基于多尺度注意力特征增强的异常流量检测方法

杨宏宇[1,2]，张豪豪[1]，成翔[3]
（1.中国民航大学安全科学与工程学院，天津 300300；
2.中国民航大学计算机科学与技术学院，天津 300300；
3.扬州大学信息工程学院，江苏 扬州 225127）

摘　要：针对现有网络异常流量检测方法存在特征冗余以及流量序列的时间依赖性，导致模型训练速度慢和检测性能不佳等不足，提出一种基于多尺度注意力特征增强的异常流量检测方法。首先，通过基于动态分组的特征选择算法从流量数据中选出最优特征集合。其次，使用密集卷积神经网络和多尺度注意力特征提取网络分别提取流量数据的局部和全局特征。最后，利用特征增强网络增强局部和全局特征的区分度和整体表达的有效性，并采用加权融合的方法进行特征融合，实现异常流量检测。实验结果表明，所提方法在CIC-IDS2017和CSE-CIC-IDS2018数据集上的F1分数分别提升0.17%~2.75%、0.43%~8.99%，具有良好的检测效果。

关键词：异常流量检测；特征选择；多尺度注意力；特征增强网络
中图分类号：TP393
文献标志码：A
DOI：10.11959/j.issn.1000–436x.2024262

Abnormal traffic detection method based on multi-scale attention feature enhancement

YANG Hongyu[1,2], ZHANG Haohao[1], CHENG Xiang[3]
1. School of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China
2. School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China
3. School of Information Engineering, Yangzhou University, Yangzhou 225127, China

Abstract: To address feature redundancy and temporal dependencies in traffic data sequences that slow down model training and degrade performance of existing network abnormal traffic detection methods, an abnormal traffic detection method based on multi-scale attention feature enhancement was proposed. Firstly, an optimal feature set was selected from traffic data using a feature selection algorithm based on dynamic grouping. Secondly, Dense-CNN and a multi-scale attention feature extraction network were employed to extract local and global features of the traffic data. Finally, a feature enhancement network was used to increase the distinctiveness and expressiveness of local and global features, which were then fused using a weighted fusion approach to achieve abnormal traffic detection. Experimental results on the CIC-IDS2017 and CSE-CIC-IDS2018 datasets show that the proposed method improves F1 score by 0.17% to 2.75% and 0.43% to 8.99%, respectively, which has good detection performance.

Keywords: abnormal traffic detection, feature selection, multi-scale attention, feature enhancement network

收稿日期：2024–07–10；修回日期：2024–11–21
通信作者：杨宏宇，yhyxlx@hotmail.com
基金项目：国家自然科学基金民航联合研究基金重点资助项目（No.2433205）；国家自然科学基金资助项目（No.U1833107）；江苏省基础研究计划自然科学基金青年基金项目（No.BK20230558）
Foundation Items：Civil Aviation Joint Research Fund Project of the National Natural Science Foundation of China (No.2433205), The National Natural Science Foundation of China (No.U1833107), The Jiangsu Provincial Basic Research Program Natural Science Foundation - Youth Fund Project (No.BK20230558)

0　引言

随着互联网技术的迅速发展，网络服务已经深入各个领域，从基础设施到社交媒体，从商业交易到政府服务，网络应用无处不在。然而，在互联网环境中，用户行为、数据传输和服务请求通常伴随着潜在的网络攻击。异常流量检测通过对网络流量行为进行分析和监控，能够及时发现恶意攻击、网络入侵、数据泄露等安全威胁，从而维护互联网信息安全和业务系统的稳定运行。因此，网络异常流量检测研究具有重要的理论价值和现实意义。

目前，网络异常流量检测方法主要分为基于机器学习的异常流量检测方法和基于深度学习的异常流量检测方法[1]。基于机器学习的异常流量检测方法通过学习和分析正常流量的固有模式检测异常行为[2]，但在处理高维度特征和复杂数据结构的网络流量时，机器学习方法在特征提取和特征选择阶段计算成本较高，并且模型泛化能力较弱。近年来，随着深度学习技术的发展，研究人员开始使用卷积神经网络（CNN, convolutional neural network）、生成对抗网络（GAN, generative adversarial network）、长短期记忆网络（LSTM, long short-term memory）和图神经网络（GNN, graph neural network）等深度学习方法进行异常流量检测[3-4]，但在应对现实环境中的复杂网络状况时仍存在以下不足。

1) 网络流量数据的高维特征导致模型的计算负担加重，特征空间的庞大使检测模型难以捕捉到流量数据中的细微差异，从而影响异常流量检测的准确率。

2) 网络流量序列的时间依赖性以及长期的相互依赖关系导致异常流量检测变得更加复杂，现有检测方法通常忽略了时序数据的动态特性，从而导致检测精度不足。

3) 部分研究方法在提取流量特征时，生成的特征向量信息量不足，未能充分表示流量数据的复杂性和差异性，从而导致检测效果不佳。

针对上述不足，本文提出一种基于多尺度注意力特征增强的异常流量检测（MSAFE-ATD, abnormal traffic detection method based on multi-scale attention feature enhancement）方法。该方法首先通过基于动态分组的特征选择算法减少流量数据的冗余特征。然后，利用密集卷积神经网络和多尺度注意力特征提取网络分别提取流量数据的局部和全局特征。多尺度注意力特征提取网络通过设置不同的膨胀率，动态调整感受野大小，使检测模型能够同时捕捉短期内的异常波动和长期依赖的全局特征。不同膨胀率的空洞卷积分别捕捉流量数据中短期突变、周期性异常和短时趋势变化，以及全局特征和长期趋势变化，通过融合不同尺度的特征信息，提升对流量序列中细微特征和复杂结构的提取能力，从而解决现有方法因固定感受野导致特征提取不足的问题。此外，本文采用改进的混合注意力机制增强局部和全局特征，提升流量特征之间的差异性和整体表达效果，从而实现对异常流量的有效检测。本文主要工作和贡献如下。

1) 提出一种基于动态分组的特征选择算法，该算法利用对称不确定性（SU, symmetrical uncertainty）分别计算流量特征之间的冗余性和特征与标签类别之间的相关性，通过提出的评分函数和动态分组算法对特征进行排序和分组，并在每个分组中选出最优特征，从而缩短异常流量检测模型的训练时间和检测时间。

2) 提出一种多尺度注意力特征提取网络（MSAFEN, multi scale attention feature extraction network），多尺度注意力指在Transformer的注意力机制中引入不同膨胀率的空洞卷积，通过改变膨胀率调整感受野的大小，从而在不同时间尺度提取流量数据的全局特征。与现有特征提取方法相比，MSAFEN在不同尺度上同时关注流量数据短期内的特征变化和长期依赖的全局特征，能够在特征提取过程中将不同时间尺度的信息融合，有效解决现有检测方法忽视时序数据动态特性的局限。通过引入多尺度特征信息，MSAFEN在高维时序数据中实现了更全面的特征表达，有效提高了异常流量检测的精度。

3) 提出一种基于改进的混合注意力特征增强网络，包括基于特征压缩的位置注意力特征增强网络（PAMFC, position attention module of feature compression）和通道注意力特征增强网络（CAMFC, channel attention module of feature compression）。通过特征增强网络将流量数据中更广泛的上下文信息编码到局部和全局特征中，进一步增强对流量特征的表达能力，提高异常流量检测的精度。

1 相关工作

1.1 基于机器学习的异常流量检测方法

基于机器学习的异常流量检测方法通过特征工程和分类算法结合，以实现异常流量检测。Wu等[5]提出一种随机森林和合成少数类过采样（SMOTE, synthetic minority over sampling technique）结合的异常流量检测方法。该方法首先利用K-means聚类算法与SMOTE采样算法生成新的少数类样本以平衡数据集，从而提高对少数类攻击的检测性能。然后，通过随机森林优化决策树之间的相似度，并根据不同攻击类型的相似性矩阵调整分类结果。然而，该方法在数据预处理阶段的计算开销较大，难以适用于实时性要求较高的场景。在此基础上，Lu等[6]提出一种基于改进随机森林算法的异常流量检测方法，该方法通过引入秃鹰搜索算法和主成分分析方法对流量数据降维。然后，采用改进的随机森林分类器进行训练。与传统随机森林算法相比，该方法缩短了模型训练时间，但在面对未知攻击时的检测精度仍需进一步提高。

1.2 基于深度学习的异常流量检测方法

为了应对日益复杂的网络攻击，基于深度学习的异常流量检测方法逐渐成为研究热点。Hou等[7]提出一种基于GAN的异常流量检测（MTDGSE, malicious traffic detection based on GAN sample enhancement）方法，通过GAN生成攻击样本并利用CatBoost算法实现异常流量检测，但该方法在模型训练过程中存在梯度消失的问题。Li等[8]提出一种基于条件生成对抗网络的异常流量检测（CGAN-IDS, conditional generative adversarial network for intrusion detection）方法，通过条件生成对抗网络生成攻击样本，同时将编码器嵌入判别器中，提高生成器生成攻击样本的质量。然而，生成样本的质量依赖于原始数据集，如果恶意样本不足，生成的攻击样本将缺乏多样化和代表性，从而降低异常流量的检测性能。

为了有效处理流量数据的高维特征，Wu等[9]提出一种基于Transformer的异常流量检测（RTIDS, robust transformer based for intrusion detection system）方法，通过Transformer的编码器和解码器将流量数据映射到低维空间中，并采用多头注意力机制提取网络流量特征之间的上下文信息，以检测网络流量中的异常行为。在此基础上，Luo等[10]提出一种基于多通道对比学习网络的入侵检测方法（MCLDM, multi-channel contrastive learning network based intrusion detection method）。该方法采用自编码器对流量数据进行特征重构，并结合多通道对比学习提取流量数据的深层次特征。上述检测方法在使用编码器对流量特征重构时，未能充分保留原始流量数据的全局特征信息，导致检测效果不佳。

为了减少流量数据的冗余特征，提高异常流量的检测效率，Kanna等[11]提出一种基于特征选择和LSTM的入侵检测方法（HID-MCLSTM, hybrid intrusion detection using mapreduce based black widow optimized convolutional long short-term memory neural network），通过特征选择算法实现特征降维，并使用基于MapReduce的参数优化算法对LSTM网络参数进行优化。但在参数优化过程的计算资源开销过大，影响异常流量的检测效率。Bhardwaj等[12]提出一种基于增强型神经网络的异常流量检测方法（ENFM, enhanced neural network-based attack investigation framework for network forensics），通过一维卷积神经网络提取流量数据的时序和局部特征，但该网络无法有效提取长时间依赖性的时序特征。为解决该问题，Liu等[13]提出一种分层注意力机制的异常流量检测方法（HAGRU, hierarchical attention gated recurrent unit）。该方法采用双向门控循环单元提取流量数据的时序特征，并使用分层注意力机制为时序特征分配不同权重，以提高检测效果。但该方法未考虑流量数据不平衡的问题，导致对样本较少的攻击类型检测效果较差。Wang等[14]提出一种结合ResNet、Transformer和BiLSTM的异常流量检测方法（Res-TranBiLSTM, an intelligent approach for intrusion detection in the Internet of things），并行提取网络流量的空间和时间特征，但该方法需要集成多个深度学习模型提取流量特征，模型复杂度较高，不利于实时检测。

