研究与开发

GhostMamba-SAN：一种高效的DDoS攻击检测模型

包晓安[1]，杨奉豪[1]，范云龙[1]，涂小妹[2]，胡天缤[3]，吴彪[4]
（1. 浙江理工大学计算机科学与技术学院，浙江 杭州 310018；
2. 浙江广厦建设职业技术大学城乡建设学院，浙江 东阳 322100；
3. 河海大学人工智能与自动化学院，江苏 常州 213000；
4. 浙江理工大学理学院，浙江 杭州 310018）

摘　要：针对软件定义网络（SDN）环境中DDoS攻击流量特征复杂、包级语义信息利用不足以及检测模型精度与效率难以兼顾的问题，设计了一种融合有效载荷信息与流级统计特征的混合检测模型。该模型首先基于改进的Mamba网络对有效载荷序列进行深度建模，以挖掘包级特征中的时序依赖与上下文信息；其次，采用Ghost卷积替代常规卷积结构，在保持特征表达能力的同时有效减少模型参数量并提升计算效率；最后，通过自注意力机制对多维融合特征进行加权，以强化关键攻击特征并抑制无关信息。实验结果表明，所设计模型在CICDDoS2019数据集上实现了99.56%的检测准确率，平均检测延迟仅为0.21 ms，优于现有主流方法。此外，在多个公开数据集上的验证结果进一步证明，该模型具有良好的泛化能力。

关键词：软件定义网络；DDoS攻击检测；深度学习；自注意力机制；网络安全
中图分类号：TP393；TM91
文献标志码：A
doi: 10.11959/j.issn.1000−0801.DXKX250676

GhostMamba-SAN: an efficient DDoS attack detection model

Bao Xiaoan[1], Yang Fenghao[1], Fan Yunlong[1], Tu Xiaomei[2], Hu Tianbin[3], Wu Biao[4]
1. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China
2. School of Urban and Rural Construction, Zhejiang Guangsha Vocational and Technical University of Construction, Dongyang 322100, China
3. School of Artificial Intelligence and Automation, Hohai University, Changzhou 213000, China
4. School of Science, Zhejiang Sci-Tech University, Hangzhou 310018, China

Abstract: To address the challenges of complex traffic characteristics, insufficient utilization of packet-level semantic information, and the trade-off between detection accuracy and efficiency in software-defined networking (SDN) environments, a hybrid detection model that integrates payload information and flow-level statistical features was proposed. Specifically, an improved Mamba network was employed to perform deep modeling of payload sequences, en-

收稿日期：2025−11−24；修回日期：2025−12−16
通信作者：吴彪，biaowuzg@zstu.edu.cn
基金项目：浙江省重点研发计划项目（No.2020C03094）；浙江省教育厅项目（No.Y202250706，No.Y202147659）
Foundation Items: Zhejiang Provincial Key Research and Development Program (No. 2020C03094), Projects of Zhejiang Provincial Department of Education (No.Y202250706, No.Y202147659)

abling the extraction of temporal dependencies and contextual semantics within packet-level features. Meanwhile, the conventional convolutional structure was replaced with Ghost convolution to effectively reduce the number of parameters and computational cost while maintaining strong feature representation capability. Finally, a self-attention mechanism was introduced to adaptively weight and fuse multi-dimensional features, which enhanced the representation of critical attack patterns and suppresses irrelevant information. Experimental results demonstrate that the proposed model achieves a detection accuracy of 99.56% on the CICDDoS2019 dataset with an average detection latency of only 0.21 ms, outperforming existing mainstream methods. Moreover, validation on multiple public datasets further confirms the strong generalization capability of the proposed model.

Key words: software defined network, DDoS attack detection, deep learning, self-attention mechanism, network security

0　引言

近年来，全球互联网用户规模以年均12.3%的增速持续扩张，5G、物联网等技术的普及驱动网络流量呈指数级增长。传统网络架构受限于静态路由策略与分布式控制机制，在带宽动态分配、策略灵活调整及安全威胁响应等方面逐渐滞后，难以满足云计算、边缘计算等新型业务对低时延、高弹性的需求。在此背景下，软件定义网络（SDN）通过控制平面与数据平面解耦的创新架构，实现了网络资源的全局可编程调度。然而，集中式控制器的设计在提升管理效率的同时，也显著放大了安全风险，从SDN的3层框架视角看，安全风险不仅局限于单一设备或单一链路，而是可能在数据层、控制层与应用层之间跨平面耦合传播，例如，数据层的异常流量与规则更新压力会向上放大为控制层的控制消息拥塞与计算资源争用，而应用层策略下发与控制逻辑缺陷又可能进一步放大网络范围内的连锁影响[1]。其中控制器作为全网流表决策的核心节点，其中央处理器（central processing unit，CPU）与内存资源极易遭受DDoS攻击的大规模伪造请求耗尽，进而引发全网流表更新停滞、服务质量断崖式下降等系统性崩溃。据Akamai全球安全态势报告显示，2023年针对SDN控制器的DDoS攻击事件数量同比激增58%[2]。这类攻击通过耗尽控制器的带宽、CPU和内存资源，导致控制消息处理严重延迟或丢失[3]，从而引发网络中断与关键业务不可用。此期间，网络资源被大量消耗，企业不仅面临高昂的恢复成本和生产力损失，还易遭勒索软件等二次攻击[4]。因此在SDN架构下对DDoS攻击检测的研究，具有重要的意义。

SDN中DDoS攻击检测所用方法根据使用的核心技术或理论框架的不同，大致可划分为基于信息熵的方法、基于机器学习的方法和基于深度学习的方法[5]。基于信息熵的方法虽然其开销较低，但随着流量规模的增大和攻击模式的多样化，其误报率和漏报率剧增，无法适应现代网络检测需求[6]。基于机器学习的DDoS攻击检测方法，在特征空间较为简单或攻击类型较为明确的情况下能够有效检测DDoS攻击[7]，例如，Alhamamik等[8]通过一个检测系统评估了多种机器学习算法，包括XGBoost、逻辑回归、支持向量机（support vector machine，SVM）、K-近邻算法（k-nearest neighbors，KNN）和朴素贝叶斯在SDN环境中实时检测DDoS攻击的效果。Ribeiro等[9]通过机器学习实时识别DDoS流量，并利用SDN的可编程性动态变换网络配置（移动目标防御（moving target defense，MTD）机制），将攻击流量重定向至虚拟陷阱，从而实现攻击的自动化检测与主动防御。在现有工作中，Dong等[10]提出的基于攻击程度的改进KNN具有一定代表

