一种基于双层模型和指标分布的恶意网络流持续检测和分类方法

陆浩天,董育宁*,全宇轩
(南京邮电大学通信与信息工程学院,江苏南京 210003)

摘 要: 开集恶意流量识别在网络安全领域发挥着重要的作用. 现有文献方法存在模型结构单一,缺乏灵活性;忽视增量训练样本选择,造成分类性能欠优等问题. 针对这些问题,本文提出了一种基于双层模型和指标分布的恶意网络流持续检测和分类方法. 该方法基于可扩展极限学习机(Scalable Extreme Learning Machine,S-ELM)输出权重与标准输出的关系,设计了改进的最接近皮尔森相关系数、归一化相对方差和归一化"其他"列距离这三个指标,通过相乘最终得到一个综合指标,并结合单分类器来进行未知类检测. 为了提高S-ELM在开集识别任务中的连续增量能力,设计了基于综合指标分布的样本筛选方法,选择最优增量训练样本集. 与代表性文献方法的对比实验表明,本方法的未知类检测NA指标能改善3%~13%,持续增量更新后的分类Acc性能可以提高约3%~7%.

关键词: 网络流量分类;入侵检测系统;开放集识别;未知类检测;增量学习;极限学习机
基金项目: 国家自然科学基金(No.61271233);江苏省研究生科研创新计划(No.KYCX23_1031)
中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112(2025)05-1637-13
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250069

A Method for Continuous Detection and Classification of Malicious Network Traffic Based on Double-Layer Model and Distribution of Indexes

LU Hao-tian, DONG Yu-ning*, QUAN Yu-xuan
(*College of Telecommunications and Information Engineering*, *Nanjing University of Posts and Telecommunications*, *Nanjing*, *Jiangsu* 210003, *China*)

Abstract: Open-set malicious traffic recognition plays an important role in the field of network security. Existing methods have limitations in single model structure and lack of flexibility; neglecting incremental training samples selection, resulting in suboptimal classification performance. To address these problems, this paper proposes a method for continuous detection and classification of malicious network flows based on double-layer model and index distribution. Based on the relationship between the output weights of scalable extreme learning machine (S-ELM) and the standard output, this method designs following three indexes: the improved closest Pearson's correlation coefficient, the normalized relative variance, and the normalized distance to "the others" column. These indexes are multiplied together to obtain a comprehensive index, which is combined with a single classifier for unknown class detection. In order to improve the continuous incremental capability of S-ELM in the open-set recognition task, a sample selection method based on the distribution of the comprehensive index is developed to select the optimal sub-dataset for incremental model training. Comparison experiments with existing representative methods show that the NA index of unknown class detection of the proposed method can be improved by 3%~13%, and the classification Acc index can be enhanced by about 3%~7% after continuous incremental updating.

Key words: network traffic classification; intrusion detection system; open-set recognition; unknown class detection; incremental learning; extreme learning machine

Foundation Item(s): National Natural Science Foundation of China (No.61271233); Postgraduate Research & Practice Innovation Program of Jiangsu Province (No.KYCX23_1031)

收稿日期:2025-01-17;修回日期:2025-04-13;责任编辑:赵克
*通讯作者:董育宁

1 引言

随着网络应用程序的不断迭代更新,大量新型应用程序和网络攻击相继出现,导致网络流量的数量和种类显著增加[1]. 据《2024 年应用安全趋势报告》[2]最新数据显示,全球将近 7% 的网络流量是恶意的,而2023 年报告数据为 6%,这表明如今恶意流量攻击呈现出上升趋势,网络安全形势严峻. 网络流量分类(Network Traffic Classification,NTC)和入侵检测系统(Intrusion Detection System,IDS)[3]在网络安全和管理中起着至关重要的作用[4]. 因此,开发先进的入侵检测系统以应对快速增长的网络恶意攻击,特别是未知的新型攻击带来的挑战,已成为当前研究的热点之一.

现有的 NTC 方法通常在封闭数据集上训练分类器[5,6]. 然而,这些方法未能适应开放网络环境的动态特性. 在这种环境中,不断受到已知应用程序、新兴(未知)应用程序,甚至恶意流量的影响,这些流量相互交织,形成复杂的流量模式[7].

开放集识别(Open Set Recognition,OSR)[8]要求分类器不仅需要准确分类已知类别,还需要有效检测未知类别. 最新研究[9,10]尝试在物联网设备指纹识别中检测未知设备,但这些方法并不适用于未知攻击检测. 鉴于未知攻击的多样性和动态性,其分布特征难以精确捕捉,从而使得细粒度攻击检测中的开集识别成为了一个棘手难题. 目前针对这一问题的解决方案在效果上尚不尽如人意. 具体而言,(1)OpenIDS[11]和 CVAE-EVT[12]基于极值理论检测未知攻击,但当真实未知攻击的分布仅部分符合极值理论假设时,检测效果显著下降. (2)部分方法[13]引入辅助样本以确定已知/未知类别边界,然而其分类性能高度依赖辅助样本的质量. 此外,该类方法在某些复杂网络环境中可能会引发负面效应,例如增加检测设备负载、降低系统稳定性,甚至干扰正常流量的识别. 同时,这些算法在需要快速响应和增量更新的场景下往往表现不佳,难以适应网络攻击的持续变化. 针对上述挑战,本文提出了一个结合多维度考量的综合指标,以应对错综多变的网络攻击场景. 值得注意的是,现有大多数方法仍依赖人工设定阈值,易受人为经验限制. 为此,引入单分类器结构,能够自适应不同攻击场景,不仅避免了繁琐的手动调参过程,还提高了未知攻击检测的精度.

随着 OSR 任务的拓展,仅仅检测未知类已不足以应对持续增加的各种网络应用. 因此,需要研究基于增量学习的持续识别能力[14]. 在开放集环境下的增量学习中,现有方法[15]更加注重未知类的检测,在模型更新时仅用聚类分配的标签进行增量,忽略了对用于增量学习样本的优化,从而导致增量过程的误差累积,进而影响整体的分类性能. 部分方法甚至默认为未知类别分配正确标签[9]. 文献[13]提出将检测到的新攻击类别交由网络安全管理员进行分析并手动标注细粒度标签,但这一流程在实际应用中过于依赖人工干预,难以满足大规模、实时检测的需求. 为此,本文通过分析检测未知攻击流量的综合指标,优先筛选误检率较低的样本用于增量更新训练,从而降低累积误差对模型性能的影响. 此外,基于深度学习的方法在训练过程中,需要通过反向传播算法迭代更新每一层的权重. 由于参数量通常较大,计算资源的需求也较高. 同时考虑到这类方法在增量学习时容易出现灾难性遗忘问题[16],引入了可扩展极限学习机(Scalable Extreme Learning Machine,S-ELM)[17]与单分类器构建双层模型架构,结合 S-ELM 的快速训练优势与单分类器的灵活适应能力,在增强检测实时性的同时,能有效提升类增量更新后的模型性能.

本文的主要贡献总结如下:

(1)提出了一种结合双层模型和指标分布的恶意网络流检测和分类方法. 该方法通过将模型输出权重与标准输出的关系相融合,引入了一种基于 S-ELM 分类模型的统计指标,同时与单分类器构成双层模型,实现未知类检测. 该方法能够有效提高未知类检测能力,并减少分类时间.

(2)设计了基于指标分布的样本筛选(Sample Selection,SS)方法. 通过综合指标在不同区间上的分布,对检测为未知类的样本进行筛选,得到用于增量更新的最优样本集,从而改善 S-ELM 模型的增量更新效果.

(3)在两个实际网络数据集上与现有方法进行了性能对比. 实验结果表明,与代表性文献方法相比,本文方法的未知类检测 NA 指标能改善 3%~13%,持续增量更新后的模型 Acc 指标可以提高约 3%~7%.

2 相关工作

目前,基于机器学习和深度学习的 NTC 方法已经被用来解决 OSR 问题.

2.1 未知流量检测相关

在 OSR 的框架下,学者们持续探索能够同时实现未知攻击检测与已知攻击细粒度分类的入侵检测方法. 然而,鉴于未知类别信息的不可预测性,传统基于单分类器[18]或单一阈值[19]的方法往往存在已知/未知类别决策边界模糊的问题,导致误分类率上升.