随着网络环境中的拓扑结构日益复杂，GNN逐渐应用于异常流量检测。Pujol-Perich等[15]提出一种基于图神经网络的异常流量检测方法（GNN-NIDS, unveiling the potential of graph neural networks for robust intrusion detection），通过GNN分析网络流量的拓扑结构实现异常检测。Lo等[16]提出一种基于GNN的入侵检测方法E-GraphSAGE，用于检测物联网环境下的异常流量。该方法将物联网设备之间的通信数据构建为图结构，采用GNN

分析设备之间的拓扑特征和时序依赖性。然而，随着物联网节点数量的增加，图结构的复杂性也随之增加，导致计算资源和时间的开销增加。Reka等[17]提出一种基于多头自注意力图卷积网络的异常流量检测方法（MSA-GCNN, multi head self-attention gated graph convolutional network based multi-attack intrusion detection）。该方法在图卷积网络中引入多头自注意力，增强了检测模型在识别多种攻击类型时的特征提取能力，但在处理动态网络拓扑和多种复杂攻击时计算资源消耗较大。

2　检测方法框架

目前，现有异常流量检测方法在处理复杂网络流量时，未能充分提取流量数据的局部特征和全局特征。同时，在特征提取过程中通常仅从单一尺度提取特征信息，忽略了流量序列的时间依赖性和长期相互依赖关系，未能充分考虑流量数据在不同尺度下的特征差异性，导致提取的特征向量信息量不足。为提高检测性能，本文面向流量数据的局部和全局特征，提出一种基于多尺度注意力特征增强的异常流量检测方法，该方法框架如图1所示。

基于多尺度注意力特征增强的异常流量检测方法由样本预处理、基于动态分组的特征选择算法、双分支特征提取、基于PAMFC和CAMFC的特征增强网络与异常流量检测4个部分组成。各部分功能设计如下。

1) 样本预处理。首先，对异常流量数据集进行数据清洗，去除重复、缺失和异常数据，以确保数据的质量和完整性。其次，针对流量样本数量较少的类型进行样本增强，以增加样本的多样性。最后，将流量样本中非数值类型特征转换为数值形式，并将转换后的特征进行标准化处理。

2) 基于动态分组的特征选择算法。首先，采用SU分别计算流量特征之间的冗余性以及流量特征与标签类别之间的相关性。其次，基于本文所提出的评分函数，对特征进行重要性排序。最后，利用动态分组算法从每个分组中选出最优特征，构成最优特征集合。

3) 双分支特征提取。将样本预处理和特征选择后的流量数据分别输入密集卷积神经网络（Dense-CNN, dense convolutional neural network）和多尺度注意力特征提取网络，分别提取流量数据的局部和全局特征。

4) 基于PAMFC和CAMFC的特征增强网络与异常流量检测。首先，使用基于PAMFC和CAMFC的特征增强网络对提取的局部特征和全局特征进行增强和优化。然后，将增强后的流量特征输入异常流量检测网络，利用全连接层和分类函数生成检测概率，从而实现异常流量检测。

3　基于动态分组的特征选择算法

为了提高异常流量检测方法中特征选择算法

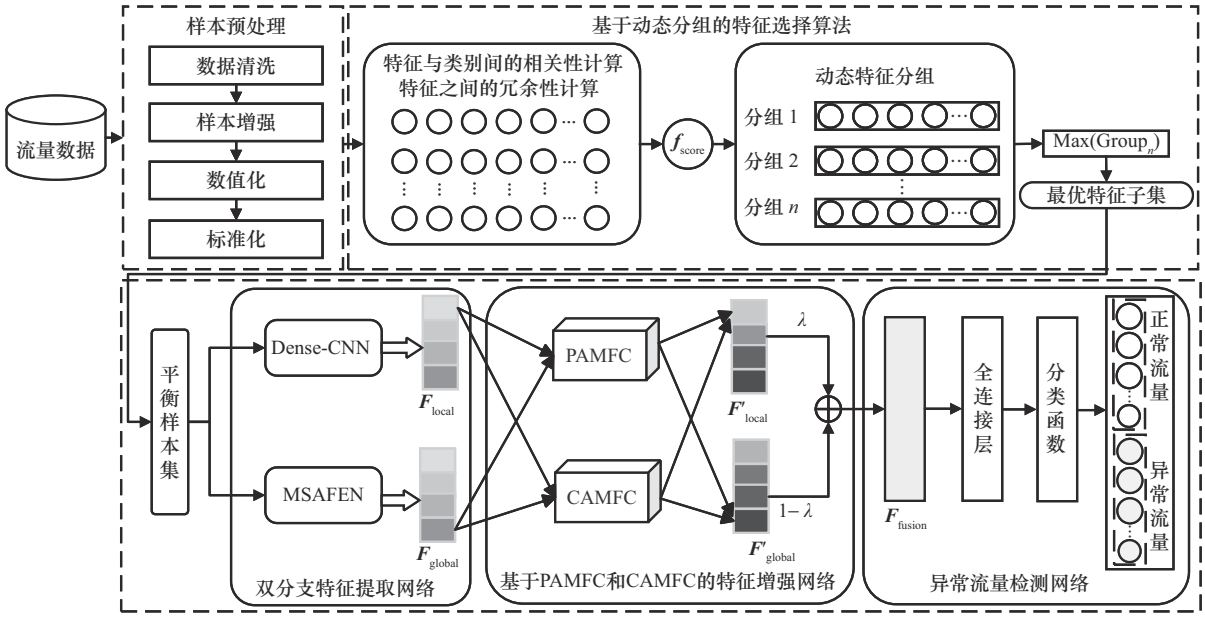


图1　基于多尺度注意力特征增强的异常流量检测方法框架

的有效性，本文提出一种基于动态分组的特征选择算法，如算法 1 所示。该算法在考虑已选特征与标签类别之间的相关性的基础上，同时分析候选特征、已选特征与标签类别之间的复杂关系。首先，采用对称不确定性计算流量数据集中特征 a_i 之间的冗余性以及特征 a_i 与标签类别 c 之间的相关性。其次，根据提出的评分函数 f_{score} 对特征进行重要性排序。最后，通过动态特征分组算法，从不同特征分组内选择最优特征作为最优特征集合。

算法 1 基于动态分组的特征选择算法

输入 原始特征集 D，特征数 K，标签类别 c，流量总数 n

输出 最优特征子集 S

1) $S\leftarrow\varnothing$, $A\leftarrow\varnothing$, $k\leftarrow 0$, $i\leftarrow 0$
2) for i in n do
3) 计算特征 a_i 与类别 c 间的相关性 $S(a_i, c)$
4) 计算特征 a_i 与 a_j 之间的冗余性 $S(a_i, a_j)$
5) $A \leftarrow \text{Rank}[f_{\text{score}}]$
6) i++
7) end for
8) 动态特征分组：
9) while $k<K$ do
10) $t_1\leftarrow A[1]$
11) $t_{k+1}\leftarrow \text{Max}[S(a_1, a_i)]$
12) $q_0\leftarrow P$
13) $R\leftarrow A\backslash S$
14) $\forall f_i\in R$, 计算 $q_i\leftarrow P(f_i)$
15) if $P> q_0$ then
16) $G_k\leftarrow t_i$
17) end if
18) $S\leftarrow S\cup\{f_{11}, f_{21}, \cdots, f_{m1}\}$
19) k++
20) end while

3.1 基于对称不确定性的特征相关-冗余性计算方法

本文采用对称不确定性分别计算网络流量数据集中特征之间的相关性、特征与标签类别之间的冗余性。基于对称不确定性的特征相关-冗余性计算方法通过对互信息（MI, mutual information）进行归一化处理，将特征与类别标签之间的相关性范围限定在 0 到 1 之间[18]，有效避免互信息在评估特征时倾向于选择多值特征的问题，从而能够全面地衡量特征集合对于流量分类的整体贡献。

流量特征 a_i 与标签类别 c 之间的相关性计算如式(1)所示，$S(a_i, c)$值越大表示相关性越大。

$$S(a_i,c) = 2\frac{I(a_i; c)}{H(a_i) + H(c)}, 0 \leqslant S(a_i,c) \leqslant 1 \quad (1)$$

其中，$I(a_i; c)$表示特征 a_i 与标签类别 c 之间的互信息，$H(a_i)$和 $H(c)$分别表示特征 a_i 和标签类别 c 的信息熵。

特征 a_i 与 a_j 之间的冗余性计算如式(2)所示，$S(a_i, a_j)$值越大表示特征间的冗余性越大。

$$S(a_i,a_j) = 2\frac{I(a_i; a_j)}{H(a_i) + H(a_j)}, 0 \leqslant S(a_i,a_j) \leqslant 1 \quad (2)$$

其中，$I(a_i; a_j)$表示特征间的互信息。

3.2 评分函数

为了综合评估流量特征与标签类别之间的相关性以及特征之间的冗余性，解决传统特征选择算法无法有效去除冗余特征的问题。本文设计了一个评分函数 f_{score}（如式(3)所示），通过计算流量数据集中每个特征的 f_{score} 值，实现对特征的排序。

$$f_{\text{score}} = S(a_i,c) - \alpha\sum_{f'\neq f} S(a_i,a_j) \quad (3)$$

其中，α 是平衡系数，用于调整相关性和冗余性在评分中的重要性。

3.3 动态分组算法

本文采用动态分组算法将排序后的特征进行分组，该算法在综合评估流量特征与类别标签之间的相关性的基础上，进一步考虑特征之间的冗余性。通过引入动态特征分组机制，能够在特征选择过程中有效减少冗余特征的影响，从而增强所选特征集的整体代表性与有效性。具体分组算法设计如下。

1) 特征分组初始化。首先，将特征排序中居首位的特征 a_1 分配至分组 t_1 中。然后，计算 a_1 与其他特征 a_i 之间的对称不确定性 SU，并将 SU 值最大的特征放入 t_1 中。最后，计算分组标准 P（如式(6)所示）值，并将其记作 q_0。

$$S_T = \sum_{a_i\in T} S(a_i,c) \quad (4)$$

$$R_T = \sum_{a_i,a_j\in T} S(a_i,a_j) \quad (5)$$

$$P = \frac{S_T}{R_T} \quad (6)$$

其中，S_T 表示特征分组 T 与标签类别 c 之间的相关性，R_T 表示特征分组 T 内特征 a_i 与 a_j 的冗余性。

2) 动态分组。对于剩余特征，选取当前排序

中位于首位的特征加入现有分组中，并重新计算该分组的P值。若新的P值小于q_0，则表明新加入的特征与分组内已有特征存在较高的冗余性，应继续加入下一特征并更新P值。循环进行该过程，直至分组的P值大于q_0。此时，该特征分组即第一个分组。对未分组的特征重复上述步骤，以形成新的特征组，依次类推，直至所有的特征均被分组。