性，该方法将流量强度相关特征映射为攻击程度，并据此对KNN的距离度量与投票机制进行加权，从而增强对不同攻击强度样本的区分能力。但面对有效载荷语义差异显著的攻击时，特征表达能力仍可能受限。基于深度学习的方法中，大部分检测算法都是基于卷积神经网络（convolutional neural network，CNN）、循环神经网络（recurrent neural network，RNN）、长短期记忆（long short term memory，LSTM）网络或它们的变体来进行构建的[11]。尽管这些算法在一些情况下取得了较好的检测精度，但在DDoS攻击流量种类增多时精度普遍较低。主要原因是这些研究大部分都只使用了流级的统计信息，当攻击流量种类增多时，流级特征的表达能力不足导致模型性能下降[12]。为了有效利用载荷信息特征，已有学者做了尝试。例如，Hosseinis等[13]基于压缩位图索引和数据筛选策略来表示部分有效载荷信息。Sohis等[14]挖掘了原始字节中的内容特征，用以识别特定类型的流量。Lotfollahim等[15]通过固定长度的有效载荷字节向量捕捉网络流量的本质特征。这些方法虽然在特定攻击场景下有效，但仅提取了少量或特定位置信息，难以应对多样化攻击。有效载荷序列往往较长，传统的序列模型在处理这类长序列的特征依赖时，由于其结构限制，容易在建模过程中遗失历史信息。Transformer能够有效建模远距离特征间的关联，但其计算复杂度会随着序列长度的增加而呈二次增长，显著增加计算负担和内存消耗，无法满足检测的实时性需求。近期，Gu等[16]提出Mamba模型，该模型利用其独特的选择性状态空间机制，能够在不使用注意力机制的情况下有效建模长程依赖关系[17]，避免了Transformer中随着序列长度增加而呈平方增长的计算开销[18]。Mamba在计算机视觉、自然语言处理等多个领域取得了显著成果，充分展现出其在长序列建模任务中的强大潜力[19]。

鉴于上述研究，针对SDN环境下DDoS攻击检测精度问题，本文设计了GhostMamba-SAN混合检测模型。本文主要贡献如下。

（1）提出面向检测的有效载荷特征构造方法：将原始有效载荷由字节级序列转化为经滑窗统计与协方差建模得到的特征序列，在输入端显式编码局部相关性与结构信息，从而降低噪声干扰并增强对攻击模式的可分性。

（2）构建统计矩阵和有效载荷序列的双分支特征表征：在流级统计特征方面，将多维统计量组织为小尺度矩阵表示以保留特征间的组合关系；在有效载荷方面，利用序列表示刻画跨字段的时序依赖，实现多粒度信息互补。

（3）针对本文输入特性对Mamba结构进行改进：基于局部信息已被统计建模显式编码的特点，移除原始Mamba前端的一维卷积块及冗余线性投影层，使模型在保持长程依赖建模能力的同时降低参数量与计算开销，从而提升推理效率。

（4）在CICDDoS2019数据集上验证模型的有效性与泛化能力：在CICDDoS2019等数据集上开展实验评估，从检测性能与复杂度指标等维度对本文方法进行系统验证，并与多种方法进行对比，实验结果表明，所提出模型在准确性与效率方面均具有优势，证明其在多类别DDoS攻击检测任务中的有效性与实用价值。

1 流量表征

1.1 统计特征

为增强流级统计特征的建模效果，本文对CICDDoS2019数据集中由专家规则构造的87个特征进行了处理。具体而言，删除了Unnamed: 0、Flow ID、Source IP、Destination IP、Source Port、Destination Port和Timestamp这7个与检测任务无关的特征。同时删除了冗余字段Fwd Header Length. 1和SimilarHTTP。Fwd URG Flags、Bwd URG Flags、CWE Flag Count、ECE Flag Count，以及Active Min和Idle Min这6个特

征在样本中的值几乎都为零，缺乏统计意义，因此删去。由于与 Subflow 相关的 4 个特征在实时检测场景中难以获取，且计算复杂度高，缺乏实际应用价值，故移除。此外，还增加流量突发强度和包长偏度这两个统计特征。流量突发强度（burst strength，BS）可以有效反映网络流量的波动性，其计算式如式（1）所示。

$$\mathrm{BS}=\log\left(1+\frac{\mathrm{VarIAT}}{\mathrm{MeanIAT}^2}\right) \tag{1}$$

其中，MeanIAT 和 VarIAT 则代表到达间隔的均值和方差。包长度偏度（packet length skewness，PLS）是衡量流量数据包长度分布相对于正态分布的偏斜程度。其计算式如下。

$$\mathrm{PS}=\frac{N}{(N-1)(N-2)}\sum_{i=1}^{N}\left(\frac{X_i-\mu}{\sigma}\right)^3 \tag{2}$$

其中，N是流中的总包数，X_i为当前包的包长，μ和σ分别代表数据包长度的均值和标准差。此外，为了确保数据集的完整性和一致性，针对统计特征中的缺失值和无限值，这里使用零进行替换，并采用 Min-Max 归一化方式对各特征值进行缩放。

1.2　有效载荷特征

Hosseinis 等[13]则提出了基于 CNN-LSTM 的 DDoS 攻击分类方法，借助 XGBoost 进行特征选择，实现了在二分类任务中达到 99.50% 的高准确率。然而，这 2 种方法在流量表征上均存在显著缺陷：它们主要仅从单个流中提取静态特征，完全忽略了流之间的时序关联以及包级别的细粒度数据。事实上，网络流量往往包含丰富的时序信息和微妙的包级变化，而这种粗糙的表征方式使得模型在应对复杂多分类任务时，无法捕捉到攻击行为的细微差别，导致整体精度下降。为了解决这一瓶颈问题，本文提出了一种全新的流量表征方案，该方案综合利用包级细粒度信息和流间时序特征，旨在构建更精准、全面的流量描述，从而显著提升多分类检测任务的准确率。

从 CICDDoS2019 数据集 pcap 文件原始数据中提取有效载荷特征过程如下。首先根据传输层五元组（源 IP 地址、目的 IP 地址、源端口、目的端口、协议类型）将各个包划分成不同的流。利用 Scapy 库编写解析脚本过滤掉校验和错误和长度异常的无效数据包，并将所有流的起始时间戳统一转换为 UTC 时间格式。然后从流中截取前 M 个包，并将每个包的前 N 个字节拼接成矩阵。为了去除无关噪声，对这些字节值进行标准化处理，得到标准化矩阵 $\boldsymbol{X}$，计算式如下所示：

$$x_{i,j}=\frac{x_{i,j}-\mu_j}{\sigma_j} \tag{3}$$

其中，μ_j和σ_j分别为行字节平均值和标准差。接下来，计算$\boldsymbol{X}$的协方差矩阵$\boldsymbol{\Sigma}$，以捕获特征间的相关性。然后对协方差矩阵$\boldsymbol{\Sigma}$进行特征值分解，将$\boldsymbol{X}$投影到前 128 个主成分构成的子空间中得到压缩后的特征矩阵$\boldsymbol{Z}$。计算式如下。

$$\boldsymbol{\Sigma}=\frac{1}{N-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} \tag{4}$$

$$\boldsymbol{\Sigma}=\boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^{\mathrm{T}} \tag{5}$$

$$\boldsymbol{Z}=\boldsymbol{X}\boldsymbol{W}_{128} \tag{6}$$

其中，$\boldsymbol{W}$为特征向量矩阵，其每一列均为协方差矩阵$\boldsymbol{\Sigma}$对应的单位特征向量。此外，为了帮助模型捕捉序列中远距离位置之间的依赖关系，对$\boldsymbol{Z}$进行正弦和余弦的位置编码，如式（7）和式（8）所示。

$$\mathrm{PE}(\mathrm{pos},2i)=\sin\left(\frac{\mathrm{pos}}{10\,000^{2i/d}}\right) \tag{7}$$

$$\mathrm{PE}(\mathrm{pos},2i+1)=\cos\left(\frac{\mathrm{pos}}{10\,000^{2i/d}}\right) \tag{8}$$