针对这一挑战,现有研究提出了多种改进方案,但均存在局限性. Yang 等人[20]通过对比学习构建低维潜在空间的距离度量函数,该函数能有效放大不同类别样本间的距离. 尽管如此,当未知攻击的分布变得复

杂时,其检测性能显著下降. 文献[12]则尝试通过极值理论对重建误差分布进行建模,提出了一种已知/未知入侵检测的两阶段学习方法. 然而,基于极值理论的方法通常难以在不同场景下保持稳定性能. Liang等人[15]提出了一种支持模型增量更新的双层应用分类系统(Double-layer Application Classification Scheme, DACS),其核心思想是对已知类别进行分组,并分别训练两个随机森林(Random Forest,RF)分类器,然后基于两个上级支持向量机(Support Vector Machine,SVM)的投票结果来检测未知类别. 虽然DACS实现了较高的检测效率,但其模型框架对已知类别的数量有额外的要求.

上述方法的性能在很大程度上依赖于模型对已知类别的学习成效. 为了突破这一局限性,一些研究者提出利用辅助样本来增强对未知类别的检测能力. Zhao等人[21]提出了一种基于原型学习的开集识别方法(Prototyping and Adaptive Thresholding,PAT). 该方法采用流形混合技术模拟未知样本特征,并通过虚拟原型和自适应阈值有效区分已知类和未知类. 此外,PAT还通过原型分类器仅保留原型和少量样本,从而缓解了灾难性遗忘的问题. Zhong等人[13]则设计了一种基于双鉴别器的生成对抗网络,通过生成对抗训练模拟未知攻击的分布,并将其与已知攻击分布共同输入异构集成模型,该方法在保持已知类细粒度分类能力的同时增强了未知类检测鲁棒性. 然而,这些方法的有效性取决于生成结果对真实未知空间的拟合能力,并且在实际应用中能否保持稳定性能仍然面临挑战.

现有未知流量检测的方法通过挖掘更多的已知类样本信息或者基于预设的规则进行流量分类,这在一定程度上提高了检测的准确性. 然而,随着网络攻击手段的不断演变和复杂化,单一模型往往难以适应多变的网络环境,从而导致模型检测性能的下降. 为此,本文提出了一种基于S-ELM的综合指标,并将其与单分类器结合构成双层模型,以进行未知类检测. 相较于依赖不可预测的生成样本方法,本文结合多个指标进行检测,提升了其在实际部署场景中的适用性,同时能够更有效地适应复杂多变的网络环境.

2.2　增量学习相关

闭集环境的增量学习是在已有真实标签的数据环境中,使用真实标签标记的新类别样本对模型进行更新,主要是为了验证模型具备增量更新的能力. 郭虎升等人[22]为解决在线模型对于概念漂移响应能力较差的问题,引入增量学习器,该模型能够随着新样本的实时流入进行动态的增量更新,从而有效地提取流数据的全局分布信息. 然而,这类研究主要聚焦于应对概念漂移现象[23],与本文探讨的研究场景和目标有所不同.

在开集增量学习中,模型需先检测未知类别,再基于检测到的未知类别进行类增量学习. 对于恶意攻击场景下的开集增量学习,除了需应对闭集增量学习中常见的灾难性遗忘等挑战外,还需要实时检测出未知攻击流量并及时更新模型,同时筛选出可能存在的误检未知攻击流样本. 然而,现有方法通常理想化地假设用于增量更新的未知类别样本完全准确,这与实际应用中的复杂性不符. Sheng等人[24]提出了一种基于密度的启发式聚类方法的自增长攻击流量分类模型,可以持续自动地实时检测和区分不同类型的未知攻击流量. 但是该模型在训练和识别阶段的耗时较长,且其识别性能存在一定的局限性. Zhou等人[9]设计了一个基于随机树的新类检测器,对于数据流中的未标记样本,该检测器可以有效地检测难以识别的新类,并对半监督流数据中的已知类实例进行分类. Fan等人[10]提出了一种名为AutoIoT的新型物联网设备识别模型,该模型能够在引入新设备时自动进行更新,其未知设备检测算法主要依赖于KS(Kolmogorov-Smirnov)检验. 值得注意的是,这两种方法均针对物联网设备的流量数据进行定制,在捕捉未知攻击分布方面存在不足.

尽管上述开集增量学习的研究试图在检测未知类别的同时利用这些样本进行增量学习,但是未充分考虑样本本身的质量,忽视了样本选择的重要性. 为此,本文设计了一种基于指标分布的样本筛选方法,旨在挑选出更具代表性的增量样本. 同时采用轻量级模型S-ELM,以确保模型的快速更新.

3　本文方法

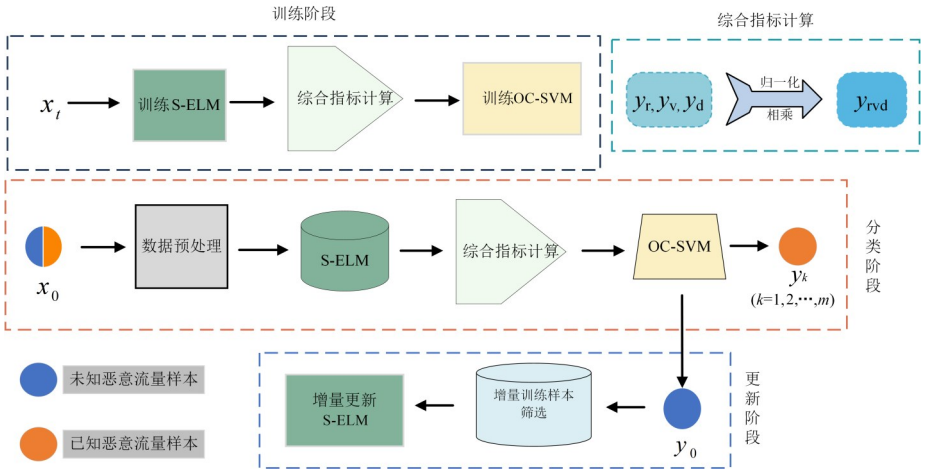
3.1　总体框架

如图1上半部分所示,训练阶段以已知的网络攻击流量样本x_t为输入,训练S-ELM模型,并根据输出计算得到综合指标y_{rvd}. 该指标由三个子指标y_{r}、y_{v}和y_{d}归一化相乘而得,具体计算方式详见3.4节. 随后,利用该综合指标训练单类支持向量机(One Class-SVM, OC-SVM)[25],以便在后续阶段识别未知攻击流量.

在分类阶段(图1中间部分),当面对已知与未知恶意流量攻击的交织挑战时,采用了由S-ELM与OC-SVM组成的双层模型进行区分. 此阶段,系统接收在线采集的网络流量样本x_0. 这些样本经过数据预处理,包括特征提取、特征降维和特征选择(详见3.3节),然后输入双层模型进行识别. OC-SVM首先依据y_{rvd}区分已知和未知攻击流,若样本被判定为已知类y_k(k=1,2,…,m),则进一步由S-ELM完成已知攻击流量的精细分类.

图1下部展示了增量更新过程. 当分类阶段检测

到未知攻击样本y_0时，这些样本会进入增量训练样本筛选模块，筛选后的优质样本用于对S-ELM模型进行类增量更新，提升模型对新型网络攻击流的持续检测能力，防止性能退化.



注：x_t是训练集样本；x_0是在线输入的网络流样本；y_k和y_0分别是分类输出的已知类和未知类样本.

图1 基于双层模型和指标分布的恶意网络流持续检测和分类框架

3.2 可扩展学习机S-ELM

与深度神经网络不同，S-ELM由单个隐藏层构成，属于浅层神经网络. 它不依赖多层特征提取，而是通过随机生成的权重实现快速学习，无需反向传播和多轮迭代训练.

S-ELM的模型架构如图2所示，$\boldsymbol{\omega}=[w_1,w_2,\cdots,w_L]$是隐藏层节点上的权值，$\boldsymbol{B}=[b_1,b_2,\cdots,b_L]$是隐藏层节点上的偏差，$x\in\mathbb{R}^{N\times D}$为输入数据，$\boldsymbol{T}=\left[t_1,t_2,\cdots,t_m\right]$为输出预测标签，其中$L$为隐藏层的节点数量，$N$为样本数量，$D$为特征维数，$m$为类别数. 模型的最终输出的公式如下：

$$\boldsymbol{O}=G(w\cdot x+b)\cdot\boldsymbol{\beta}=\boldsymbol{H}\cdot\boldsymbol{\beta} \tag{1}$$

其中，$G(x)$为激活函数，$\boldsymbol{H}$是隐藏层输出，$\boldsymbol{\beta}$为隐藏层输出权值.