3) 构建最优特征集合。从特征分组集$T=\{t_1, t_2, t_3,\cdots, t_i\}$的每个特征组$t_i$中选择排在首位的特征，构成最优特征集合。

4　双分支特征提取

4.1　基于 Dense-CNN 的局部特征提取

在网络流量分析中，不同网络层的特征对应不同的流量模式和行为。为了深入挖掘不同层级局部特征的细微特性，本文采用 Dense-CNN 提取流量数据的局部语义特征[19]。与传统 CNN 相比，Dense-CNN 在处理网络流量数据时能够更快地聚焦于关键局部特征。首先，Dense-CNN 的每一层在接收输入时，并非仅接收前一层的输出，而是接收之前所有层的输出，每层的输出都包含之前所有层输出的特征信息，使其能够直接利用前面所有层提取的特征信息，实现局部特征的跨层传递和特征的最大化共享，确保模型在不同层次上提取到的流量特征能够相互补充和加强。其次，由于 Dense-CNN 每一层直接连接到损失函数，在反向传播过程中能够有效缓解梯度消失问题，提高模型的训练效率和稳定性。最后，Dense-CNN 通过共享特征的方式，无须重复学习相似的特征信息，从而能够有效减少模型的参数数量。基于 Dense-CNN 的局部特征提取过程如图 2 所示，具体方法设计如下。

1) 卷积。将经过特征选择后的流量数据输入卷积层进行运算，该卷积层通过卷积核进行卷积操作，从输入数据中提取特征信息。

2) Dense-CNN 的多层特征复用。将特征向量输入密集块（Dense Block），每个 Dense Block 由 5 个密集层（DenseLayer）组成，每层均使用 3×3 的卷积核，并保持 64 个输出通道数。具体计算过程如下

$$\boldsymbol{f}_i = \mathrm{ReLU}\left(\mathrm{BN}\left(\boldsymbol{W}_i \mathrm{Concat}\left(\boldsymbol{f}_0,\cdots \boldsymbol{f}_{i-1}\right) + \boldsymbol{b}_i\right)\right) \quad (7)$$

$$\boldsymbol{F}_{\mathrm{local}} = \mathrm{Linear}\left(\mathrm{Concat}\left(f_0, f_1,\cdots f_{i-1}\right)\right) \quad (8)$$

其中，$\boldsymbol{f}_i$表示第i层的输出，$\boldsymbol{W}_i$和$\boldsymbol{b}_i$分别表示 Dense Block 第i层的权重矩阵和偏置矩阵，$\boldsymbol{F}_{\mathrm{local}}$表示输出的局部特征向量。

4.2　基于 MSAFEN 的全局特征提取

由于网络流量数据通常包含长序列信息，Transformer 在处理长序列数据时计算成本显著增加，并且难以有效地捕捉长距离依赖关系[20]。为深入提取不同时间尺度的全局流量特征，本文提出一种基于 MSAFEN 的全局特征提取网络。该网络在 Transformer 中引入多尺度扩展注意力（MSDA, multi-scale dilated attention）机制，通过在 MSDA 中设置不同膨胀率的空洞卷积增大卷积核的感受野。首先，通过不同膨胀率的空洞卷积层提取不同语义层次上的特征信息，使模型在不同尺度上能够更广泛地关注流量上下文信息，从而有效地捕捉流量数据中长距离依赖关系。其次，由于网络流量数据具有高维度和复杂的时序结构，传统方法在从不同尺度提取流量特征时，通常需要多个独立的模型或多层次的特征融合，而 MSAFEN 通过引入不同膨胀率的空洞卷积，从不同尺度上提取全局特征，以确保从不同尺度提取的流量全局特征具有多样性和互补性。基于 MSAFEN 的全局特征提取过程如图 3 所示，具体方法设计如下。

首先，在多尺度注意力特征提取网络中将流量数据$X \in R^{C\times H\times W}$输入 1×1 卷积层进行初始化。

其次，在 MSDA 中使用膨胀率分别为 1、2、3 的空洞卷积计算不同尺度下的查询向量$\boldsymbol{Q}$。其中，膨胀率为 1 的卷积操作中仅作用于紧邻元素，能够提取流量数据中短时间内的局部异常特征；膨胀率为 2 的卷积核具有更广泛的感受野，能够有效检测周期性的小规模异常流量和短期流量模式中的异常，从而提升对流量数据短时动态变化的识别能力；膨胀率为 3 的卷积核进一步扩大感

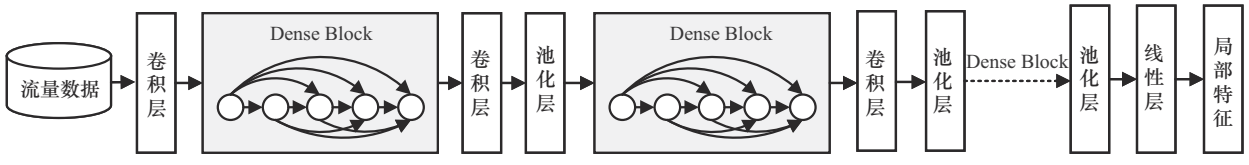


图 2　基于 Dense-CNN 的局部特征提取过程

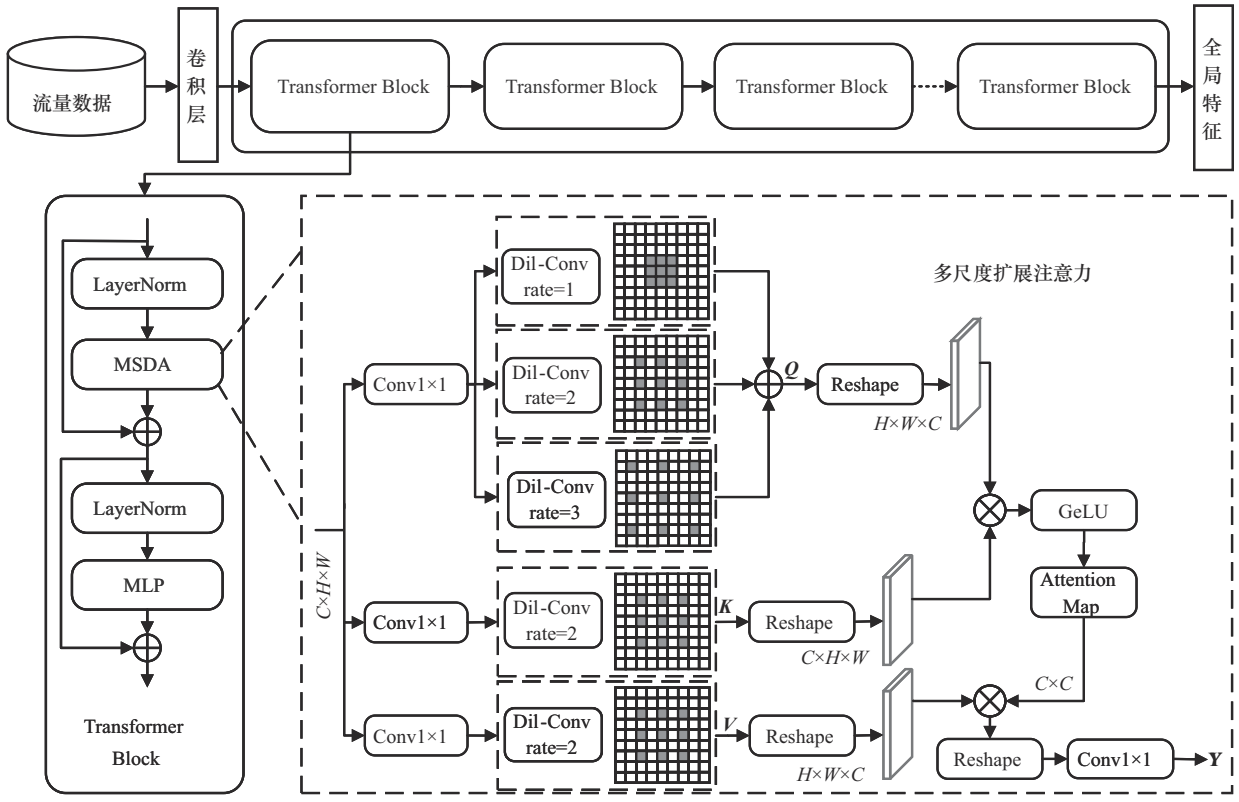


图3　基于MSAFEN的全局特征提取过程

受野，以提取大规模异常流量的持续性特征和长期流量趋势变化的全局特征。然而，当膨胀率过大时，模型关注区域过于分散，导致对局部和短时变化特征的提取能力降低[21]。因此，本文将膨胀率的最大值设置为3。最后，通过加权求和的方式将不同尺度下的查询向量$\boldsymbol{Q}$进行融合，从而增强特征之间的互补性和特征表达的多样性，其计算式如(9)所示。

$$\boldsymbol{Q}(i)=\sum_{l=1}^{L}\boldsymbol{f}(i+rl)g(l) \tag{9}$$

其中，$\boldsymbol{f}(i)$为输入向量，$\boldsymbol{Q}(i)$为输出向量，$g(l)$为长度为l的卷积核，r为$\boldsymbol{f}(i)$进行采样的空洞率。

然后，利用多尺度数据计算不同位置之间的查询向量$\boldsymbol{Q}$和键向量$\boldsymbol{K}$的相似度，但并非所有的查询向量$\boldsymbol{Q}$都与键向量$\boldsymbol{K}$相关，因此使用全部查询向量$\boldsymbol{Q}$和键向量$\boldsymbol{K}$之间的相似关系并不能有效地对全局流量特征重构。为解决该问题，本文使用GeLU激活函数代替Softmax，GeLU函数能够减少流量数据中无关特征的影响，同时保证注意力权重的稀疏性。

最后，通过Reshape函数分别得到$\boldsymbol{Q}\in R^{HW\times C}$、$\boldsymbol{K}\in R^{HW\times C}$和$\boldsymbol{V}\in R^{HW\times C}$，再通过注意力机制的计算，得到流量数据的全局特征向量$\boldsymbol{F}_{\text{global}}$，具体计算过程如下

$$\boldsymbol{x}_{i,j}=\text{GELU}\left(\frac{\boldsymbol{Q}^{\text{T}}\boldsymbol{K}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{10}$$

$$\boldsymbol{h}_i=\text{MSDA}(\text{Norm}(\boldsymbol{x}))+\boldsymbol{x} \tag{11}$$

$$\boldsymbol{H}=\text{Linear}(\text{Concat}(h_1,h_2,\cdots,h_n)) \tag{12}$$

$$\boldsymbol{F}_{\text{global}}=\text{MLP}(\text{Norm}(\boldsymbol{H}))+\boldsymbol{H} \tag{13}$$

其中，$\boldsymbol{x}_{i,j}$表示输入流量特征向量$\boldsymbol{X}$经过注意力计算的全局特征分量，d_k表示键向量维度，$\boldsymbol{h}_i$表示多头扩展注意力的输入结果，$\boldsymbol{H}$表示特征向量$\boldsymbol{h}_i$拼接后经过线性层的输出结果。