其中，pos 和 i 分别为$\boldsymbol{Z}$的行和列的编号，d为矩阵行数总和。上述过程中的M和N为超参数，通过实验确认其具体的值以达到最佳的检测效果也是本文的工作内容。

2 方法介绍

GhostMamba-SAN 检测模型架构如图 1 所示。

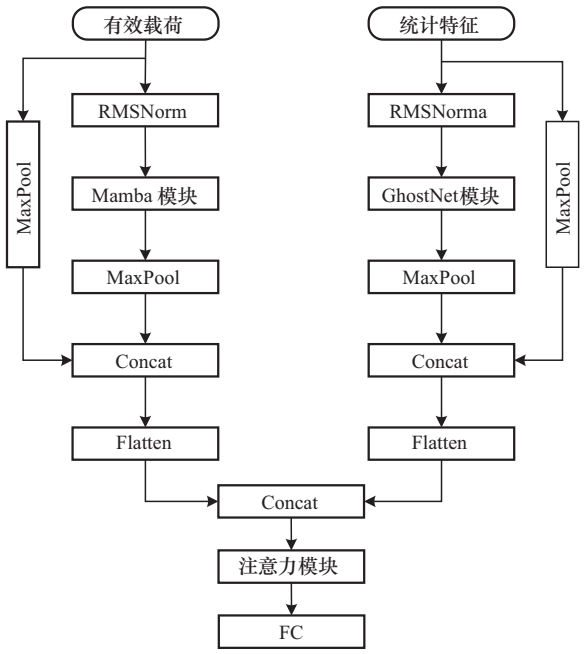


图1 GhostMamba-SAN检测模型架构

GhostMamba-SAN 由 3 部分组成：有效载荷特征分支、统计特征分支和注意力融合模块。对于单条流量样本，首先构建 2 类输入：一是经协方差建模、PCA 压缩与位置编码处理后的有效载荷特征序列，记为$\boldsymbol{X}_p$；二是统计特征向量，按 9×9 重组为矩阵输入，记为$\boldsymbol{X}_s$。在有效载荷特征分支$\boldsymbol{X}_p$输入改进 Mamba 块进行序列建模得到序列表示$\boldsymbol{H}_p$。随后沿序列维度进行池化获得载荷分支向量表示$\boldsymbol{h}_p$，以减少序列长度并保留关键语义特征。统计特征分支中，$\boldsymbol{X}_s$输入 GhostNet 模块提取统计矩阵的关联模式，得到统计分支向量表示$\boldsymbol{h}_s$。为避免简单拼接导致两类特征贡献不均，本文将$\boldsymbol{h}_p$与$\boldsymbol{h}_s$视为 2 个节点构成长度为 2 的特征序列$\boldsymbol{H}=[\boldsymbol{h}_p;\boldsymbol{h}_s]$，输入注意力融合模块进行自适应加权融合得到融合表示$\boldsymbol{h}_f$。最后将$\boldsymbol{h}_f$输入全连接分类器并经 Softmax 输出类别概率，实现对流量类型的判别。

2.1 Mamba模块

Mamba 是以选择性状态空间模型（SSM）为核心的序列建模网络。Mamba 网络结构如图 2 所示。

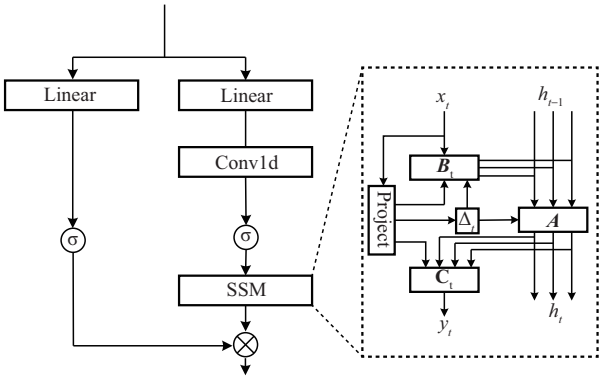


图2 Mamba网络结构

传统 SSM 的参数矩阵在时间维度上是固定不变的，导致其无法通过输入的内容调整信息的传播路径。

Mamba 在 SSM 的基础上引入了选择性机制，具体来说就是首先通过输入序列x_t的线性投影实时更新网络参数矩阵$\boldsymbol{B}_t$、$\boldsymbol{C}_t$和离散化步长Δ_t。

$$\boldsymbol{B}_t=\boldsymbol{W}_{\mathrm{B}}\boldsymbol{x}_t+\boldsymbol{b}_{\mathrm{B}} \tag{9}$$

$$\boldsymbol{C}_t=\boldsymbol{W}_{\mathrm{C}}\boldsymbol{x}_t+\boldsymbol{b}_{\mathrm{B}} \tag{10}$$

$$\Delta_t=\mathrm{Softplus}(\boldsymbol{W}_{\Delta}\boldsymbol{x}_t+\boldsymbol{b}_{\Delta}) \tag{11}$$

其中，Softplus 函数确保$\Delta>0$。接着利用零阶保持法将连续状态方程离散化为递归形式如式（12），实现状态更新。

$$\boldsymbol{h}'(t)=\boldsymbol{A}\boldsymbol{h}(t)+\boldsymbol{B}\boldsymbol{x}(t) \tag{12}$$

$$\bar{\boldsymbol{A}}_t=\mathrm{e}^{\Delta_t\boldsymbol{A}} \tag{13}$$

$$\bar{\boldsymbol{B}}_t=(\Delta_t\boldsymbol{A})^{-1}(\mathrm{e}^{\Delta_t\boldsymbol{A}}-\boldsymbol{I})\Delta_t\boldsymbol{B}_t \tag{14}$$

$$\boldsymbol{h}_t=\bar{\boldsymbol{A}}_t\boldsymbol{h}_{t-1}+\bar{\boldsymbol{B}}_t\boldsymbol{x}_t \tag{15}$$

$$\boldsymbol{y}_t=\boldsymbol{C}_t\boldsymbol{h}_t \tag{16}$$

在上述过程中，$\boldsymbol{B}_t$根据输入的序列特征有选择性地将关键信息注入状态更新，$\boldsymbol{C}_t$从隐藏状态中筛选最有判别力的维度用于输出。而Δ_t则充当