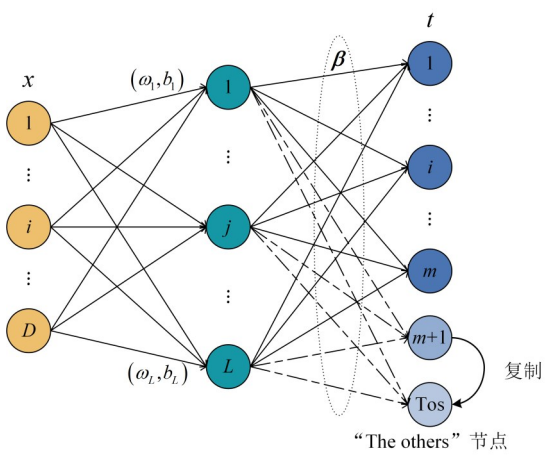


图2 S-ELM网络结构图

由于隐藏层的权重$\boldsymbol{\omega}$与偏置$\boldsymbol{B}$在生成后就不再调整，因此唯一的目标就是找到最优的输出权重$\hat{\boldsymbol{\beta}}$，其可以通过计算隐藏层输出$\boldsymbol{H}$与输出标签$\boldsymbol{T}$的伪逆乘积来获得. 计算公式如下：

$$\hat{\boldsymbol{\beta}}=\arg\min_{\boldsymbol{\beta}}\left\|\boldsymbol{H}\cdot\boldsymbol{\beta}-\boldsymbol{T}\right\|=\boldsymbol{H}^{\dagger}\boldsymbol{T}=\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{T} \tag{2}$$

S-ELM在初始训练阶段会增加一个“其他”(The others，Tos)节点，用于保留已知类作为输出节点的负信息，并在增量过程中，始终保留一个包含模型全部类别的负信息的输出节点. 在训练时，将每个样本在Tos节点的训练标签t_{m+1}均置为−1，即该节点的输出不属于任何已知类.

当增量训练到时刻s时，输出节点个数(即此阶段的类别数)为m_s+1. 假设s+1时刻有c个新类出现，模型需要增加c个节点，则此时的输出权重可以表示为

$$\left[\beta_t\right]_{L\times(m_{s+1}+1)}=\left[\left[\beta_t\right]_{L\times(m_s+1)}\ \left[\triangle\beta_t\right]_{L\times c}\right] \tag{3}$$

其中，$\left[\triangle\beta_t\right]_{L\times c}$为增量添加$c$列后的$\boldsymbol{\beta}$.

3.3 数据预处理

为了加快模型处理在线输入网络流样本的速度，需要进行特征降维和特征选择(Feature Selection，FS).

首先对流量特征进行计算时间复杂度分析，删除复杂度大于$O(n)$的特征子集，其中n表示数据包数量.

随后计算各特征之间以及每个特征与标签之间的皮尔森相关系数(Pearson Correlation Coefficient，PCC)[26]，对于PCC大于0.9的高相关特征对，保留其中与标签PCC值较大的特征，删除相关性较低的另一个.

最后使用极限树特征重要性得分[27]、随机森林特

征重要性得分[28]和特征与标签的PCC得分相加的FS方法,根据得分逐个增加特征,观察分类准确率的变化并寻找拐点,以此获得最优特征子集,对于不同流量数据集,拐点处的特征数量不一定相同,一般在15~30个之间.

3.4 基于S-ELM输出权重指标分布的未知类检测

在基于神经网络的网络流分类任务中,分类结果通常由SoftMax函数转换为概率分布,而未经过激活函数处理的模型输出被称为logits,即神经网络对每个类别的原始预测值. 在知识蒸馏等领域,logits被认为比最终的概率分布更具信息价值[29]. 与直接使用概率输出相比,logits能够更加细致地反映模型对每个类别的判别,尤其是在处理复杂网络流分类任务时,它有助于捕捉类别间的细微差异和潜在的关联. 虽然S-ELM输出权重$\boldsymbol{O}$与logits的语义并不相同,但是都可以被解释为模型对每个类别的"分数". 受这一认识的启发,本方法比较基于S-ELM的输出权重,而非依赖于模型的置信度信息.

一般来说,区分两个权重矩阵的常用方法是通过距离函数. 然而S-ELM输出权重矩阵$\boldsymbol{O}$的维度与数据集类别的数量正相关. 当处理多类别网络流数据集时,随着类别维度的增加,距离函数往往会在高维空间中失效[20]. 因此,为了改善未知攻击检测效果,同时使模型具备增量更新的能力,提出了基于S-ELM输出权重统计分布的未知类检测方法.

由于ELM的输出结果是预测权重,若为该类别,则对应的标准输出权重为1,其余为–1. S-ELM则比ELM多一个输出权重节点,该输出节点对于所有的已知类,标准输出权重均为–1. 根据上述S-ELM的输出权重特性,设计了三个指标以及一个综合指标,用来衡量真实输出权重与标准输出权重的偏差.

(1)改进的最接近皮尔森相关系数(y_r)

对于不同类别样本的真实输出权重$\boldsymbol{X}$,已知类样本的$\boldsymbol{X}$通常更接近对应类别的标准输出权重向量$\boldsymbol{Y}$(定义见下),而当某个样本的$\boldsymbol{X}$值与所有类别的$\boldsymbol{Y}$值均存在较大差距时,可将其判定为未知类样本. 基于这一认识,首先采用PCC来衡量$\boldsymbol{X}$与$\boldsymbol{Y}$之间的相关性,以此作为判断依据. 具体来说,PCC是衡量两个连续变量之间线性相关程度的统计量. 然而,要与每个类别的$\boldsymbol{Y}$计算PCC显然计算量很大,其复杂度为$O(m^2)$. 为此,设计了一个最接近PCC的指标$\mathrm{pcc}_{\max}$. 根据定义,预测标签是由$\boldsymbol{X}$中最大值对应的类别决定的. 为了降低计算开销,选取$\boldsymbol{X}$中与1最接近的输出类别,构造该类别的$\boldsymbol{Y}$(=1),并计算其相应的PCC,取绝对值后得到$\mathrm{pcc}_{\max}$,从而使计算复杂度降至$O(m)$. 计算公式如下:

$$\mathrm{pcc}_{\max}=\left|\frac{\sum_{i}^{N}\left(\boldsymbol{X}_i-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{Y}_i-\overline{\boldsymbol{Y}}\right)}{\sqrt{\sum_{i}^{N}\left(\boldsymbol{X}_i-\overline{\boldsymbol{X}}\right)^2\cdot\sum_{i}^{N}\left(\boldsymbol{Y}_i-\overline{\boldsymbol{Y}}\right)^2}}\right| \tag{4}$$

其中,N为总样本数;$\boldsymbol{X}_i$表示单个样本x_i在S-ELM中的输出权重行向量(即输出权重矩阵$\boldsymbol{O}$的一个行向量);$\boldsymbol{Y}_i$为x_i对应的标准输出权重向量,定义如下:若$\boldsymbol{X}_i$最接近于1的列为j,则令$\boldsymbol{Y}_i[j]=1$,其余的值取为–1;$\overline{\boldsymbol{X}}$和$\overline{\boldsymbol{Y}}$分别是$\boldsymbol{X}$与$\boldsymbol{Y}$的平均值.

