neural network based intrusion detection system for IoT[C]//Proceedings of the NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium. Piscataway: IEEE Press, 2022: 1-9.

[31] 罗国宇, 汪学舜, 戴锦友. 物联网入侵检测的随机特征图神经网络模型[J]. 计算机工程与应用, 2024, 60(21): 264-273.
Luo G Y, Wang X S, Dai J Y. Random feature graph neural network for intrusion detection in Internet of Things[J]. Computer Engineering and Applications, 2024, 60(21): 264-273.

[32] Tseng S M, Wang Y Q, Wang Y C. Multi-class intrusion detection based on Transformer for IoT networks using CIC-IoT-2023 dataset[J]. Future Internet, 2024, 16(8): 284.

[作者简介]

张万琪（2000– ），男，中国人民公安大学信息网络安全学院硕士生，主要研究方向为网络信息安全。

王家兴（2000– ），男，中国人民公安大学信息网络安全学院硕士生，主要研究方向为自然语言处理、情感方面级分析等。

宋振峰（1980– ），男，博士，中国人民公安大学信息网络安全学院副教授，主要研究方向为警务信息技术、网络安全。