遗忘门，通过控制$\boldsymbol{A}$和$\boldsymbol{B}_t$的离散化强度，从而使模型能够在历史上下文与当前输入之间动态选择。由于$\boldsymbol{B}_t$、$\boldsymbol{C}_t$和Δ_t随输入变化而动态生成，模型从线性时不变系统转化为非线性时变系统，从而能够自适应地捕捉长序列中的关键信息。原始的Mamba结合了选择性状态空间模型与H3架构，利用两者的互补优势提升序列建模能力。然而，在本文任务中Mamba模块接收经滑窗统计与协方差建模后的有效载荷特征序列作为输入，该序列已显式编码局部邻域信息；同时，选择性SSM本身能够在时间维度实现从局部到长程的依赖建模。因此，原始Mamba中用于局部模式提取的一维卷积以及额外线性投影在本任务中存在功能重叠，既带来冗余计算与参数开销，也可能因固定感受野破坏协议字段结构的完整性。为缓解上述问题，本文对Mamba进行了结构精简。具体而言，本文移除了输入选择性状态空间模型前的一维卷积块和一个线性投影层。改进后的Mamba模块如图3所示。在改进后的架构中，输入的有效载荷序列经过单次线性投影映射后直接传入选择性SSM模块，实现对序列的建模。同时该线性映射结果还作为门控序列，用于调整模型的输出，使模型聚焦于关键输入区域，降低噪声干扰。该设计能够使模型更加专注于对全局序列语义的捕获。

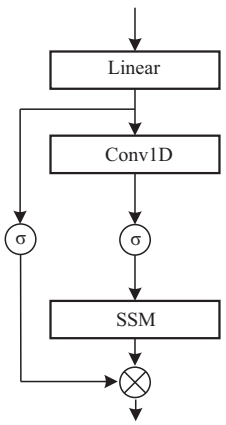


图3　改进后的Mamba模块

2.2　GhostNet模块

为了减少模型的参数量并提升计算效率，本文使用GhostNet来代替常规卷积对输入的统计特征进行特征提取。在常规卷积的结果中，不同通道之间存在大量类似的特征图，GhostNet的核心思想是利用常规卷积仅生成部分关键特征图，随后利用计算成本更低的深度可分离卷积操作生成剩余的冗余特征图，从而达到减少参数量和计算量的目的。常规卷积和Ghost卷积过程对比如图4所示。

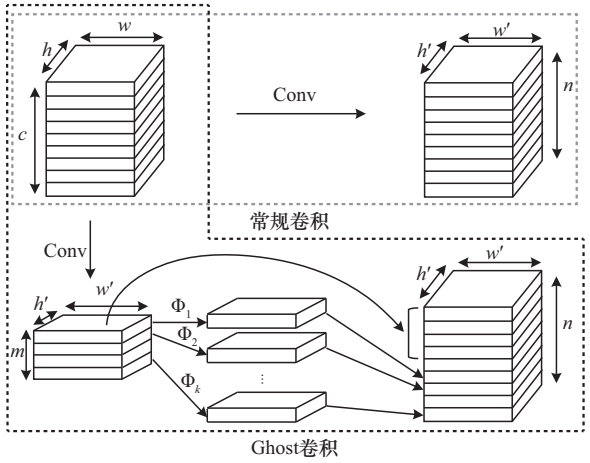


图4　常规卷积和Ghost卷积对比

Ghost卷积首先对输入的统计特征矩阵$\boldsymbol{X}\in R^{h\times w\times c}$使用1×1进行常规卷积得到结果$\boldsymbol{Y}'\in R^{h'\times w'\times m}$其中，$m=n/s$，$s$是每个基础特征要进行线性变换的次数。接着对$\boldsymbol{Y}'$进行深度卷积，计算如下。

$$\boldsymbol{Y}_{ij}''=\boldsymbol{\Phi}_{i,j}(\boldsymbol{Y}_i')=\text{DepthwiseConv}_{d\times d}(\boldsymbol{Y}_i') \quad (17)$$

其中，$\boldsymbol{Y}''\in R^{h'\times w'\times m(s-1)}$，将得到$\boldsymbol{Y}'$和$\boldsymbol{Y}''$进行特征拼接即Ghost卷积结果。本文令深度卷积的卷积核大小d为3，特征变换次数s为2。常规卷积的计算FLOPs如式（18）所示，Ghost卷积第一步和第二步的计算FLOPs如式（19）和式（20）所示。

$$\text{FLOPs}_{\text{Conv}}=h'\times w'\times n\times c\times k^2 \quad (18)$$

$$\mathrm{FLOPs}_{\mathrm{GhostStep1}}=\frac{n\times\left(c\times k^2\times h'\times w'\right)}{2} \quad (19)$$

$$\mathrm{FLOPs}_{\mathrm{GhostStep2}}=\frac{n\times\left(h'\times w'\times 9\right)}{2} \quad (20)$$

由式（19）和式（20）可得，Ghost卷积的运算参数量约为常规卷积的一半。利用Ghost卷积优势，GhostNet模块如图5所示。

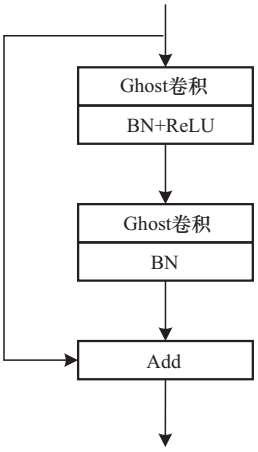


图5 GhostNet模块

该模块由2个串联的Ghost卷积构成，旨在通过通道扩展–压缩机制提升特征表达能力。首先，输入的9×9统计特征矩阵经第1个Ghost卷积层处理，通道数扩展至8以增强特征多样性。随后进行批量归一化（BN）和ReLU激活，稳定特征分布并引入非线性表达能力。接着，通过第2个Ghost卷积层将通道数压缩回1，实现特征蒸馏并抑制浅层冗余信息对建模过程的干扰。根据MobileNetV2的设计经验，在该压缩层中仅保留BN，省略ReLU。最后，输出的特征矩阵与输入矩阵相加，防止梯度发散。