然而,$\mathrm{pcc}_{\max}$对数值变化率相同但幅度较大的情况并不敏感,对于某些与$\boldsymbol{Y}$存在明显差异的$\boldsymbol{X}$,$\mathrm{pcc}_{\max}$仍有可能给出较高的相关性值,无法反映实际的差异. 对于未知类的输出,由于其不可预测性并且不同的数据集的输出存在不规律性,单凭$\mathrm{pcc}_{\max}$并不能有效地应对这种不确定性. 为此,引入一个归一化平均绝对偏差因子r,以优化$\mathrm{pcc}_{\max}$的局限性,其计算公式如下:

$$r=2\times\left\{1-\left[1+\exp\left(-\left|\overline{\sum_{i}^{N}\left|\boldsymbol{X}_i-\boldsymbol{Y}_i\right|}\right|\right)\right]\right\} \tag{5}$$

式(5)中采用e的负指数对$\boldsymbol{X}$与$\boldsymbol{Y}$的绝对偏差值进行归一化处理,从而可以更好地捕捉幅度差异的影响. 此改进方案不仅弥补了$\mathrm{pcc}_{\max}$对幅度差异不敏感的不足,还能不受数据集输出模式的影响,更精确地反映已知和未知恶意流量的区别. 最后,y_r表示为

$$y_r=r\times\mathrm{pcc}_{\max} \tag{6}$$

y_r的值越接近于1,表明$\boldsymbol{X}$与$\boldsymbol{Y}$之间的线性相关性越强;反之,则表明两者具有较弱的线性相关性,更像未知类.

(2)归一化相对方差(y_v)

为了进一步捕捉异常和离群的输出权重,引入了归一化相对方差. 具体来说,相对方差是一种衡量数据分散程度的统计量,其值越大表示数据的离散程度越高,数据点相对于平均值的分布越广,并且方差对异常值格外敏感. 在归一化处理后,设计了归一化相对方差y_v以完善y_r对异常值不敏感的缺陷. 计算公式如下:

$$y_v=\exp\left(-\left|\frac{\sum_{i}^{N}\left(\boldsymbol{X}_i-\overline{\boldsymbol{Y}}\right)^2-\sum_{i}^{N}\left(\boldsymbol{Y}_i-\overline{\boldsymbol{Y}}\right)^2}{m}\right|\right) \tag{7}$$

y_v的值越大,代表$\boldsymbol{X}$与$\boldsymbol{Y}$的离散程度越小,越接近标准输出,样本更可能是已知类;反之,更可能是未知类.

(3)归一化Tos列距离(y_d)

为了后续的增量学习,S-ELM增加了Tos列. 该列的设计初衷是将所有已知类样本视为负类,因此已知

类的输出权重设定为–1. 基于这一特性,设计了y_{d}来衡量输出权重与标准结果的偏离距离. 计算公式如下:

$$y_{\mathrm{d}}=\exp\left(-\left|X_T+1\right|\right) \tag{8}$$

其中,X_T为S-ELM输出权重矩阵$\boldsymbol{O}$最后一列.

当y_{d}的值越接近1时,真实Tos输出权重越接近–1,表明样本可能属于已知类;反之,样本更可能是未知类.

为了更直观地展示上述三个指标在区分已知与未知恶意流量输出权重方面的有效性,将它们分别映射为三维坐标系中的三个坐标轴,绘制了已知类与未知类的三维空间分布图,如图3所示. 由图3可知,随着三个指标值趋近于1,已知类的数据点呈现出显著的聚集趋势;与此同时,指标值靠近0时,未知类的数据点在空间中更为集中,反映了这些指标在区分已知和未知类时的有效性与可靠性.

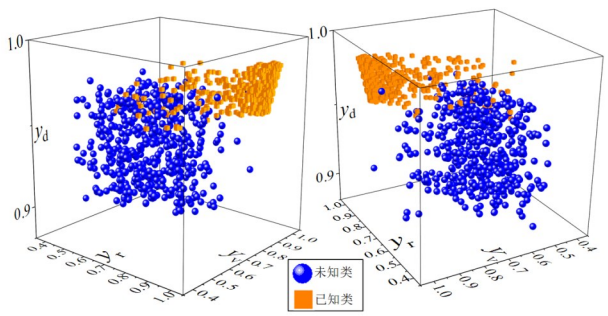


图3 基于三个指标的已知(黄色方框)和未知(蓝色圆点)类别样本分布示意图

为了全面度量偏差情况,将上述三个独立指标相乘,在综合各自影响的基础上,减少了平衡参数调整的难度. 综合指标的计算公式如下:

$$y_{\mathrm{rvd}}=y_{\mathrm{r}}\times y_{\mathrm{v}}\times y_{\mathrm{d}} \tag{9}$$

与其他依赖固定且预先设定的人工阈值的方法相比,DMDI通过基于已知类的综合指标来训练单分类器. 即已知类的y_{rvd}指标趋近于1,而未知类则趋向于0,单分类器学习到这种区分效应,从而增强了在处理新流量样本时的持续有效性. 除此之外,经验证,上述y_{r}和y_{v}指标同样适用于ELM模型.

3.5 基于指标分布的增量训练样本筛选

在选取增量训练的样本时,传统上常采用随机抽样的方法. 然而,这种随机性可能会导致样本代表性不足,进而使得增量学习的效果不稳定,且在开集增量中存在识别错误的样本.

为了解决上述问题,本文设计了一种基于综合指标分布的样本筛选(SS)方法. 该方法根据检测到的未知类样本在不同综合指标区间上的分布进行SS,用于S-ELM模型的增量更新. 具体而言,首先根据每个样本的综合指标值将其划分至对应的指标区间;接着,通过统计各区间内样本数量的分布情况,评估不同指标值范围内样本的丰富程度与代表性. 最后,优先选取分布较集中的样本用于增量更新,从而提高更新样本的质量,更有效地提升模型的分类性能.

记区间数为gn,第i个区间上的样本数为sn_i,进行增量训练的总样本数为$\mathrm{sn}_{\mathrm{total}}$,保留区间的样本数量满足式(10)条件下的区间样本:

$$\mathrm{sn}_i\geqslant\frac{\mathrm{sn}_{\mathrm{total}}}{\mathrm{gn}},\quad 1\leqslant i\leqslant\mathrm{gn} \tag{10}$$

如图4所示,即去除黄色方框中的样本. 这样能够降低增量训练样本中的噪声干扰(即因错误的开集识别带来的标签噪声样本),使得用于增量更新的未知类样本更加可靠,从而提升增量学习的效果并提高模型性能的稳定性.

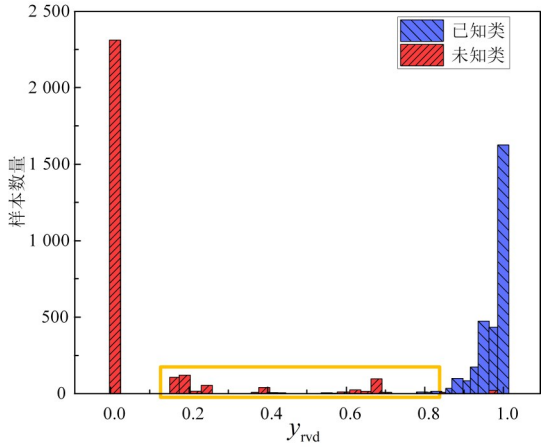


图4 去除样本示意图

在保留的区间中,根据区间样本数量进行随机SS,区间样本数量多的区间筛选的样本数就多,如图5中右边黄框筛选得到的样本数要比左边黄框筛选得到的样本数要略高. 最后共选择$\mathrm{sn}_{\mathrm{total}}$条未知类样本用于增量.

进行增量样本的筛选是为了能优化选取不同分布

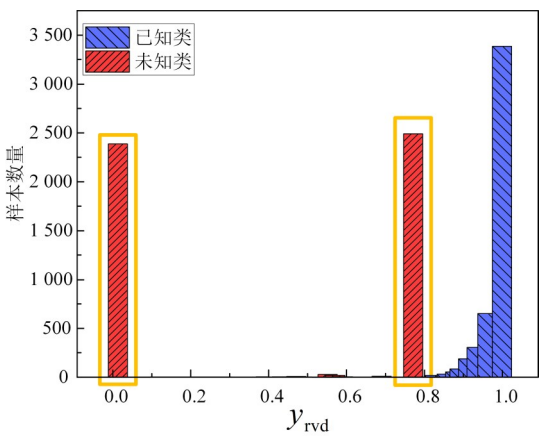


图5 筛选样本示意图

的未知恶意流量样本,同时去除错分到未知类中的少量已知类样本,使用于增量训练的样本更加纯净,分布更加全面.

在增量更新后,新加入的已知恶意流样本经过S-ELM分类器后,计算得到的综合指标仍然保持接近1的特性. 因此,单分类器在增量过程中依然发挥作用. 这也间接证明了设计的统计指标的优越性,不需要依赖复杂的分类器,仅通过简单的单分类器即可识别未知类,并在增量场景下同样适用.