5　基于PAMFC和CAMFC的特征增强网络与异常流量检测

在流量数据的局部和全局特征信息中，部分关键信息未能充分表示，同时冗余特征信息会影响异常流量检测的准确性。为了增强检测模型对局部和全局特征的捕捉能力，并抑制冗余特征信息的干扰，本文提出一种基于PAMFC和CAMFC的特征增强网络，对提取后的流量局部和全局特征进行增强，进一步提高流量特征的表达能力，提高异常流

量检测的精度。

在异常流量检测网络部分，为了充分利用局部和全局特征的互补性，将经过PAMFC和CAMFC增强后的全局和局部特征进行加权融合，融合后的特征向量同时包含流量数据的全局信息和局部信息；再将融合后的特征向量通过全连接层和Softmax函数，实现异常流量检测。

5.1　基于PAMFC的特征增强网络

在原始位置注意力机制中，通过特征向量间的内积运算获取不同向量之间的关系。然而，在高维网络流量数据中，直接计算特征向量之间的关系会导致巨大的计算量和内存消耗。因此，本文在PAMFC中采用特征图降维策略，将提取的特征向量映射到低维的空间中，并计算低维特征向量与位置特征中心的相关性，动态地调整不同位置的注意力权重，使模型能够自动聚焦于对异常流量检测贡献最大的空间区域。该特征压缩机制有效减少了计算和内存开销，同时突出局部和全局特征的关键位置信息，抑制冗余特征信息的干扰，以实现对关键特征的有效增强。基于PAMFC的特征增强网络如图4所示，具体方法设计如下。

首先，对流量的全局和局部特征进行卷积操作，获得特征图$\boldsymbol{Y}_1 \in R^{C\times H\times W}$，再通过由1×1、2×2和3×3不同大小的池化层组成的特征中心对特征图$\boldsymbol{Y}_1$进行处理，获得3个不同大小的特征图。为了便于后续处理，将特征图的维度调整为$C\times L^2$，并将池化后的特征图进行拼接，得到特征中心特征图$\boldsymbol{P}\in R^{C\times M}$，其中$\boldsymbol{P}$的第二维度之和被标记为$M$。

其次，为了从特征图$\boldsymbol{Y}_1$中提取关键位置信息，将其输入1×1的卷积层，连接一个全连接层，从而得到特征图$\boldsymbol{G}\in R^{\hat{C}\times Z}$，并对特征中心特征图$\boldsymbol{P}$进行同样的操作得到$\boldsymbol{N}\in R^{\hat{C}\times M}$。

再次，在获得特征图$\boldsymbol{G}$和$\boldsymbol{N}$后，计算特征图$\boldsymbol{G}$的转置与特征图$\boldsymbol{N}$的矩阵乘法，生成位置注意力权重$\boldsymbol{L}\in R^{Z\times M}$。为了确保权重图的有效性，采用Softmax函数对其进行标准化处理。标准化处理后的位置注意力权重图$\boldsymbol{l}$的计算如式(14)所示。

$$\boldsymbol{l}_{j,i}=\frac{\exp\left(\boldsymbol{G}_j^{\mathrm{T}}\boldsymbol{N}_i\right)}{\sum_{i=1}^{M}\exp\left(\boldsymbol{G}_j^{\mathrm{T}}\boldsymbol{N}_i\right)} \tag{14}$$

其中，$\boldsymbol{l}_{j,i}$为第i个位置特征中心与第j个特征之间的关系，反映位置信息在局部和全局特征中的重要性；$\boldsymbol{G}_j$表示第j个位置特征图中的特征向量，$\boldsymbol{G}_j^{\mathrm{T}}$表示特征图$\boldsymbol{G}_j$的转置。

最后，将特征中心特征图$\boldsymbol{P}$输入全连接层得到特征图$\boldsymbol{Z}\in R^{C\times M}$，通过式(15)得到最终输出特征图$\hat{\boldsymbol{Y}}_1$。通过这种方式，PAMFC能够精确地捕捉数据中的关键位置信息，有效地提高计算效率。

$$\hat{\boldsymbol{Y}}_1=\mu\sum_{i=1}^{M}\left(\boldsymbol{l}_{j,i}\boldsymbol{Z}_i\right)+\boldsymbol{Y}_{1j} \tag{15}$$

其中，μ表示位置注意力系数，用于衡量不同位置特征在生成输出特征图$\hat{\boldsymbol{Y}}_1$的重要性；$\boldsymbol{Y}_{1j}$表示特征图$\boldsymbol{Y}_1$中的第j个特征向量。

5.2　基于CAMFC的特征增强网络

在原始通道注意力机制中，为了获取通道注意力权重，通常需要计算各个通道之间的相关性。然而，当特征图通道数较大时，这种计算方式会显著增加计算复杂度。因此，本文在CAMFC中采用特征图降维策略，将每个通道视为一个独立的特征单元。首先，对每个通道进行聚合，形成代表该通道整体特性的特征中心。然后，计算低维空间中每个通道与通道特征中心的相关性，获取通道注意力权重，并通过计算得到的相关性动态调整每个通道的权重。该特征压缩机制有效地减少了计算复杂度，同时使模型能够更直接地关注对整体流量特征贡献较大的通道，以实现特征的局部与全局增强。基于CAMFC的特征增强网络如图5所示，具体方法设计如下。

首先，对流量的局部和全局特征进行卷积操作，得到特征图$\boldsymbol{Y}_2\in R^{C\times H\times W}$，再通过1×1的卷积层对其通道维度进行降维处理，得到特征图$\boldsymbol{D}\in$

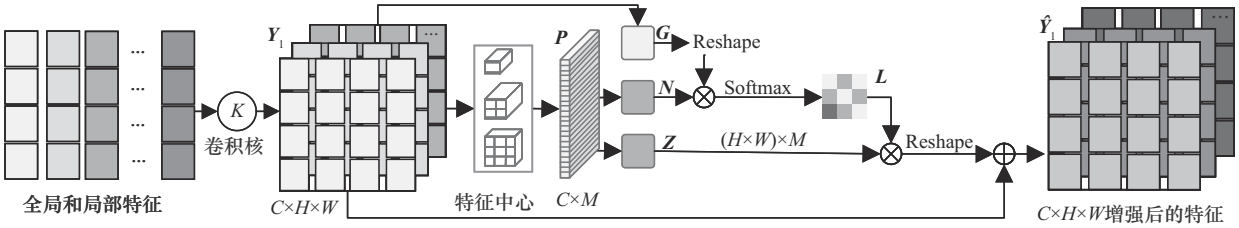


图4　基于PAMFC的特征增强网络

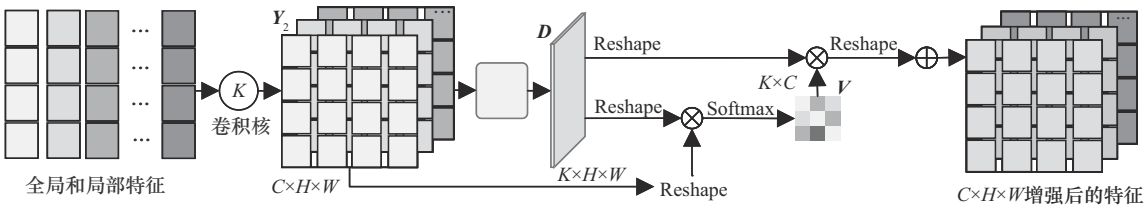


图5　基于CAMFC的特征增强网络

$R^{C\times H\times W}$，C为通过1×1卷积层降维后的通道数。特征图$\boldsymbol{D}$的每个通道映射视为一个通道特征中心。

然后，通过式(16)计算特征图$\boldsymbol{D}$中每个通道的重要性权重$\boldsymbol{V}\in R^{C\times K}$，该过程与PAMFC操作相同，通过一个简化的注意力分配实现该过程，使用1×1的卷积层生成通道重要性权重图，权重图经过Softmax处理，得到各通道的标准化注意力权重$\boldsymbol{V}$。

$$\boldsymbol{v}_{j,i}=\frac{\exp\left(\boldsymbol{Y}_j^{\mathrm{T}}\boldsymbol{D}_i\right)}{\sum_{i=1}^{K}\exp\left(\boldsymbol{Y}_j^{\mathrm{T}}\boldsymbol{D}_i\right)} \tag{16}$$

其中，$\boldsymbol{v}_{j,i}$为第i个通道特征中心与第j个通道之间的关系，反映通道信息在局部和全局特征中的重要性；$\boldsymbol{Y}_j$表示第j个通道特征图中的特征向量，$\boldsymbol{Y}_j^{\mathrm{T}}$表示特征图$\boldsymbol{Y}_j$的转置。

最后，通过式(17)计算得到最终的输出特征图$\hat{\boldsymbol{Y}}_2$。

$$\hat{\boldsymbol{Y}}_2=\beta\sum_{i=1}^{M}\left(\boldsymbol{v}_{j,i}\boldsymbol{D}_i\right)+\boldsymbol{Y}_{2j} \tag{17}$$

其中，β表示通道注意力系数，用于衡量不同通道特征在生成输出特征图$\hat{\boldsymbol{Y}}_2$的重要性；$\boldsymbol{Y}_{2j}$表示特征图$\boldsymbol{Y}_2$中的第j个特征向量。

5.3　异常流量检测

异常流量检测网络由特征融合层、全连接层和Softmax函数组成。首先，将增强后的局部特征和全局特征输入特征融合层进行加权融合。然后，利用全连接神经网络层将提取的特征映射到异常流量的类别标记空间。最后，通过Softmax函数计算异常流量的检测概率。异常流量检测网络如图6所示，具体方法设计如下。

1) 加权特征融合。为充分利用局部和全局特征的互补性，提高异常流量检测的准确率，将经过PAMFC和CAMFC增强后的全局和局部特征进行加权融合，生成融合特征向量$\boldsymbol{F}_{\mathrm{fusion}}$，其计算式为

$$\boldsymbol{F}_{\mathrm{fusion}}=\lambda\boldsymbol{F}'_{\mathrm{local}}+(1-\lambda)\boldsymbol{F}'_{\mathrm{global}} \tag{18}$$

其中，λ为加权融合系数，$\boldsymbol{F}'_{\mathrm{local}}$和$\boldsymbol{F}'_{\mathrm{global}}$分别为增强后的局部和全局特征。