2.3 注意力模块

建模后的有效载荷特征序列长度远大于统计特征序列，长有效载荷特征序列包含大量流量局部细节信息，而短统计特征序列携带全局抽象信息。直接拼接后，全连接层的固定权重矩阵难以区分这2类特征的重要性，短序列统计特征容易被长序列的噪声淹没。因此在输入全连接层进行分类前加上一个自注意力模块对拼接后的序列进行自适应加权处理，这有助于对重要的特征赋予更大的注意力权重，从而提高模型的检测精度。注意力模块如图6所示。处理过程如下：首先对拼接后长度为L特征序列$\boldsymbol{X}$进行异构子空间映射，通过线性投影生成查询矩阵$\boldsymbol{Q}$、键矩阵$\boldsymbol{K}$、值矩阵$\boldsymbol{V}$。随后，通过跨节点注意力建模计算序列单元间的相关性，得到注意力分数。计算序列节点i和节点j之间的注意力分数的计算式如下。

$$e_{ij}=\frac{\boldsymbol{Q}_i\boldsymbol{K}_j^{\mathrm{T}}}{\sqrt{d_k}}(i,j\in[1,L]) \quad (21)$$

其中，d_k为键的子空间维度，该操作可以衡量序列节点i与j的时空关联强度。为了动态抑制异常特征值的非线性畸变同时增强注意力机制对不同攻击模式的适应性，本文采用PReLU函数对注意力分数进行非线性变换，计算式如下。

$$e'_{ij}=\mathrm{PReLU}(e_{ij})=\begin{cases}e_{ij}, & e_{ij}\geqslant 0\\ a\cdot e_{ij}, & \text{其他}\end{cases} \quad (22)$$

其中，a为可学习参数。然后通过Softmax函数对分数进行归一化得到注意力权重。

$$\alpha_{ij}=\mathrm{Softmax}(e'_{ij})=\frac{\exp(e'_{ij})}{\sum_{k=1}^{L}\exp(e'_{ik})} \quad (23)$$

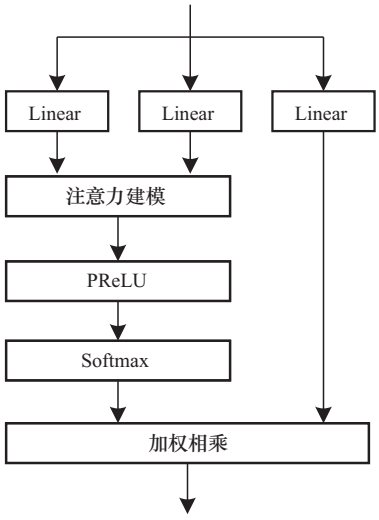


图6 注意力模块

最后通过注意力权重与值向量相乘生成加权后的输出特征$\boldsymbol{Z}$，计算式如下。

$$\boldsymbol{Z}_i = \sum_{j=1}^{L} w_{ij} \boldsymbol{V}_j \tag{24}$$

3　实验与分析

本实验基于Linux环境下的PyTorch深度学习框架开展，采用Python 3.8版本及CUDA 12.0进行高效计算加速，硬件平台选用NVIDIA GeForce RTX 4080 SUPER显卡，内存为16 GB。为避免实现层面的加速策略对耗时结果造成干扰，未额外启用混合精度训练与编译级优化，数据加载采用标准DataLoader配置，并保持各对比方法使用一致的数据预处理与加载流程。

3.1　评价指标

为了评估模型在不同场景下的表现，本文采用了多种常见的性能分类指标，包括准确率(Accuracy)、精确率（Precision）、召回率（Recall）、F1值（F1-score）以及时间开销。具体的计算式如下：

$$\mathrm{Acc} = \frac{\mathrm{T_p} + \mathrm{T_N}}{\mathrm{T_p} + \mathrm{F_N} + \mathrm{F_P} + \mathrm{T_N}} \tag{25}$$

$$\mathrm{Precision} = \frac{\mathrm{T_p}}{\mathrm{T_p} + \mathrm{F_P}} \tag{26}$$

$$\mathrm{Recall} = \frac{\mathrm{T_p}}{\mathrm{T_p} + \mathrm{F_N}} \tag{27}$$

$$\mathrm{F1\text{-}score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{28}$$

其中，$\mathrm{T_p}$表示被正确分类为正类的样本数，$\mathrm{T_N}$表示被正确分类为负类的样本数；而$\mathrm{F_P}$是指被错误分类为正类的负样本数，$\mathrm{F_N}$则表示被错误分类为负类的正样本数。此外，为更直观地量化模型的高效特性，本文在时间开销指标之外统计模型可训练参数量P（单位：K），并定义综合效率指标（efficiency cost index，ECI）为ECI=P×T，其中，T为单样本平均检测延迟（ms，batch=1）。ECI越小表示模型在参数规模与推理时延上越高效。

3.2　数据集描述

CICDDoS2019数据集作为网络安全领域的权威基准，由加拿大网络安全研究所（CIC）与新布伦瑞克大学联合构建，该数据集包括12个攻击类别和1个正常流量类别。原始包数据以PCAP格式存储，完整保留网络层至应用层的协议头部与负载信息，支持细粒度流量解析与行为分析。原数据集中部分样本存在时间戳紊乱、字段缺失或包含异常字符等问题。为保证样本质量，删除了部分格式不规范的样本。CICDDoS2019数据集样本数据见表1。

表1　CICDDoS2019数据集样本数据

日期	正常流量/个	攻击流量/个	总数/个
1月12日	56 000	20 100 500	20 156 500
3月11日	56 863	50 006 249	50 063 122

鉴于原始数据集样本量庞大，为提高实验的可行性与效率，本文仅选取1月12日的流量数据进行处理和分析。由于WebDDoS攻击类样本数量极少，难以支撑有效建模与分析，因此在实验中不考虑该攻击类型。此外，为了进一步评估模型的泛化能力，本研究还选取了Maple-IDS、CICIDS2017及CICIDS2018这3个经典数据集进行实验。

3.3　实验结果

针对不同截断包数M和有效载荷字节长度N的配置，模型的检测效果存在差异。本文设置实验分析不同流量截断长度（128 byte、256 byte、384 byte、512 byte、1 024 byte）与包数（20~30）的组合下检测准确率的变化趋势。不同截断包数和有效载荷长度下检测准确率如图7所示。截断长度为256 byte时在多个截断包的情况下表现最佳，尤其是当截断包数为26时，分类准确率达到最高99.56%。这是因为这一配置能够在提供充足流量信息的同时有效避免冗余信息的引入。相比之下，截断长度为128 byte时，由于无法提供足够的流量信息，其分类准确率普遍较低。而截断

长度为512 byte和1 024 byte时，虽然包含更多信息，但过多的冗余反而降低了分类的准确率。随着截断包数增加，模型能够获取更多的信息，然而当截断包数目超过25后，准确率提升趋于饱和甚至略有下降，表明引入了过多的冗余特征，反而对分类效果产生干扰并显著增加计算开销。综上所述，本文选取的截断包数M为26，有效载荷字节N为256。