3.6 DMDI算法复杂度分析

DMDI算法的时间复杂度主要来源于S-ELM模型的训练与更新,以及综合指标的计算. 设L为隐藏层的节点数量,N为训练样本数量,D为特征维数,式(1)和式(2)的复杂度分别为$O(N\times D\times L)$与$O\left(N\times L^2\right)$,即S-ELM的训练复杂度为$O\left(N\times D\times L+N\times L^2\right)$. S-ELM增量训练时,设增量的样本大小为$\Delta N$,对应的隐藏层输出为$\Delta \boldsymbol{H}$,其主要依靠矩阵分块逆公式进行快速更新,公式如下:

$$\left(\boldsymbol{A}+\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}\right)^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I}+\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{U})^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1} \tag{11}$$

其中,$\boldsymbol{A}=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$,$\boldsymbol{U}=\Delta\boldsymbol{H}^{\mathrm{T}}$,$\boldsymbol{V}=\Delta\boldsymbol{H}$,则S-ELM增量更新的复杂度为$O\left(\Delta N\times L^2+\Delta N^3\right)$,详细推导见下.

当S-ELM增量训练时,不再从头计算隐藏层输出权值$\boldsymbol{\beta}$. 计算$\boldsymbol{\beta}_{\mathrm{new}}$的公式为

$$\boldsymbol{\beta}_{\mathrm{new}}=\left(\boldsymbol{H}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{new}}\right)^{-1}\boldsymbol{H}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{new}} \tag{12}$$

其中,$\boldsymbol{H}_{\mathrm{new}}=\begin{bmatrix}\boldsymbol{H}\\ \Delta\boldsymbol{H}\end{bmatrix}$,$\boldsymbol{Y}_{\mathrm{new}}=\begin{bmatrix}\boldsymbol{Y}\\ \Delta\boldsymbol{Y}\end{bmatrix}$,$\Delta\boldsymbol{H}$是大小为$\Delta N$的增量样本对应的隐藏层输出.

接下来运用矩阵分块逆公式加速增量更新,令$\boldsymbol{A}=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$,$\boldsymbol{U}=\Delta\boldsymbol{H}^{\mathrm{T}}$,$\boldsymbol{V}=\Delta\boldsymbol{H}$,计算式(12)只需要更新原来的$\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}$即可. 可以得到:

$$\boldsymbol{A}_{\mathrm{new}}=\boldsymbol{H}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{new}}=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\Delta\boldsymbol{H}^{\mathrm{T}}\Delta\boldsymbol{H}=\boldsymbol{A}+\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}} \tag{13}$$

运用矩阵分块逆公式更新逆矩阵:

$$\begin{aligned}\boldsymbol{A}_{\mathrm{new}}^{-1}&=\left(\boldsymbol{A}+\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}}\right)^{-1}\\&=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{I}+\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1}\end{aligned} \tag{14}$$

其中,$\boldsymbol{A}^{-1}\boldsymbol{U}$与$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1}$的时间复杂度都是$O\left(\Delta N\times L^2\right)$,$\left(\boldsymbol{I}+\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}$的时间复杂度则是$O\left(\Delta N^3\right)$. 因此,S-ELM增量更新的复杂度为$O\left(\Delta N\times L^2+\Delta N^3\right)$.

综合指标计算式(6)~式(8)的时间复杂度皆为$O(N)$. 综上所述,DMDI算法的总时间复杂度为$O\left(N\times D\times L+N\times L^2+\Delta N\times L^2+\Delta N^3+N\right)$. 当DMDI处于大规模网络流量环境中,模型可以通过减少隐藏层节点L的数量来提升运行效率.

4 实验结果与分析

4.1 数据集

本实验采用两个广泛用于网络入侵检测的公共数据集CSE-CIC-IDS2018(IDS2018)[30]和NF-UQ-NIDS-v2(NIDS)[31]. 这两个数据集覆盖了多种攻击场景,能够有效地验证模型在复杂真实网络环境中的泛化能力.

IDS2018数据集包含了7种不同的攻击场景,如表1所示,分别由Brute-Force、Heartbleed、Botnet、DoS、DDoS、Web Attacks和Infiltration组成,其中有3种攻击流量类别的占比低于1%. 显然,多数类与少数类的差距极大,一些攻击流量样本数量稀少,导致分类器的训练样本不足,这也是识别攻击流量的挑战之一.

表1　数据集IDS2018:对应标签、流量类型及样本数

流量类型	样本数
Bot, Brute Force-Web, Brute Force-XSS	20 000, 611, 230
DDOS-HOIC, DDoS-LOIC-UDP, DDoS-LOIC-HTTP	20 000, 1 730, 20 000
DoS-GoldenEye, DoS-Hulk, DoS-SlowHTTPTest	20 000
DoS-Slowloris, FTP-BruteForce, Infiltration28	10 990, 20 000, 20 000
SQL Injection, SSH-Bruteforce, Infiltration1	87, 20 000, 20 000

对于特征提取(Feature Extraction,FE)部分,IDS2018使用CICFlowMeter-V3[32]工具从捕获的流量中提取如持续时间、数据包数量、字节数等80个特征. 然而,这些原始特征集中于流量基础统计量,缺乏对时序行为模式的刻画,为此,本研究在实验阶段额外添加了如最小数据包时间间隔与最大数据包时间间隔比、流持续时间与正反向数据包总时间间隔比等40个特征,将特征维度扩展至120维,以增强模型对流量动态行为模式的捕捉能力.

NIDS数据集由UNSW-NB15[33]、BoT-IoT[34]、ToN-IoT[35]和CSE-CIC-IDS2018数据集整合而成. 相较于IDS2018,该数据集包含更丰富的攻击流量类型,如表2所示,包含了DDoS、DoS、Scanning、Reconnaissance、XSS、Injection、Bot等20种攻击,模拟了更加真实、复杂的网络环境,为模型的检测性能带来了更大的挑战.

NIDS数据集的FE方法是由Sarhan等人[31]提出的基于NetFlow特征的新变体,原始特征包含如流持续时间、数据包数量等43个特征. 为了避免过拟合特定网络配置,本实验中移除了如IP地址与端口号等环境相关特征,同时添加了如平均每毫秒(ms)字节数、平均每毫秒包数目等22个额外特征,使最终的特征数达到61维.

表2 数据集NIDS:对应标签、流量类型及样本数

流量类型	样本数
DDoS, DoS, Reconnaissance, Scanning	6 000
XSS, Analysis, Backdoor, Bot	6 000
Brute Force, Exploits, Fuzzers, Generic	6 000
Infilteration, Injection, Mitm, Password	6 000
Ransomware, Shellcode, Theft, Worms	3 288, 1 369, 2 324, 158

4.2 评价指标

鉴于本文方法涉及未知攻击流量检测与类增量学习两个核心任务,模型评估将从这两个维度展开. 针对开集识别,不仅需关注未知类的检测能力,还需全面评估已知类的细分类性能与整体分类效率. 具体评估指标如下文所述.

(1)开集总体准确率(NA)

NA表示所有分类正确的样本占全部样本的比例,代表了未知类检测的能力,计算公式如下:

$$\mathrm{NA}=\lambda\mathrm{AKS}+(1-\lambda)\mathrm{AUS} \tag{15}$$

其中,λ为开集样本中已知类所占的比例,$0<\lambda<1$,本文实验中λ设置为0.5;AKS表示已知类的分类准确率,即Known-Acc;AUS表示未知类的分类准确率,即Unknown-Acc,计算公式如下:

$$\mathrm{AKS}=\mathrm{Acc}=\frac{T_P+T_N}{T_P+F_P+T_N+F_N} \tag{16a}$$

$$\mathrm{AUS}=\frac{T_U}{T_U+F_U} \tag{16b}$$

其中,T_P和F_P分别表示某类被正确分类和错误分类的样本数;T_N和F_N分别表示其他类被正确分类和被误分类的样本数;T_U和F_U表示未知类被正确和错误分类的样本数. 由式(12)可知,NA能同时反映模型对未知类检测能力与已知类分类能力.