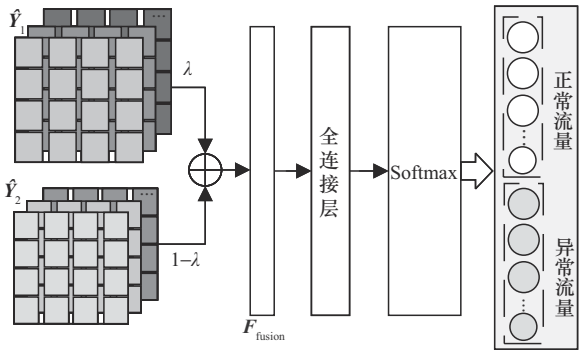


图6　异常流量检测网络

2) 异常流量分类。将融合后的特征向量$\boldsymbol{F}_{\mathrm{fusion}}$输入全连接层，并使用ReLU激活函数提取非线性特征，得到特征向量$\boldsymbol{z}$。此时，特征向量$\boldsymbol{z}$具有局部和全局特征的有效信息，同时包含经过非线性变换后提取的关键特征。最后，将经过ReLU激活函数的特征向量输入Softmax函数，从而得到异常流量的分类概率，其计算式为

$$\boldsymbol{z}=\mathrm{ReLU}\left(\boldsymbol{W}\boldsymbol{F}_{\mathrm{fusion}}+\boldsymbol{b}\right) \tag{19}$$

$$\mathrm{Softmax}\left(z\right)=\frac{\mathrm{e}^{z}}{\sum_{j=1}^{S}\mathrm{e}^{z_j}} \tag{20}$$

其中，$\boldsymbol{W}$和$\boldsymbol{b}$分别表示检测网络中全连接层的权重矩阵和偏置项，S为异常流量类别的数量。

6　实验与结果分析

6.1　实验数据集

CIC-IDS2017数据集[22]包含大量来自真实网络环境的网络流量数据，该数据集涵盖正常流量和多种类型的网络攻击流量，具体样本分布如表1所示。

CSE-CIC-IDS2018数据集[23]记录了10天的网络流量数据，包含7种不同场景下的攻击类型，如分布式拒绝服务攻击、僵尸网络和跨站脚本攻击等，具体样本分布如表2所示。

表 1　CIC-IDS2017 数据集样本分布

攻击类型	样本数量/个	训练集/个	测试集/个
Benign	2 273 097	1 818 478	454 619
DoS Hulk	231 073	184 858	46 215
PortScan	158 930	127 144	31 786
DDoS	128 027	102 422	25 605
DoS GoldenEye	10 293	8 234	2 059
FTP-Patator	7 938	6 350	1 588
SSH-Patator	5 897	4 718	1 179
DoS Slowloris	5 796	4 637	1 159
DoS Slowhttptest	5 499	4 399	1 100
Bot	1 966	1 573	393
Brute Force	1 507	1 206	301
XSS	652	522	130
Infiltration	36	29	7
SQL Injection	21	17	4
Heartbleed	11	9	2
总计	2 830 743	2 264 596	566 147

表 2　CSE-CIC-IDS2018 数据集样本分布

攻击类型	样本数量/个	训练集/个	测试集/个
Benign	2 247 402	1 797 922	449 480
DDoS attack-HOIC	686 012	548 810	137 202
DDoS LOIC-HTTP	576 191	460 953	115 238
DoS attacks-Hulk	461 912	369 530	92 382
Bot	286 191	228 953	57 238
FTP-BruteForce	193 354	154 683	38 671
SSH-Bruteforce	187 589	150 071	37 518
Infilteration	160 639	128 511	32 128
DoS SlowHTTPTest	139 890	111 912	27 978
DoS attacks-GoldenEye	41 508	33 206	8 302
DoS attacks-Slowloris	10 990	8 792	2 198
DDoS LOIC-UDP	1 730	1 384	346
Brute Force -Web	611	489	122
Brute Force -XSS	230	184	46
SQL Injection	87	70	17
总计	4 825 536	3 995 470	998 866

6.2　数据预处理

为了提升模型的训练效率和数据质量，对 CIC-IDS2017 和 CSE-CIC-IDS2018 数据集进行预处理，具体步骤如下。

1) 数据清洗。从数据集中删除重复的记录以及包含 Nan 和 Infinity 值的无效数据，以降低数据的冗余性。对于缺失值，采用加权平均法进行填补。此外，由于数据集中正常流量的样本量显著高于异常流量，且不同类型的异常流量样本存在明显的不平衡，使用数据下采样策略对样本数据进行增强，有效缓解数据不平衡的问题。

2) 数值化。对 CIC-IDS2017 数据集和 CSE-CIC-IDS2018 数据集中的标签和非数值类型进行 one-hot 编码，将其转换成相应的离散数值形式。

3) 标准化。由于数据集中不同流量特征值之间差异较大，为减少这种差异对模型训练的影响，本文对数据集中的特征值进行标准化处理。通过标准化处理，将数据集中特征值的均值调整为 0，标准差调整为 1，从而避免因特征取值不同导致权重偏差较大的问题。标准化的计算方式如下

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{21}$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{22}$$

$$x' = \frac{x - \bar{x}}{S} \tag{23}$$

其中，n 为样本数，x 为原始特征值，$\bar{x}$ 为特征值的均值，S 为特征值的标准差，x' 为标准化值。

6.3　实验设置与评估指标

实验使用的 CPU 为 Intel(R) Xeon(R) Gold 5212 处理器，16 GB 内存，NVIDIA GeForve RTX 4060Ti GPU。使用编程语言 Python3.9 和深度学习框架 PyTorch 1.11 实现本文检测方法。模型具体训练参数如表 3 所示。

表 3　训练参数

参数	值
位置注意力系数 μ	0.7
通道注意力系数 β	0.5
加权融合系数 λ	0.5
优化器	Adam
学习率	1×10^{-5}
批次大小	256
训练轮次	100
损失函数	Focal Loss

为了验证本文方法的有效性，在CIC-IDS2017和CSE-CIC-IDS2018数据集上设计并开展特征选择算法实验、异常流量检测二分类实验、异常流量检测多分类实验、超参数实验、算法时间复杂度实验和消融实验。

为了评估本文方法在异常流量方面的检测性能，本文分别使用准确率（Accuracy）、精确率（Precision）、召回率（Recall）、F1分数（F1-score）作为异常流量检测任务的评估指标。上述评估指标的计算式分别为

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \tag{24}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{25}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{26}$$

$$\text{F1-score} = \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}} \tag{27}$$

6.3.1　特征选择算法实验与结果分析

根据本文提出的基于动态分组的特征选择算法对CIC-IDS2017和CSE-CIC-IDS2018数据集中的特征进行筛选。为了确定最优的特征集合，通过设置评分函数f_{score}中的平衡系数α为[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]，以评估其对特征选择效果的影响，不同α的特征选择实验结果如表4所示。本文分别从特征数量、检测准确率和检测时间3个方面分析不同平衡系数α值对特征选择效果的影响，具体分析如下。

表4　　不同α的特征选择实验结果

平衡系数α	CIC-IDS2017数据集			CSE-CIC-IDS2018数据集		
	特征数量/个	检测准确率	检测时间/s	特征数量/个	检测准确率	检测时间/s
0.1	47	0.980 9	2.37	47	0.924 0	3.99
0.2	42	0.980 5	2.00	46	0.923 5	3.85
0.3	37	0.981 8	1.96	38	0.936 7	3.73
0.4	35	0.980 6	1.92	34	0.935 3	3.59
0.5	34	0.981 8	1.84	33	0.934 6	3.70
0.6	32	0.981 4	1.89	33	0.934 2	3.64
0.7	29	0.981 4	1.85	34	0.933 9	3.67
0.8	25	0.981 1	1.95	32	0.935 7	3.62
0.9	23	0.979 9	1.89	30	0.933 9	3.60

1) 当α取较低值时，评分函数更注重特征与标签之间的相关性，而忽略了特征之间的冗余性，导致模型的复杂度和检测时间增加。当α为0.1时，在CIC-IDS2017数据集上，筛选后的特征数量为47个，准确率为0.980 9，但检测时间为2.37 s，计算成本较高；在CSE-CIC-IDS2018数据集上，筛选后的特征数量为47个，准确率为0.924 0，但检测时间为3.99 s，计算成本较高。因此，虽然较低的α值提高了准确率，但由于保留了较多冗余特征，在计算资源有限和实时性要求较高的实际应用中存在局限。

2) 当α取中等值时，评分函数在特征与标签的相关性和特征间的冗余性之间达到了较好的平衡效果，检测模型实现了较高的准确率和较低的计算成本。当α为0.5时，在CIC-IDS2017数据集上，筛选后的特征数量减少至34个，检测准确率为0.981 8，检测时间为1.84 s，实现了较高的检测准确率并降低了计算开销；当α为0.4时，在CSE-CIC-IDS2018数据集上，筛选后的特征数量为34个，检测准确率为0.935 3，检测时间为3.59 s，检测模型在准确率与计算成本之间实现了较好的平衡。因此，本文在CIC-IDS2017数据集和CSE-CIC-IDS2018数据集上将α分别设置为0.5和0.4，以在准确率和计算成本之间实现的最佳的权衡效果。

3) 当α取较高值时，评分函数倾向于选择与标签高度相关的特征，但忽略了特征间的相关性，导致筛选后的特征数量减少。当α为0.9时，在CIC-IDS2017数据集上，特征数量减少至23个，准确率为0.979 9，检测时间为1.89 s；在CSE-CIC-IDS2018数据集上，准确率为0.933 9，检测时间为3.60 s。较高的α值虽然减少了特征数量，提高了检测效率，但部分关键特征未能保留，导致检测准确率下降。

6.3.2　异常流量检测二分类实验与结果分析

为验证本文方法的检测性能，将本文方法分别与MTDGSE[7]、MCLDM[10]、Res-TranBiLSTM[14]和MSA-GCNN[17]这4种方法在CIC-IDS2017数据集上进行二分类实验，检测结果如表5所示。

由表5可知，与其他4种二分类方法相比，本文方法在准确率、召回率和F1分数上均有明显提升。其中，准确率、召回率和F1分数的最大提升幅度分别为0.021 2、0.131 7和0.072 0。Res-TranBiLSTM在所有方法中性能较弱，这是因为其在特征提取过程中主要依赖于LSTM层的时间序列建模

能力，难以有效提取流量数据的局部和全局特征。同时，Res-TranBiLSTM 在应对高维度、复杂结构的网络流量数据时容易出现信息丢失，导致特征表达不充分，从而影响异常流量检测的准确性。

表 5　CIC-IDS2017 数据集二分类检测结果

方法	准确率	召回率	精确率	F1 分数
MTDGSE	0.979 6	0.966 0	0.968 3	0.967 1
MCLDM	0.984 3	0.971 7	0.986 5	0.979 0
Res-TranBiLSTM	0.989 9	0.989 9	0.858 0	0.919 2
MSA-GCNN	0.971 2	**0.994 5**	0.984 7	0.989 5
MSAFE-ATD	**0.992 4**	0.991 7	**0.990 7**	**0.991 2**

为了进一步验证本文方法对异常流量检测的泛化性，将本文方法分别与 CGAN-IDS[8]、HID-MCLSTM[11]、ENFM[12] 和 GNN-NIDS[15] 这 4 种方法在 CSE-CIC-IDS2018 数据集上进行二分类实验，检测结果如表 6 所示。