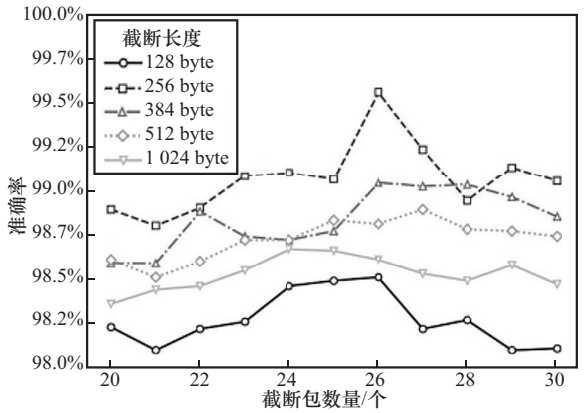


图7　不同截断包数和有效载荷长度下检测准确率

在实验中选择Adam作为优化器，并使用交叉熵作为损失函数。学习率是模型能否高效收敛到高质量解的重要超参数。为确定合适的学习率，本文分别将其设置为10^{-2}、10^{-5}、10^{-6}和10^{-7}进行对比实验。不同学习率下的损失值如图8所示。结果表明，学习率为10^{-2}时损失曲线存在明显波动，说明训练过程不稳定。学习率为10^{-7}时学习率过小，模型收敛缓慢且最终损失较高。学习率为10^{-6}时在稳定性方面表现良好，但收敛速度仍较慢。综合比较下，学习率设置为10^{-5}时模型收敛速度较快且过程稳定。因此本文将10^{-5}模型训练的初始学习率设置为10^{-5}。

为评估不同批处理大小对模型性能的影响，本文分别设置批处理大小为8、16和32进行对比实验。不同批处理大小下的损失值如图9所示。结果表明，当批处理大小设置为8时，模型在训练过程中虽然能获得更细粒度的更新，但收敛速度较慢且波动较大。批处理大小为32时，在训练初期损失值下降最快，但在25轮后陷入局部极小值状态，直到120轮后损失值才再次下降，表明整体收敛速度较慢。相比之下，批处理大小为16时模型在训练初期即可稳定下降至较低的损失值，并在后续训练过程波动不超过0.001，表明模型已基本收敛。因此，本文将模型的批处理大小设置为16。

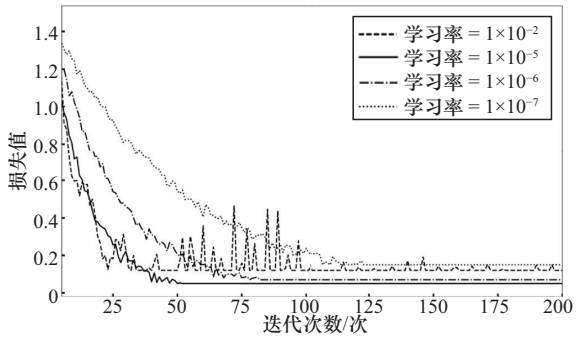


图8　不同学习率下的损失值

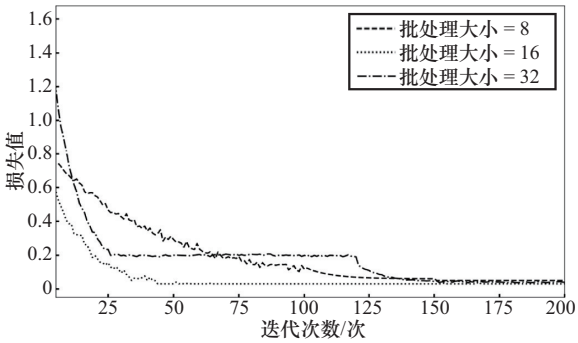


图9　不同批处理大小下的损失值

为了验证模型的性能，这里使用本文模型和其他深度学习模型在CICDDoS2019数据集上进行实验。本文对数据集进行预处理后按照数据采集时间进行划分。具体而言，将所有样本按时间戳升序排序后，依次划分为训练集70%、验证集10%与测试集20%。在该设置下，模型仅使用历史数据进行训练与调参，最终在未来时间段的测试集上进行评估，从而保证结果的可泛化性与评估的公平性。进一步地，本文以flow_id、session_id、attack_id为分组单位进行划分，确保同

一组样本不会同时出现在不同子集中，以避免重复样本带来的数据泄露。在验证集上所有的 epoch 中将 F1 值最高的模型保存为最终模型。不同模型的评估指标见表 2。

本文模型在精确率、召回率和 F1 值上分别达到了 99.62%、99.57% 和 99.59%，均优于其他对比模型。较次优模型 AC-MKNet 分别提升 1.49%、1.59% 和 1.54%。同时，检测耗时仅需 0.21 ms/条，与最快的 CNN-LSTM 模型几乎持平，但 F1 值提高了 8.19 个百分点。Transformer 与 AC-MKNet 的 ECI 分别达到 10.63 与 11.02，反映出其较大的模型规模与较高的推理耗时。相比之下，本文模型的 ECI 为 1.09，在深度学习方法中处于较低水平，同时又取得最高的检测性能，说明本文模型在模型容量与推理速度之间实现了更优的平衡，体现了其高效的特性。这证明本文模型突破了传统基于时序特征建模方面的局限性，也有效克服了 Transformer 类模型所面临的高计算复杂度问题。此外，传统的机器学习模型虽然在推理速度上占优，但其检测性能受限于对有效载荷序列与协议字段语义的刻画能力不足，难以满足多类别高精度检测需求。不同模型对各个类型流量的检测准确率见表 3。

实验结果表明，整体上，传统机器学习方法在多类攻击场景下表现明显弱于深度学习方法，其加权平均准确率分别为 89.05%、86.80% 和 84.95%，说明仅依赖统计特征的传统模型对协议

表 2　不同模型的评估指标

模型	精确率	召回率	F1 值	参数量/K	耗时/ms	ECI/(K·ms)
RNN-AE[20]	90.72%	90.43%	90.57%	2.10	0.47	0.99
CNN-LSTM[21]	91.54%	91.27%	91.40%	1.85	0.19	0.35
CNN-BiSRU[22]	92.91%	92.74%	92.82%	2.65	0.35	0.93
Transformer[23]	96.14%	95.82%	95.98%	8.50	1.25	10.63
AC-MKNet[24]	98.13%	97.98%	98.05%	7.20	1.53	11.02
KNN[10]	92.35%	91.80%	92.07%	0.00	0.10	0.00
XGBoost[25]	89.63%	88.50%	89.06%	0.00	0.15	0.00
LightGBM[26]	87.37%	86.21%	86.78%	0.00	0.12	0.00
本文模型	99.62%	99.57%	99.59%	5.21	0.21	1.09