(2)F1得分(F1)

$$\mathrm{F1}=\frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \tag{17}$$

其中,$\mathrm{Precision}=\frac{T_P}{T_P+F_P}$,$\mathrm{Recall}=\frac{T_P}{T_P+F_N}$. 如前所述,网络攻击流量往往存在数据分布不均衡问题. F1得分作为多类别分类任务中衡量模型性能的平衡性指标,能够更全面地反映模型在所有已知类别上的整体表现,尤其适用于已知恶意流的精细分类评估. 上述指标值越大,说明模型在开集识别中的检测精度与分类性能越优越.

在开集类增量学习场景中,随着类增量任务的持续进行,模型易受到灾难性遗忘与未知类检测误差的影响,导致性能逐步下降. 因此,引入了多维度指标综合评估模型的增量学习能力,包括准确率差值(Acc差)、训练时间和增量更新时间等指标.

(3)准确率差值(Acc差)

$$\mathrm{Acc差}=\mathrm{Acc}_t-\mathrm{Acc}_0 \tag{18}$$

其中,Acc_t为第t次增量训练后的准确率,Acc_0为模型初始准确率. Acc差越小,说明模型在增量学习中的性能保持越稳定,对已学知识遗忘程度越低,对新知识的增量适应能力越强.

4.3 实验环境

分类器模型采用S-ELM和OC-SVM,其中S-ELM中隐含神经元的个数设为100,激活函数为Sigmoid函数;OC-SVM中规则化参数设为0.07,所有模型均采用5折交叉验证进行评估. 实验平台为Linux操作系统,CPU为酷睿 i5-13400f@2.5 GHz内核,内存为16 GB,GPU为RTX3060 Ti.

4.4 不同方法的性能对比

4.4.1 不同未知类检测方法的性能对比

为了验证DMDI的有效性,在对比实验阶段选取了DACS[15]、PAT[21]、RFG[13]和AutoIoT[10]四个文献方法进行性能比较. 这四种方法均涉及未知攻击检测与开集增量学习场景. 其中,DACS是一种基于传统机器学习算法的模型,具有较高的分类实时性,为评估DMDI在保持检测性能的同时能否具备更优的实时处理能力提供了参考;PAT与RFG分别基于原型学习和生成对抗网络构建,通过辅助样本生成机制优化决策边界,为当前提升未知类识别精度的主流方法;AutoIoT和PAT则在模型增量学习时引入迁移学习框架和原型学习,旨在缓解灾难性遗忘现象,由此对比可以评估DMDI提出的基于S-ELM增量机制的表现. 需要特别说明的是,AutoIoT原本采用多任务学习框架区分物联网设备,为实验公平性,本文在检测未知恶意网络流量时省略了多任务学习模块,并统一采用与RFG相同的DNN网络结构进行训练,以保证性能比较的严谨性与公正性.

经过在上述数据集中二十余种不同类别组合下的实验,表3统计了本文方法与其他四种方法的平均分类时间、NA和已知类Known-F1指标.

为了更直观地比较不同方法在IDS2018数据集所有实验组合中的NA表现,图6展示了各个方法在NA上的数据分布. 图6中的每个子图通过核密度估计来反映了不同NA的相对密集程度,图形的宽度表示该区域数据的密集度,其中较宽的区域表示该值域的数据分布更集中,而较窄的部分表明数据较为稀疏. 每个子图中的上下两条黑线代表最大值和最小值,中间的黑线表示均值,虚线表示中位数. 由图6可知,尽管所有方法均遭遇了NA表现不佳的情况,但DMDI在较低NA区域的分布相对较少,且其最低NA依然保持在0.5以上.

对于Known-F1指标,更能体现分类模型对已知类的分类性能. 从图7中可以看出,本文方法的Known-F1

表3　不同方法平均分类性能对比

性能指标(数据集)	DMDI	DACS	RFG	AutoIoT	PAT
分类时间/(μs/样本)	59.5	76.3	2 355.0	89.8	105.3
NA(IDS2018)/%	88.1	76.2	85.5	82.0	75.6
Known-F1(IDS2018)/%	92.3	89.1	90.3	88.1	83.2
NA(NIDS)/%	90.1	86.1	87.6	86.1	80.3
Known-F1(NIDS)/%	83.9	78.9	83.5	81.8	76.4

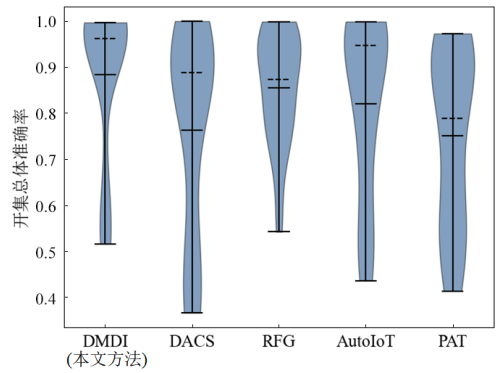


图6　IDS2018中NA分布图

主要集中在0.95左右,而其他对比方法在低于0.9的区域也有所集中. 基于深度学习的方法则往往对训练样本的数量有较高要求,如PAT在样本不足时难以生成具有足够代表性的原型,从而影响分类性能. RFG在训练中加入了对比学习的方法,提高了整体的准确性. 相比之下,尽管DMDI并非在所有实验组合中都表现最佳,但其表现更为均衡,总体性能较优. 这是因为所使用的单分类器OC-SVM具有较强的普适性,表现更为稳定. 此外,OC-SVM不需要额外的辅助样本或复杂的经验来设置阈值参数,使得模型更加自动化.

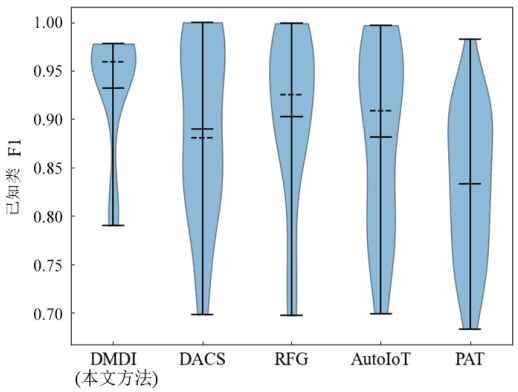


图7　IDS2018中Known-F1分布图

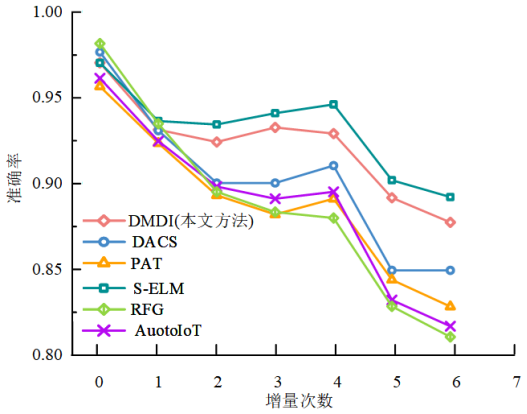
在分类时间方面,DMDI继承了ELM分类速度快的优势,同时其计算指标的复杂度较低,因此整体分类速度较快. 对于AutoIoT和PAT,由于深度学习方法在测试阶段仍需要执行多层的前向传播计算,分类时间上不具备优势,尤其是PAT还需要额外构造虚拟原型,导致其深度残差网络对每个样本的平均分类时间超过了100 μs. RFG在检测未知类时需生成未知类分布并与已知类分布集成,过程复杂,分类时间达到ms级别,且模型更新时需要进行费时的重训练. DACS则凭借传统机器学习模型实现较快的分类速度,但由于需将已知类划分为两部分,当已知类数量少于4个时无法进行未知类检测. 相比之下,DMDI不仅避免了上述局限,还具备快速增量更新能力. 综合而言,DMDI在分类时间和适用性上优于四种对比方法.

4.4.2　不同增量更新方法的性能对比

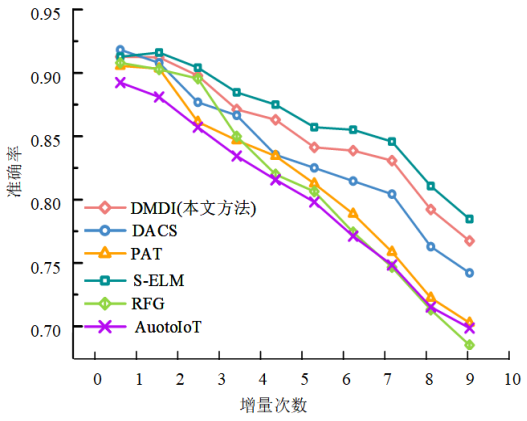
S-ELM模型本身不具备未知类检测的能力,在文献[17]中,仅是在闭集中使用未知类样本对模型进行连续增量训练. 因此,本节将闭集环境下的S-ELM设立为对照组,在该组实验中,每次增量更新时加入的样本均不含噪声,以此作为理想情况下模型精度的参考上限. 另外四个对比方法则是开集环境下的DACS、RFG、AutoIoT和PAT.