表 6　CSE-CIC-IDS2018 数据集二分类检测结果

方法	准确率	召回率	精确率	F1 分数
CGAN-IDS	0.982 1	0.982 1	0.982 4	0.982 2
HID-MCLSTM	0.982 5	0.986 7	0.974 8	0.980 7
ENFM	0.983 0	0.890 0	0.940 0	0.914 3
GNN-NIDS	0.948 9	0.839 2	0.969 1	0.899 4
MSAFE-ATD	**0.989 4**	**0.989 8**	**0.988 9**	**0.989 3**

由表 6 可知，与其他 4 种二分类方法对比，本文方法在准确率、召回率、精确率和 F1 分数均有明显提升。其中，准确率、召回率、精确率和 F1 分数的最大提升幅度分别为 0.040 5、0.150 6、0.048 9 和 0.089 9。在上述 4 种二分类对比方法中，GNN-NIDS 方法的检测性能较弱，其原因为 GNN 主要关注节点间的静态关系，对局部特征和不同尺度时序特征的提取能力不足，且该方法在处理动态流量数据时难以提取其中细微异常特征，导致在高维度和复杂结构的流量场景中将流量数据转换成图结构时，丢失部分关键流量特征信息。

上述二分类实验结果表明，本文方法在准确率、召回率、精确率以及 F1 分数 4 个评价指标上均获得较好的结果，证实了其在异常流量二分类检测中的有效性和实用性。

6.3.3　异常流量检测多分类实验与结果分析

为了验证本文方法在异常流量检测多分类检测任务中的有效性，在 CIC-IDS2017 和 CSE-CIC-IDS2018 数据集进行多分类实验，结果分别如表 7 和表 8 所示。

表 7　CIC-IDS2017 数据集多分类检测结果

攻击类型	准确率	召回率	精确率	F1 分数
Benign	0.993 1	0.993 1	0.997 8	0.995 4
Bot	0.981 7	0.981 7	0.935 3	0.957 9
DDoS	0.988 3	0.988 3	0.999 3	0.993 7
DoS GoldenEye	0.998 1	0.998 1	0.995 2	0.996 6
DoS Hulk	0.999 0	0.999 0	0.973 8	0.986 2
DoS Slowhttptest	0.997 2	0.997 2	0.998 2	0.997 7
DoS Slowloris	0.997 1	0.997 1	0.994 0	0.995 5
FTP-Patator	0.993 4	0.993 4	0.996 4	0.994 8
Heartbleed	0.999 9	0.999 9	0.999 9	0.999 9
Infiltration	0.995 3	0.995 3	0.997 4	0.996 3
PortScan	0.998 4	0.998 4	0.999 0	0.998 7
SSH-Patator	0.993 6	0.993 6	0.758 9	0.860 5
Brute Force	0.987 9	0.987 9	0.957 0	0.972 2
SQL Injection	0.692 3	0.692 3	0.993 7	0.816 0
XSS	0.977 8	0.977 8	0.984 4	0.989 1

表 8　CSE-CIC-IDS2018 数据集多分类检测结果

攻击类型	准确率	召回率	精确率	F1 分数
Benign	0.999 2	0.999 2	0.988 8	0.989 0
Bot	0.999 9	0.999 9	0.999 9	0.999 9
Brute Force-Web	0.969 5	0.969 5	0.993 9	0.981 5
Brute Force-XSS	0.963 2	0.963 2	0.963 0	0.963 1
DDoS-HOIC	0.999 9	0.999 9	0.999 9	0.999 9
DDoS-LOIC-UDP	0.999 9	0.999 9	0.999 8	0.999 9
DDoS-LOIC-HTTP	0.999 9	0.999 9	0.997 3	0.998 5
DoS-GoldenEye	0.999 8	0.999 8	0.996 4	0.998 0
DoS-Hulk	0.999 6	0.999 6	1.000 0	0.989 7
DoS-SlowHTTPTest	0.994 3	0.994 3	0.995 6	0.994 9
DoS-Slowloris	0.999 6	0.999 6	0.999 8	0.999 7
FTP-BruteForce	0.998 2	0.998 2	0.957 6	0.957 9
Infilteration	0.960 4	0.960 4	0.958 3	0.949 2
SQL Injection	0.968 4	0.968 4	0.960 1	0.959 2
SSH-Bruteforce	0.999 9	0.999 9	1.000 0	0.999 9

由表7可知，在CIC-IDS2017数据集上，本文方法在准确率、召回率、精确率和F1分数4个评价指标上均达到了较好的检测效果，特别是在检测Heartbleed类型的攻击流量时，精确率和F1分数均达到了0.999 9，表明本文方法对此类型攻击具有较高的敏感性，但对于SQL Injection攻击的检测效果较弱，准确率和召回率仅为0.6923，其原因分析如下。

1) Heartbleed攻击利用SSL/TLS协议中的漏洞，通过伪造连接请求窃取服务器内存中的敏感数据，该攻击模式通常在流量层面表现为不符合常规的SSL/TLS数据包特征，并且具有显著的单向数据流。本文方法通过多尺度特征提取网络能够有效识别异常的数据流方向，同时能够充分提取SSL/TLS协议异常特征，从而实现对Heartbleed攻击的有效检测。

2) SQL Injection攻击通过在Web请求中注入SQL语句以干扰数据库操作，该攻击通常嵌入在正常的HTTP请求中。因此，SQL Injection流量包的整体结构并无明显异常，异常特征主要体现在SQL关键字和特定字符等请求内容的细节层面，导致在特征提取过程中难以在不同尺度上有效提取到关键特征。此外，SQL Injection的样本数量较少，导致检测模型在训练时难以获取足够的特征信息，使其对SQL Injection的检测能力进一步受限。

由表8可知，在CSE-CIC-IDS2018数据集上，本文方法对大部分攻击类型上的准确率和F1分数均达到了0.99以上。对于DDoS-HOIC和DDoS-LOIC-UDP类型的攻击，本文方法实现了0.999 9的准确分类，表明本文方法在处理大流量分布式拒绝服务攻击方面具有较好的性能。对于Brute-Force-Web、Brute-Force-XSS和SQL-Injection攻击类型，本文方法的F1分数平均为0.972 6，略低于其他攻击类型的检测效果。其原因分析如下。

1) DDoS攻击通常伴随大量异常流量，并且流量速率和并发连接显著增加。本文方法能够有效提取高频请求和短时流量峰值特征，对DDoS攻击中的异常模式具有较高的敏感性，从而实现对DDoS攻击的有效检测。

2) Brute-Force-Web、Brute-Force-XSS和SQL Injection等Web攻击类型的共同点在于依赖HTTP请求的特定模式变化，并且涉及复杂的请求−响应交互。同时，数据集中样本数量相对较少，特征分布较为稀疏，导致检测模型在训练过程中未能充分学习到Web攻击的细微差异。本文方法虽然能够在不同尺度上提取流量特征，但对Web请求内容层面微小变化的敏感性较弱，导致检测效果不佳。

将本文方法分别与DT、DNN、LSTM、RTIDS[9]、HAGRU[13]和E-GraphSAGE[16]这6种方法进行多分类对比实验，在CIC-IDS2017和CSE-CIC-IDS2018数据集上的F1分数实验结果分别如表9和表10所示。

由表9可知，在CIC-IDS2017数据集上，本文方法对Heartbleed攻击类型达到了0.999 9的检测效果，表现出较好的检测效果。同时，在对Brute Force和Infiltration攻击检测时，与其他方法相比，本文方法分别提升0.386 9和0.371 3。对于Bot、DDoS和FTP-Patator攻击类型，本文方法的F1分数均超过0.99，表明其在处理这些攻击类型时具有较好的鲁棒性和稳定性。

由表10可知，本文方法对Bot、DDoS-HOIC、DDoS-LOIC-UDP和SSH-Bruteforce攻击的F1分数均达到了0.999 9。与其他方法相比，本文方法对Infilteration、SQL Injection和SSH-Bruteforce的检测效果提升明显，进一步验证了本文方法的有效性，以及对多类别网络攻击流量较好的检测性。

上述多分类对比结果表明，本文方法在异常流量检测方面具有较好的分类效果，其原因分析如下。

1) 在异常流量检测多分类实验中，DT、DNN和LSTM在提取流量特征时能力较弱，仅提取某个层次的流量特征信息，未能充分提取流量数据的局部和全局特征，存在特征信息丢失问题，从而导致异常流量的检测性能不佳。

2) RTIDS、HAGRU和E-GraphSAGE方法没有对重要的特征图信息进行增强，导致特征表达的区分度和整体表达的有效性不足。此外，上述4种方法缺乏对特征冗余性的处理，容易导致模型训练时计算复杂度增加，同时降低了模型的检测效率和准确性。

3) 本文方法通过Dense-CNN和MSAFEN分别提取流量数据的局部和全局2种不同层次的特征信

表 9　CIC-IDS2017 数据集的 F1 分数实验结果

攻击类型	DT	DNN	LSTM	RTIDS	HAGRU	E-GraphSAGE	**MSAFE-ATD**
Benign	0.905 1	0.936 0	0.785 9	0.999 7	0.995 2	0.981 1	**0.995 5**
Bot	**0.997 4**	0.750 7	—	—	0.796 8	0.928 9	0.957 9
DDoS	0.989 5	0.999 0	0.965 0	0.999 1	**0.999 6**	0.830 1	0.993 7
DoS GoldenEye	0.972 8	0.988 3	0.839 9	0.973 3	0.996 6	—	**0.996 6**
DoS Hulk	0.996 8	0.995 6	0.978 8	0.997 2	**0.996 6**	0.878 0	0.986 3
DoS Slowhttptest	0.953 6	0.982 1	0.798 2	0.967 5	0.996 6	0.035 8	**0.997 7**
DoS Slowloris	0.513 2	0.976 5	0.797 2	0.985 9	0.996 6	0.024 7	**0.997 5**
FTP-Patator	**0.997 7**	0.970 0	0.884 2	0.996 2	0.996 5	0.975 2	0.994 9
Heartbleed	0.999 9	—	—	0.818 6	—	0.999 9	**0.999 9**
Infiltration	0.625 0	0.266 7	—	0.959 7	—	0.087 0	**0.996 3**
PortScan	0.997 4	0.996 8	0.985 2	0.999 1	**0.999 6**	0.993 9	0.998 7
SSH-Patator	0.992 9	0.860 7	0.821 2	0.984 0	—	**0.975 2**	0.860 5
Brute Force	0.261 4	0.721 3	0.585 3	**0.996 7**	—	0.072 4	0.972 2
SQL Injection	—	—	—	0.741 7	**0.985 2**	—	0.816 1
XSS	0.039 6	—	—	0.988 4	0.985 2	—	**0.989 1**