表 3　不同模型对各个类型流量的检测准确率

类别	RNN-AE	CNN-LSTM	CNN-BiSRU	Transformer	AC-MKNet	KNN	XGBoost	LightGBM	本文模型
BENIGN	92.65%	93.56%	96.63%	100%	99.73%	94.80%	92.10%	90.50%	100%
TFTP	96.73%	98.35%	98.37%	98.16%	98.57%	96.20%	94.85%	93.70%	99.87%
DrDoS_NetBIOS	93.73%	97.93%	98.59%	98.96%	97.97%	92.50%	90.30%	88.90%	100%
Syn	98.93%	96.76%	98.95%	98.73%	100%	95.60%	93.45%	91.80%	100%
DrDoS_UDP	68.78%	61.96%	66.72%	87.63%	93.39%	65.30%	60.20%	58.00%	99.35%
DrDoS_MSSQL	75.66%	75.27%	77.95%	87.75%	98.37%	74.50%	72.80%	70.60%	98.75%
DrDoS_NTP	98.68%	97.35%	99.36%	99.59%	99.47%	96.85%	95.20%	93.95%	99.68%
DrDoS_LDAP	73.72%	85.73%	89.35%	92.62%	98.53%	82.00%	78.50%	75.90%	100%
UDP-lag	98.78%	97.95%	96.93%	97.88%	98.93%	95.40%	93.10%	91.45%	99.67%
DrDoS_SSDP	80.76%	82.37%	95.82%	92.73%	96.95%	85.70%	81.50%	79.30%	98.58%
DrDoS_DNS	97.35%	98.35%	99.73%	99.31%	99.35%	96.50%	94.75%	93.20%	99.85%
DrDoS_SNMP	87.38%	88.78%	85.85%	96.24%	96.95%	86.20%	83.90%	81.60%	98.73%
Weighted Average	90.22%	91.50%	92.91%	96.07%	98.08%	89.05%	86.80%	84.95%	99.56%

敏感型与语义相关攻击的区分能力有限。进一步从不同攻击类型的结果来看，传统基于流统计特征的模型在协议敏感型攻击检测中存在明显的局限性。例如，针对DrDoS_UDP攻击，基于统计特征的模型最高准确率仅为87.63%，而本文模型和AC-MKNet能够通过解析有效载荷中UDP反射包长度异常等协议字段特征，将检测精度提升至99.35%和93.39%。对于DrDoS_LDAP攻击，基于统计特征的模型最高准确率为92.62%，本文模型则通过识别LDAP嵌套查询结构等语义特征实现100%的精准检测。在更复杂的交互型攻击场景中，本文模型同样表现优异，在DrDoS_MSSQL和DrDoS_SSDP检测中分别取得了98.75%和98.58%的准确率，较AC-MKNet模型分别提升0.38%和1.63%。即便在以统计特征为主的Syn洪水攻击检测中，本文模型仍能保持100%的准确率，充分证明了特征融合机制的高度自适应性。本文模型在12类攻击中达到99.56%加权平均准确率，且单类检测精度均高于99.3%，证明了所提方法具有出色的检测性能。

3.4 消融实验

为验证各个模块对检测效果的影响，本文设计了模块消融实验。其中，M1代表有效载荷特征处理分支，M2代表统计特征处理分支，M3代表自注意力模块。模块消融实验结果见表4。各模块对检测效果的提升有明显的协同效应。单独使用M1和M2分别能达到93.35%和95.43%的准确率，说明载荷语义特征与统计特征各自具备一定的判别能力。同时使用M1和M2后准确率提升至99.14%，验证了统计特征和有效载荷信息融合的互补优势。进一步引入M3后，模型仅付出较小的额外计算开销就实现了99.56%的准确率。这证明了自注意力模块的有效性。

为验证所提出的Mamba结构精简策略的有效性，在保持统计特征分支与跨节点自注意力（M3）不变的情况下，仅对有效载荷分支中的Mamba块进行结构对比。设定4种变体：V0，原始Mamba；V1，去除Conv1D；V2，去除线性投影；V3，同时去除Conv1D与线性投影。所有变体采用相同的数据划分与训练配置，在相同测试环境下统计参数量、FLOPs、准确率及单样本推理耗时。原始Mamba与改进Mamba的结构消融实验结果见表5。

表4 模块消融实验结果

模型结构	准确率	耗时/ms
M1	93.35%	0.11
M2	95.43%	0.08
M1+M2	99.14%	0.12
M1+M2+M3	99.56%	0.21

表5 原始Mamba与改进Mamba的结构消融实验结果

模型结构	参数量/K	FLOPs	准确率	耗时/ms
V0	5.21	0.96	99.53%	0.35
V1	4.62	0.58	99.52%	0.24
V2	3.74	0.75	99.54%	0.27
V3	3.15	0.52	99.56%	0.21

由表5可见，相比原始Mamba，仅移除Conv1D的V1在准确率几乎不变的情况下，FLOPs和推理耗时明显降低，说明在本文特征输入形式下Conv1D的局部卷积带来的收益有限且存在冗余。仅移除线性投影的V2导致准确率略有提升且显著降低了参数量，表明该投影层同样可能引入冗余映射。综合精简后的V3在保持最高Accuracy的前提下取得最佳效率。结果表明，移除Mamba架构中的Conv1D与线性投影能够在有效提取有效载荷特征的同时，显著提升了模型的性能。

为验证9×9小尺度统计特征矩阵输入下，采用Ghost卷积是否合理，在保持主干网络结构、训练策略与其余模块完全一致的前提下，仅替换统计特征分支中的卷积单元进行对比实验，具体设置为：C1采用Ghost卷积，C2采用常规卷积，C3采用深度可分离卷积。不同卷积方式消融实验结果见表6。

表6　不同卷积方式消融实验结果

模型结构	参数量/K	FLOPs	准确率
C1	5.12	0.16	99.56%
C2	12.61	0.42	98.21%
C3	4.77	0.15	96.05%

从结果可以看出，在9×9的小尺度输入条件下，常规卷积并未带来明显的性能优势，但其参数量与计算量显著更高。深度可分离卷积进一步压缩了复杂度，但在准确率上存在一定回落，说明该分支仍需要一定的跨通道表达与局部联合建模能力。相比之下，Ghost卷积在保持较低复杂度的同时取得了最佳准确率，表明统计特征矩阵中依然存在可被高效利用的特征冗余，Ghost通过主特征和廉价生成的方式在轻量化与表征能力之间取得了折中。实验结果表明，本文在统计特征分支中采用Ghost效果最佳。

4　结束语

针对软件定义网络（SDN）环境中DDoS攻击流量特征语义信息利用不足以及检测模型精度与效率难以兼顾的问题，本文设计了GhostMamba-SAN混合检测模型。该模型通过改进的Mamba对有效载荷特征进行高效建模，通过Ghost卷积的方式提取流级统计特征，最后通过自注意力机制对两类特征进行加权融合，实现高效分类。在CICDDoS2019数据集上的实验结果表明，本文模型在检测准确率和效率上均优于传统基于统计特征的检测模型，展示了较强的综合性能。此外，本文模型在Maple-IDS、CICIDS2017及CICIDS2018数据集上亦取得了较高的检测指标，体现了良好的泛化性。