在增量学习对比实验中,初始模型训练和每个增量的未知类均使用每个类别500个样本. 最终得到图8(*a*)和图8(*b*)所示增量不同次数后Acc对比. 由图8可知,DMDI增量效果仅次于闭集S-ELM,相比之下,DACS和RFG虽然采用了重训练策略,但因增量训练样本中包含噪声,模型表现受到一定影响. AutoIoT通过迁移学习更新模型,其精度略优于重训练方法,但在源数据与目标数据相似性较低时,灾难性遗忘问题更为严重[36]. PAT则通过上一阶段保留的原型和部分样本防止灾难性遗忘,但未能有效应对增量训练过程中噪声逐步增加的问题,导致生成的原型逐渐失准,影响了模型性能. 而DMDI的SS方法可以将检测为未知类的样本过滤的更纯净,结合S-ELM模型使得增量效果更接近重训练效果,使得DMDI的增量更新效果在开集环境中优于其余对照方法,在闭集环境中接近闭集S-ELM的增量效果.

表4记录了在两个数据集上,不同实验组合的增量前后Acc差、训练时间以及增量更新时间. 训练时间表示的是模型单轮训练2 500个样本的时长,增量更新时间则是各个增量学习方法新检测到500个未知类样本后的单次模型更新时间.



(*a*) IDS2018数据集



(*b*) NIDS数据集

图8 不同增量次数时增量后Acc对比

如表4所示,在开集环境下,使用IDS2018数据集时,DMDI在6次增量后模型性能降低的最少,优于DACS约3.5%;使用NIDS数据集时,同样DMDI在9次增量后模型性能降低的最少,可见DMDI在开集环境下的增量效果优于4种文献对比方法. 闭集S-ELM在增量训练中采用了随机选取样本的策略,未能优先选择对增量学习更有益的样本,因此影响了其增量学习效果. 相比之下,DMDI通过SS方法在开集环境下的增量效果仅降低了约1.5%,并且在学习完所有类别后,DMDI的平均Acc也仅比S-ELM低约2%,进一步反映出筛选纯净样本的重要性.

在时间性能方面,由于闭集S-ELM不涉及未知类检测,不参与时间对比. DACS和RFG都采用了模型重训练的策略,因此每次增量更新时间接近于训练时间. AutoIoT通过迁移学习,仅需加载增量样本即可完成更新,显著缩短了增量更新时间. 然而,由于深度神经网络具有较大的参数规模,其更新速度仍不及基于RF和SVM的DACS. 此外,RFG、AutoIoT和PAT的整体训练耗时可能远超表中所列时间,因为表4中仅统计了模型单轮训练的时间,而深度学习模型通常需要多轮迭代才能收敛,这进一步限制了其在资源受限环境中的应用. 相比之下,提出的DMDI方法基于ELM模型的轻量级S-ELM网络,训练与更新时间远低于其余对比方法,具备更高的效率.

4.4.3 实时性分析

在实际网络入侵检测系统的部署中,边缘设备(如网关、路由器)往往作为第一道防线,负责实时监测网络通信中的异常行为. 然而,这类设备通常存储空间有限,难以承载参数规模庞大的深度学习模型. 同时,由于缺乏高性能计算硬件如GPU的支持,进一步限制了模型运行所需的计算能力.

针对这一挑战,本文引入的S-ELM模型作为一种浅层神经网络,具备参数量小、训练方式高效的优势,对设备存储空间和计算资源的需求低,推理过程仅需普通CPU即可完成,适合在大多数边缘设备上轻量化部署. 此外,考虑到边缘路由器处理网络流量的时间性能要求,如表3所示,DMDI能够在约60 μs内处理一条网络流,而一条数据流中通常包含数十到上百个数据包. 这意味着DMDI可以实现0.1~1 Mpps(百万包每秒)的吞吐量,与主流边缘路由器的性能指标大体契合[37],从而基本可以实现检测的实时性. 相比之下,基于深度神经网络的方法由于推理时间过长,导致包吞吐量显著下降,难以满足边缘设备对低延迟、高吞吐的需求,因此通常不适合在边缘场景中部署[38].

4.5 消融实验

4.5.1 不同指标分布区间数对比

根据SS的定义可知,SS效果受到区间数gn的影响. 理论上,区间数的减少会导致样本分布较为集中,进而使得筛选出的样本分布不好,集中在某几个区间;相反,区间数的增加会使得样本分布较为均衡,从而使筛选出的样本分布更加全面,涵盖较多的区间. 为了进一步验证该理论,并寻找最合适的区间数,本小节对区间数为10、20、40、50、80和100的情况进行了实验,图9

表4 不同实验组合平均性能对比

性能(数据集)	DMDI	DACS	RFG	AutoIoT	PAT	S-ELM
Acc差(0→6)(IDS2018)/%	−9.1	−12.5	−16.8	−12.6	−13.5	**−7.7**
Acc差(0→9)(NIDS)/%	−14.0	−17.0	−21.5	−18.7	−19.6	**−12.4**
单轮训练时间/ms	**245**	389	720	694	938	—
单次增量更新时间/ms	**95**	306	689	365	445	—

注:加粗数据为最优结果.

比较了每个数据集的实验组合增量后模型的 Acc 以及两个数据集的平均(Ave)结果.

由图 9 可以清晰地看出,随着区间数 gn 的增加,增量后模型的 Acc 明显上升,特别是在 gn=50 时,到达拐点,模型的增量达到最佳效果;当 gn>50 时,模型的性能趋于平稳,甚至略有下降. 因此最终选择区间数 gn 为 50.

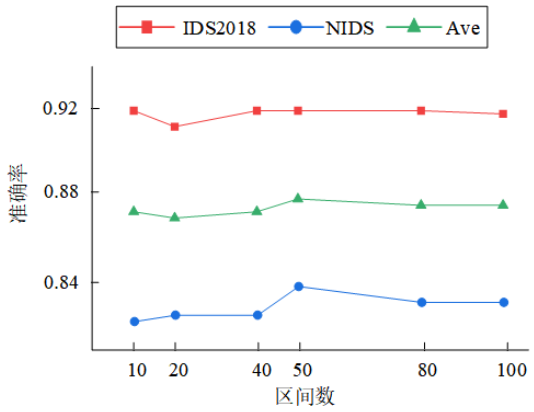


图 9　不同区间数对增量样本筛选的影响

4.5.2　不同增量更新样本对比

为比较使用不同的样本对 S-ELM 增量更新的效果的影响,分别对使用不同样本进行增量更新进行对比. 其中,RS 表示使用随机筛选检测到的未知类进行增量更新;RS-A 表示使用随机筛选检测到的未知类和识别到的已知类进行增量更新;SS-A 表示使用 SS 检测到的未知类进行筛选. 对识别到的已知类进行随机筛选,然后使用筛选到的样本进行增量更新. 多个实验组合下的平均结果如图 10 所示.

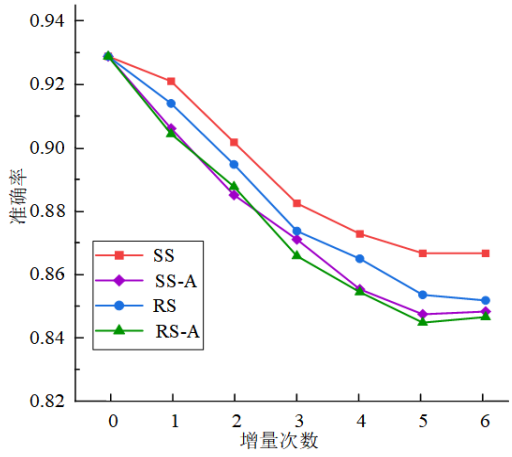


图 10　使用不同样本增量更新下 Acc 对比

可以看出,随着增量次数的增加,使用 SS 方法增量更新后模型的 Acc 最好. 这是因为已知类的分类效果并不理想,存在一定的错分现象,若在增量更新过程中使用含有随机筛选的已知类样本,会导致模型在更新后的分类性能出现下降,甚至低于不使用已知类进行更新的情况. 同时,SS 方法选择的未知类样本中包含更少错分的已知类样本,并且样本分布的更加全面. 因此,使用 SS 方法得到的样本对 S-ELM 进行增量更新是一种更为有效的方法,能够在不引入额外噪声的情况下,提升模型的分类性能.