表 10　CSE-CIC-IDS2018 数据集的 F1 分数实验结果

攻击类型	DT	DNN	LSTM	RTIDS	HAGRU	E-GraphSAGE	**MSAFE-ATD**
Benign	0.964 0	0.972 8	0.970 8	0.978 5	0.945 2	0.947 4	**0.989 0**
Bot	0.999 9	0.461 5	0.958 0	0.979 4	0.999 9	0.999 9	**0.999 9**
Brute Force-Web	0.230 8	0.228 2	0.963 6	0.979 3	0.883 2	0.148 3	**0.981 5**
Brute Force-XSS	—	0.217 4	0.967 5	**0.976 7**	0.883 2	0.020 9	0.963 1
DDoS-HOIC	0.998 6	0.983 3	0.962 9	0.969 8	0.999 9	0.998 6	**0.999 9**
DDoS-LOIC-UDP	0.988 9	0.989 7	0.632 1	0.732 9	0.999 9	0.990 5	**0.999 9**
DDoS-LOIC-HTTP	0.983 9	0.983 3	0.971 8	0.986 7	**0.999 9**	0.984 1	0.998 5
DoS-GoldenEye	0.998 5	0.945 2	0.972 7	0.982 7	**0.999 6**	0.979 7	0.998 0
DoS-Hulk	0.999 4	0.999 2	0.975 2	0.981 5	**0.999 9**	0.962 8	0.989 7
DoS-SlowHTTPTest	0.521 2	0.583 7	0.971 8	0.984 3	0.665 4	—	**0.994 9**
DoS-Slowloris	0.987 8	0.838 1	0.975 7	0.985 5	0.999 3	0.867 2	**0.999 7**
FTP-BruteForce	—	0.366 1	—	0.452 7	0.835 3	0.786 1	**0.957 9**
Infilteration	0.096 6	0.461 5	0.402 6	0.474 6	0.885 7	0.144 7	**0.949 2**
SQL Injection	—	0.128 6	—	—	0.883 2	0.144 7	**0.959 2**
SSH-Bruteforce	0.999 7	0.999 7	—	0.219 1	0.999 9	0.908 8	**0.999 9**

息，实现了流量局部特征和全局特征间的交互，减少了特征信息丢失，并使用 PAMFC 和 CAMFC 对提取的特征进行增强，提升了流量特征之间的区分度，从而实现了对异常流量的有效检测。

6.3.4　超参数实验与结果分析

为分析位置注意力系数 μ、通道注意力系数 β 和加权融合系数 λ 对异常流量检测效果的影响，本节将分别从理论分析和实验验证 2 个方面探讨不同

参数设置对检测结果的影响。

位置注意力系数μ用于动态调整特征向量在不同位置上的权重，使检测模型能够更有效地关注流量数据中的时间序列变化特征，从而增强时间序列中位置信息的表达能力。当位置注意力系数μ较大时，检测模型将加强对关键位置特征的关注，使其在面对短时内网络连接频次的异常波动和瞬间流量激增等网络行为时，能够更好地捕捉到流量特征中的位置突变信息；当位置注意力系数μ较小时，检测模型将均衡地分配各位置特征的权重，减少部分位置信息对整体特征的突出作用，使检测模型在全局范围内平衡位置特征的影响。

通道注意力系数β用于调整检测模型在不同通道维度上的权重，通过强调特征间的关联性突出关键通道特征。较大的通道注意力系数β强化了关键通道特征在整体特征图中的表达，使检测模型能够在多通道中捕捉细微的异常模式，有利于增强对关联性较高的特征提取能力；当通道注意力系数β较小时，检测模型会弱化通道特征在全局信息中的影响，从而更均衡地融合各通道的特征，同时减少对特定通道特征的依赖，增强检测模型的鲁棒性。

加权融合系数λ用于在局部特征和全局特征之间进行加权平衡。局部特征能够反映流量数据中位置上的细粒度信息，而全局特征反映整体趋势。在特征融合过程中，加权融合系数λ较大时，表示模型更侧重局部特征，导致融合后的特征更倾向于保留局部位置的细粒度信息，有利于检测突发性或短时异常流量；加权融合系数λ较小时，全局特征的权重则相对较大，表示在特征融合过程中更偏向于全局特征的表达，在长时间跨度和高维空间上稳定分布的异常流量检测任务中，较小的加权融合系数λ能够有效识别较为隐蔽的全局异常特征。

为验证上述原理分析的合理性，分别将位置注意力系数μ、通道注意力系数β和加权融合系数λ设定为0~1，并在CIC-IDS2017和CSE-CIC-IDS2018数据集上进行异常流量检测多分类实验，通过实验分析不同参数设置下的F1分数变化，以确定位置注意力系数μ、通道注意力系数β和加权融合系数λ对检测效果的影响。CIC-IDS2017和CSE-CIC-IDS2018数据集上不同超参数实验结果分别如图7和图8所示。

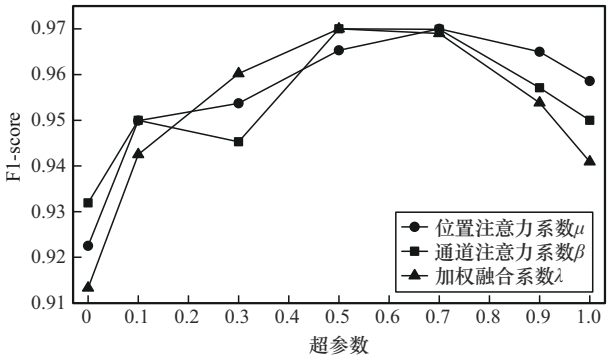


图7　CIC-IDS2017数据集上不同超参数实验结果

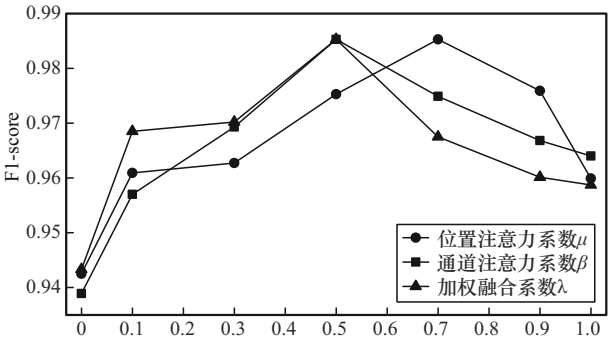


图8　CSE-CIC-IDS2018数据集上不同超参数实验结果

由图7和图8可知，随着位置注意力系数μ、通道注意力系数β和加权融合系数λ的取值变化，本文方法在CIC-IDS2017和CSE-CIC-IDS2018数据集上的F1分数呈现先上升后下降的趋势。当位置注意力系数μ小于0.7，通道注意力系数β和加权融合系数λ小于0.5时，随着参数值的增加，检测性能逐步提升；当位置注意力系数μ为0.7，通道注意力系数β和加权融合系数λ为0.5时，本文方法在2个数据集上的F1分数达到最高；当位置注意力系数μ大于0.7，通道注意力系数β和加权融合系数λ大于0.5时，随着参数大小的增加，检测性能开始下降。其原因为当位置注意力系数μ较低时，检测模型未能充分利用位置信息增强特征表达，当μ增大到临界点后，过度强化位置信息导致模型对训练数据的特定特征过于敏感，而忽略了流量数据的空间特征。当加权融合系数λ为0.5时，检测模型对局部与全局特征的融合较为均衡，有效避免了对单一特征的过度依赖，因此，本文将位置注意力系数μ设置为0.7，通道注意力系数β和加权融合系数λ设置为0.5。

为提升本文方法在不同数据集场景下的适用性，采用以下通用性策略确定位置注意力系数μ、

通道注意力系数β和加权融合系数λ的最佳取值。

1) 初始设置。在应用于未知数据集时，将位置注意力系数μ、通道注意力系数β和加权融合系数λ设置在中等偏低范围，以保持位置和通道特征的平衡，避免模型初始阶段对部分单一特征产生过度的依赖。

2) 逐步调整。根据数据集特征分布进行微调，若数据集中局部异常显著，可逐步提升位置注意力系数μ，使模型在特定位置上有更高的关注度。若通道特征关联性较强或差异性显著，则逐步增加通道注意力系数β，以突出通道特征的表达；若数据集中的全局特征较为稳定，则逐步调整加权融合系数λ以实现更优的局部和全局特征平衡。

3) 协同优化。位置注意力系数μ、通道注意力系数β和加权融合系数λ的设置需要协同调整。若位置注意力系数μ和通道注意力系数β设置较高，表示模型已偏向于位置和通道特征，应适当降低加权融合系数λ以保证全局特征的平衡性；若加权融合系数λ较高，应降低位置注意力系数μ和通道注意力系数β，防止模型仅关注全局特征而忽略局部特征。

6.3.5　算法时间复杂度实验与结果分析

为了验证本文方法在时间复杂度方面的性能，分别在CIC-IDS2017和CSE-CIC-IDS2018数据集上通过训练时间和检测时间对比不同检测方法的时间消耗，分别如表11和表12所示。

表11　CIC-IDS2017数据集上时间消耗对比

方法	训练时间/min	检测时间/s
MTDGSE	37	3.70
MCLDM	45	4.19
Res-TranBiLSTM	34	3.51
MSA-GCNN	33	3.43
MSAFE-ATD	**31**	**2.74**

表12　CSE-CIC-IDS2018数据集上时间消耗对比

方法	训练时间/min	检测时间/s
CGAN-IDS	51	4.99
HID-MCLSTM	63	6.32
ENFM	46	4.07
GNN-NIDS	52	5.35
MSAFE-ATD	**45**	**3.98**

由表11和表12可知，本文方法在2个数据集上的训练时间分别为31 min和45 min，检测时间分别为2.74 s和3.98 s。与MTDGSE[7]、CGAN-IDS[8]、MCLDM[10]、HID-MCLSTM[11]、ENFM[12]、Res-TranBiLSTM[14]、GNN-NIDS[15]和MSA-GCNN[17]检测方法相比，本文方法具有较好的检测效果。其原因分析如下。

1) 本文采用基于动态分组的特征选择算法去除流量数据中的冗余特征，降低了流量数据的特征维度，减少了计算资源的消耗，从而有效缩短了模型的训练时间和检测时间。

2) 本文采用基于PAMFC和CAMFC的特征增强网络对局部和全局特征增强。该特征增强网络能够突出局部和全局特征中的关键信息，抑制冗余信息的干扰，减少了检测模型对高维特征的依赖，从而提高了异常流量检测效率。

6.3.6　消融实验与结果分析

由于本文方法包含多个关键部分，因此，采用消融实验验证本文方法中各个部分的有效性。消融实验包括以下3种对比方法。

1) MSAFE-ATD\D：在MSAFE-ATD方法的基础上删除Dense-CNN，仅利用MSAFEN提取流量数据的全局特征，并使用基于PAMFC和CAMFC的特征增强网络对全局特征进行增强。该对比方法用于验证局部特征对异常流量检测的有效性。

2) MSAFE-ATD\M：在MSAFE-ATD方法的基础上删除MSAFEN，仅利用Dense-CNN提取流量数据的局部特征，并使用基于PAMFC和CAMFC的特征增强网络对局部特征进行增强。该对比方法用于验证全局特征对异常流量检测的有效性。