然而，当前的GhostMamba-SAN模型仍存在一些不足之处。在复杂网络环境和未知攻击场景下的鲁棒性仍有待进一步提升。其在应对动态流量分布变化及新型攻击特征时的自适应性尚显不足。未来的研究将引入多模态特征融合与时序自适应机制，以增强模型对异构流量和新型攻击模式的识别能力。同时扩展训练数据集，从而进一步提升模型在真实场景下的检测稳定性与泛化性能。

参考文献：

[1] 董仕 . 软件定义网络安全问题研究综述[J]. 计算机科学, 2021, 48(3): 295-306.
Dong S. Survey on software defined networks security[J]. Computer Science, 2021, 48(3): 295-306.

[2] Akamai. Facing the surge of security threats: attack trends in the financial services industry [EB]. (2024-12-05).

[3] Neres Carvalho R, Luiz Bordim J, Adilio Pelinson Alchieri E. Entropy-based DoS attack identification in SDN[C]//Proceedings of the 2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW). Piscataway: IEEE Press, 2019: 627-634.

[4] Hemmati Z, Mirjalily G, Mohtajollah Z. Entropy-based DDoS attack detection in SDN using dynamic threshold[C]//Proceedings of the 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS). Piscataway: IEEE Press, 2021: 1-5.

[5] Ali Ujjan R M, Pervez Z, Dahal K, et al. Entropy based features distribution for anti-DDoS model in SDN[J]. Sustainability, 2021, 13(3): 1522.

[6] Li R Y, Wu B. Early detection of DDoS based on φ-entropy in SDN networks[C]//Proceedings of the 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Piscataway: IEEE Press, 2020: 731-735.

[7] Yang Z, Han L. Research on DDoS attack detection and vulnerability mechanism based on Entropy of destination IP address in SDN environment[J]. Journal of Tianjin University of Technology, 2020, 36(4): 39-44, 59.

[8] Alhamami K, Albermany S. DDoS attack detection using machine learning algorithm in SDN network[C]//Proceedings of the 2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT). Piscataway: IEEE Press, 2023: 97-102.

[9] Ribeiro M A, Pereira Fonseca M S, De Santi J. Detecting and mitigating DDoS attacks with moving target defense approach based on automated flow classification in SDN networks[J]. Computers & Security, 2023, 134: 103462.

[10] Dong S, Sarem M. DDoS attack detection method based on improved KNN with the degree of DDoS attack in software-defined networks[J]. IEEE Access, 2019, 8: 5039-5048.

[11] Li C H, Wu Y, Qian Z Z, et al. DDoS attack detection and defense based on hybrid deep learning model in SDN[J]. Journal on Communications, 2018, 39(7): 176-187.

[12] Mohammad L, Mahdi J S, Ramin S H Z, etd. Deep packet: a novel approach for encrypted traffic classification using deep

learning[J]. Soft Comput. 24, 3 (Feb 2020), 1999 – 2012. 10.1007/s00500-019-04030-2

[13] Hosseini S M, Jahangir A H. An effective payload attribution scheme for cybercriminal detection using compressed bitmap index tables and traffic downsampling[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(4): 850-860.

[14] Sohi S M, Seifert J P, Ganji F. RNNIDS: Enhancing network intrusion detection systems through deep learning[J]. Computers & Security, 2021, 102: 102151.

[15] Lotfollahi M, Jafari Siavoshani M, Shirali Hossein Zade R, et al. Deep packet: a novel approach for encrypted traffic classification using deep learning[J]. Soft Computing, 2020, 24(3): 1999-2012.

[16] Gu A, Dao T. Mamba: linear-time sequence modeling with selective state spaces[PP]. arXiv(2023-11-01) [2024-05-01]. arXiv: 2312.00752,2023.

[17] He W, Han K, Tang Y H, et al. DenseMamba: state space models with dense hidden connection for efficient large language models[PP]. V2. arXiv (2024-03-05) [2024-05-01]. 10.48550/arXiv.2403.00818.

[18] Li K, Chen G, Yang R X, et al. SPMamba: State-space model is all you need in speech separation[PP]. V2. arXiv (2024-09-10) [2024-05-01]. arXiv: 2404.02063.

[19] Qiao Y Y, Yu Z, Guo L T, et al. VL-mamba: exploring state space models for multimodal learning[PP]. arXiv (2024-03-20) [2024-05-01]. arXiv: 2403.13600.

[20] El Sayed M S, Le-Khac N A, Azer M A, et al. A flow-based anomaly detection approach with feature selection method against DDoS attacks in SDNs[J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(4): 1862-1880.

[21] Zainudin A, Ahakonye LAC, Akter R, et al. An efficient hybrid-DNN for DDoS detection and classification in software-defined IIoT networks[J]. IEEE Internet of Things Journal, 2023, 10(10): 8491-8504.

[22] Cao L, Wen M, He W, et al. Deep learning based dos and ddos attack detection method in the highway monitoring system of iov[J]. Computer Applications and Software, 2025, 42(1): 303-311.

[23] Wang H, Li, W. DDosTC: a transformer-based network attack detection hybrid mechanism in SDN[J]. Sensors, 2021(21): 5047.

[24] Diallo A F, Patras P. Adaptive clustering-based malicious traffic classification at the network edge[C]//Proceedings of the IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2021: 1-10.

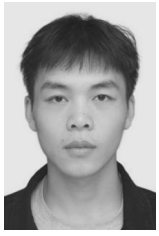
[25] Chen T Q, Guestrin C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2016: 785-794.

[26] Ke g, Meng q, Finley T, et al. LightGBM: A highly efficient gradient boosting decision tree[C]//Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017: 3146-3154.

[作者简介]

包晓安（1973– ），男，浙江理工大学计算机科学与技术学院教授，主要研究方向为网络安全、软件可靠性和深度学习。

杨奉豪（2001– ），男，浙江理工大学计算机科学与技术学院硕士生，主要研究方向为网络与系统安全、网络流量分类。



范云龙（2000– ），男，浙江理工大学计算机科学与技术学院硕士生，主要研究方向为网络安全、网络流量分类。

涂小妹（1995– ），女，浙江广厦建设职业技术大学城乡建设学院讲师，主要研究方向为多模态网络入侵检测、信息处理。

胡天缤（1998– ），女，河海大学人工智能与自动化学院博士生，主要研究方向为人工智能、软件定义网络。

吴彪（1989– ），男，博士，浙江理工大学理学院讲师，主要研究方向为软件定义网络、深度学习。