3) MSAFE-ATD\A：在MSAFE-ATD方法的基础上删除PAMFC和CAMFC，利用Dense-CNN和MSAFEN提取流量数据的局部和全局特征，并对特征进行加权融合检测分类。该对比方法用于验证特征增强网络对异常流量检测的有效性。

基于CIC-IDS2017和CSE-CIC-IDS2018数据集的消融实验结果分别如图9和图10所示。由图9和图10可得以下结论。

1) 局部特征的影响：通过比较MSAFE-ATD与MSAFE-ATD\D在2个数据集上的检测结果可知，MSAFE-ATD方法的检测性能优于MSAFE-ATD\D，表明Dense-CNN提取的流量局部特征能够提供与全

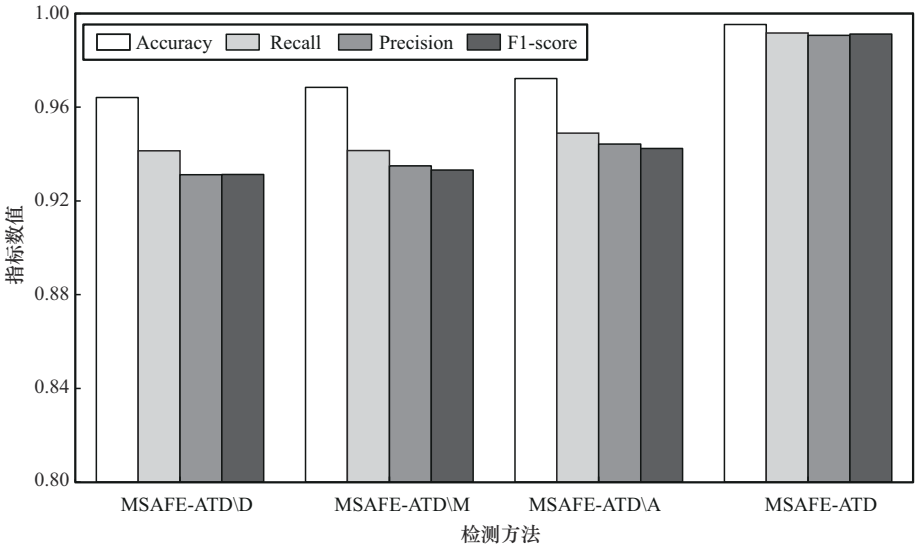


图9　基于CIC-IDS2017数据集的消融实验结果

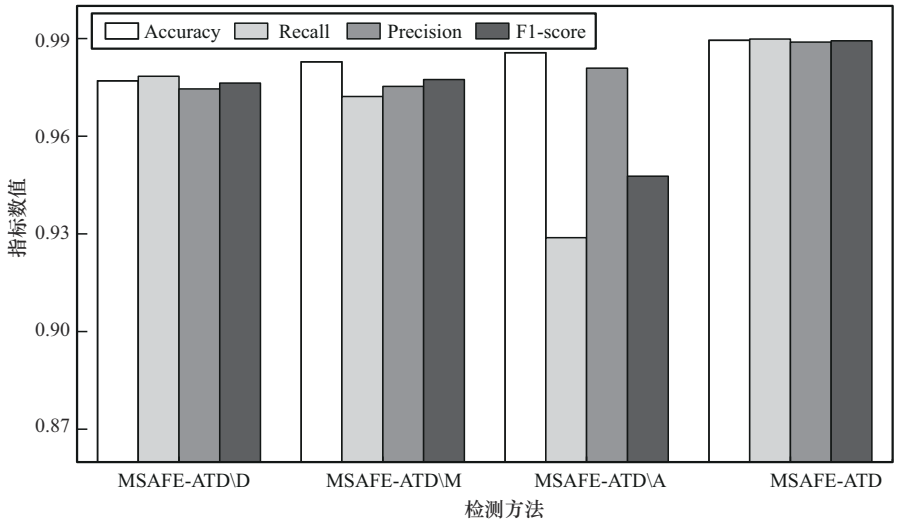


图10　基于CSE-CIC-IDS2018数据集的消融实验结果

局特征紧密相关的补充信息，从而提高检测性能。

2) 全局特征的影响：通过比较MSAFE-ATD与MSAFE-ATD\M方法的实验结果发现，MSAFE-ATD方法的F1分数在2个数据集上均优于MSAFE-ATD\M方法，表明MSAFEN能够提取流量数据中丰富的全局特征，为异常流量检测提供更多的有效信息，从而提升检测性能。

3) 特征增强网络的影响：通过比较MSAFE-ATD与MSAFE-ATD\A方法的检测准确率和F1分数的实验结果表明，MSAFE-ATD方法检测性能优于MSAFE-ATD\A，表明PAMFC和CAMFC能够有效地对提取的局部和全局特征进行增强，提升特征间的区分度和整体表达的有效性。

7　结束语

针对流量数据特征冗余、网络流量序列的时间依赖性以及长期相互依赖关系导致异常流量检测精度较低等不足，本文提出一种基于多尺度注意力特征增强的异常流量检测方法。首先，通过基于动态分组的特征选择算法去除流量数据中的冗余特征，减少异常流量检测模型的训练时间。然后，采用多尺度注意力特征提取网络从不同尺度提取流量数据的全局特征，提高检测模型对复杂和隐蔽攻击的检测能力，同时结合Dense-CNN提取局部特征，进一步丰富流量数据的特征表达。最后，通过基于PAMFC和CAMFC的特征增强网络对局部和全局特征增强，以提升整体特征信息的有效性。在CIC-

IDS2017 和 CSE-CIC-IDS2018 数据集上的实验结果表明，本文方法的异常流量检测性能优于现有方法。

在未来研究中，将考虑设计更加高效的特征提取和特征增强算法，以进一步降低计算复杂度。同时，将引入更多样化和真实的流量数据进行训练，提升异常流量检测模型在不同网络环境下的检测效率和泛化能力。

参考文献：

[1] 刘奇旭, 陈艳辉, 尼杰硕, 等. 基于机器学习的工业互联网入侵检测综述[J]. 计算机研究与发展, 2022, 59(5): 994-1014.
LIU Q X, CHEN Y H, NI J S, et al. Survey on machine learning-based anomaly detection for industrial Internet[J]. Journal of Computer Research and Development, 2022, 59(5): 994-1014.

[2] 任家东, 刘新倩, 王倩, 等. 基于 KNN 离群点检测和随机森林的多层入侵检测方法[J]. 计算机研究与发展, 2019, 56(3): 566-575.
REN J D, LIU X Q, WANG Q, et al. An multi-level intrusion detection method based on KNN outlier detection and random forests[J]. Journal of Computer Research and Development, 2019, 56(3): 566-575.

[3] DING H W, SUN Y, HUANG N N, et al. TMG-GAN: generative adversarial networks-based imbalanced learning for network intrusion detection[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 1156-1167.

[4] DUAN X Y, FU Y, WANG K. Network traffic anomaly detection method based on multi-scale residual classifier[J]. Computer Communications, 2023, 198: 206-216.

[5] WU T, FAN H H, ZHU H J, et al. Intrusion detection system combined enhanced random forest with SMOTE algorithm[J]. EURASIP Journal on Advances in Signal Processing, 2022(1): 39.

[6] LU C W, CAO Y X, WANG Z B. Research on intrusion detection based on an enhanced random forest algorithm[J]. Applied Sciences, 2024, 14 (2): 714.

[7] HOU B T, ZHANG K, ZUO X J, et al. PIoT malicious traffic detection method based on GAN sample enhancement[J]. Security and Communication Networks, 2022, 2022: 9223412.

[8] LI F, SHEN H, MAI J A, et al. Pre-trained language model-enhanced conditional generative adversarial networks for intrusion detection[J]. Peer-to-Peer Networking and Applications, 2024, 17(1): 227-245.

[9] 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.

[10] LUO J, ZHANG Y Y, WU Y N, et al. A multi-channel contrastive learning network based intrusion detection method[J]. Electronics, 2023, 12 (4): 949.

[11] KANNA P R, SANTHI P. Hybrid intrusion detection using MapReduce based black widow optimized convolutional long short-term memory neural networks[J]. Expert Systems with Applications, 2022, 194: 116545.

[12] BHARDWAJ S, DAVE M. Enhanced neural network-based attack investigation framework for network forensics: identification, detection, and analysis of the attack[J]. Computers & Security, 2023, 135: 103521.

[13] LIU X Y, LIU J M. Malicious traffic detection combined deep neural network with hierarchical attention mechanism[J]. Scientific Reports, 2021, 11(1): 12363.

[14] WANG S Y, XU W X, LIU Y W. Res-TranBiLSTM: an intelligent approach for intrusion detection in the Internet of Things[J]. Computer Networks, 2023, 235: 109982.

[15] PUJOL-PERICH D, SUAREZ-VARELA J, CABELLOS-APARICIO A, et al. Unveiling the potential of graph neural networks for robust intrusion detection[J]. ACM SIGMETRICS Performance Evaluation Review, 2022, 49(4): 111-117.

[16] 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.

[17] REKA R, KARTHICK R, SARAVANA RAM R, et al. Multi head self-attention gated graph convolutional network based multi-attack intrusion detection in MANET[J]. Computers & Security, 2024, 136: 103526.

[18] 肖利军, 郭继昌, 顾翔元. 一种采用冗余性动态权重的特征选择算法[J]. 西安电子科技大学学报, 2019, 46(5): 155-161.
XIAO L J, GUO J C, GU X Y. Algorithm for selection of features based on dynamic weights using redundancy[J]. Journal of Xidian University, 2019, 46(5): 155-161.

[19] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 2261-2269.

[20] XIANG Z L, LI X W. RETRACTED ARTICLE: Fusion of transformer and ML-CNN-BiLSTM for network intrusion detection[J]. EURASIP Journal on Wireless Communications and Networking, 2023(1): 71.

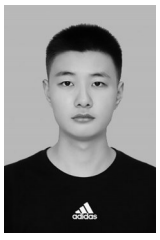
[21] ZHOU W, ZHENG F J, ZHAO Y H, et al. MSDCNN: a multiscale dilated convolution neural network for fine-grained 3D shape classification[J]. Neural Networks, 2024, 172: 106141.

[22] KHAN Z I, AFZAL M M, SHAMSI K N. A comprehensive study on CIC-IDS2017 dataset for intrusion detection systems[J]. International Research Journal on Advanced Engineering Hub (IRJAEH), 2024, 2 (2): 254-260.

[23] LEEVY J L, KHOSHGOFTAAR T M. A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 Big Data[J]. Journal of Big Data, 2020, 7(1): 104.

[作者简介]

杨宏宇（1969– ），男，吉林长春人，博士，中国民航大学教授、博士生导师，主要研究方向为网络与系统安全、软件安全检测、网络安全态势感知。



张豪豪（1999– ），男，河南焦作人，中国民航大学硕士生，主要研究方向为网络与信息安全。

成翔（1988– ），男，新疆乌鲁木齐人，博士，扬州大学讲师、硕士生导师，主要研究方向为网络与系统安全、网络安全态势感知、APT 攻击检测。