5　结论

为提高含有恶意流量的开集流分类性能,本文提出了一种可增量更新的 OSR 模型. 通过对 S-ELM 输出权重与标准权重进行的对比分析,设计了一种基于 S-ELM 输出权重的统计分布的未知类检测方法,计算 S-ELM 输出权重与标准权重的改进的最接近皮尔森相关系数、归一化相对方差和归一化 Tos 列距离三个指标,通过相乘最终得到综合指标,再通过 OC-SVM 进行未知类检测;通过对综合指标在不同区间上的样本分布进行的分析,设计了基于指标分布的增量训练样本筛选方法;根据未知类样本的综合指标分布情况进行筛选,使用筛选后的样本对 S-ELM 进行增量更新,优化 S-ELM 在实际 OSR 任务中的增量更新能力. 在两个公共数据集上进行了实验验证,对比文献方法,在未知类检测方面 NA 指标能改善 3%~13%,持续增量更新后模型的 Acc 性能可以提高约 3%~7%.

下一步的工作是提高模型整体的增量更新能力,探索综合指标是否适用于其他分类器模型,以及缩短指标计算时间.

参考文献

[1] Statista. Number of apps available in leading app stores as of August 2024[EB/OL]. (2024-08-30) [2025-01-01]. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores.

[2] CLOUDFLARE. 2024 Application security trends report[R/OL]. (2024-07-12)[2025-01-01]. https://www.cloudflare.com/zh-cn/2024-application-security-trends.

[3] BORKAR A, DONODE A, KUMARI A. A survey on intrusion detection system (IDS) and internal intrusion detection and protection system (IIDPS)[C]//2017 International Conference on Inventive Computing and Informatics (ICICI). Piscataway: IEEE, 2017: 949-953.

[4] OBASI T, SHAFIQ M O. An experimental study of different machine and deep learning techniques for classification of encrypted network traffic[C]//2020 IEEE International Conference on Big Data (Big Data). Piscataway: IEEE, 2020: 4690-4699.

[5] 周奕涛, 张斌, 刘自豪. 基于多模态深度神经网络的应用层 DDoS 攻击检测模型[J]. 电子学报, 2022, 50(2):

508-512.
ZHOU Y T, ZHANG B, LIU Z H. Application layer DDoS detection model based on multimodal deep learning neural network[J]. Acta Electronica Sinica, 2022, 50(2): 508-512. (in Chinese)

[6] 胡向东, 张琴. 基于特征组合优化的工业互联网恶意行为实时检测方法[J]. 电子学报, 2024, 52(9): 3075-3085.
HU X D, ZHANG Q. Real-time detection method of malicious behaviors in industrial Internet based on feature combination optimization[J]. Acta Electronica Sinica, 2024, 52(9): 3075-3085. (in Chinese)

[7] LE S Q, LAI Y X, WANG Y P, et al. An adaptive classification and updating method for unknown network traffic in open environments[J]. Computer Networks, 2024, 238: 110114.

[8] BENDALE A, BOULT T. Towards open world recognition[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 1893-1902.

[9] ZHOU P, WANG N, ZHAO S, et al. Difficult novel class detection in semisupervised streaming data[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(10): 6872-6886.

[10] FAN L N, HE L, WU Y C, et al. AutoIoT: Automatically updated IoT device identification with semi-supervised learning[J]. IEEE Transactions on Mobile Computing, 2023, 22(10): 5769-5786.

[11] PING G L, YE X J. Open-set intrusion detection with MinMax autoencoder and pseudo extreme value machine[C]//2022 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE, 2022: 1-8.

[12] YANG J, CHEN X, CHEN S W, et al. Conditional variational auto-encoder and extreme value theory aided two-stage learning approach for intelligent fine-grained known/unknown intrusion detection[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 3538-3553.

[13] ZHONG Y, WANG Z L, SHI X G, et al. RFG-HELAD: A robust fine-grained network traffic anomaly detection model based on heterogeneous ensemble learning[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 5895-5910.

[14] WANG T T, LV Q J, HU B, et al. A few-shot class-incremental learning approach for intrusion detection[C]//2021 International Conference on Computer Communications and Networks (ICCCN). Piscataway: IEEE, 2021: 1-8.

[15] LIANG Y L, WANG F, CHEN S H. DACS: A double-layer application classification scheme for hybrid zero-day traffic[C]//2022 IEEE 22nd International Conference on Communication Technology (ICCT). Piscataway: IEEE, 2022: 1380-1387.

[16] ZHANG J T, ZHANG J, GHOSH S, et al. Class-incremental learning via deep model consolidation[C]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2020: 1120-1129.

[17] LEE C L, CHEN Y T, WU A Y. A scalable extreme learning machine (S-ELM) for class-incremental ECG-based user identification[C]//2021 IEEE International Symposium on Circuits and Systems (ISCAS). Piscataway: IEEE, 2021: 1-5.

[18] SCHEIRER W J, DE REZENDE ROCHA A, SAPKOTA A, et al. Toward open set recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(7): 1757-1772.

[19] ZHANG J L, LI F H, YE F, et al. Autonomous unknown-application filtering and labeling for DL-based traffic classifier update[C]//IEEE INFOCOM 2020 - IEEE Conference on Computer Communications. Piscataway: IEEE, 2020: 397-405.

[20] YANG L, GUO W, HAO Q, et al. CADE: Detecting and explaining concept drift samples for security applications[C]//30th USENIX Security Symposium (USENIX Security 21). California: USENIX Association, 2021: 2327-2344.

[21] ZHAO Z X, ZHANG H Y, MIN H, et al. Towards recognition of open-set speech forgery algorithms by using prototype learning[C]//Third International Conference on Algorithms, Microchips, and Network Applications (AMNA 2024). Xian: SPIE, 2024: 1317102.

[22] 郭虎升, 丛璐, 高淑花, 等. 基于在线集成的概念漂移自适应分类方法[J]. 计算机研究与发展, 2023, 60(7): 1592-1602.
GUO H S, CONG L, GAO S H, et al. Adaptive classification method for concept drift based on online ensemble[J]. Journal of Computer Research and Development, 2023, 60(7): 1592-1602. (in Chinese)

[23] 韩光洁, 赵腾飞, 刘立, 等. 基于多元区域集划分的工业数据流概念漂移检测[J]. 电子学报, 2023, 51(7): 1906-1916.
HAN G J, ZHAO T F, LIU L, et al. Concept drift detection of industrial data flow based on multivariate region set partition[J]. Acta Electronica Sinica, 2023, 51(7): 1906-1916. (in Chinese)

[24] SHENG C, YAO Y, LI W X, et al. Unknown attack traffic classification in SCADA network using heuristic clustering technique[J]. IEEE Transactions on Network and

Service Management, 2023, 20(3): 2625-2638.
[25] ZHANG L, CUSHING R, DE LAAT C, et al. A real-time intrusion detection system based on OC-SVM for containerized applications[C]//2021 IEEE 24th International Conference on Computational Science and Engineering (CSE). Piscataway: IEEE, 2021: 138-145.
[26] SAPUTRA RANGKUTI F R, ALI FAUZI M, SARI Y A, et al. Sentiment analysis on movie reviews using ensemble features and Pearson correlation based feature selection[C]//2018 International Conference on Sustainable Information Engineering and Technology (SIET). Piscataway: IEEE, 2018: 88-91.
[27] PERNA G, MARKUDOVA D, TREVISAN M, et al. Online classification of RTC traffic[C]//2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). Piscataway: IEEE, 2021: 1-6.
[28] YU J, XIA C M, XIE J Z, et al. Research on feature importance of gait mechanomyography signal based on random forest[C]//2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL). Piscataway: IEEE, 2020: 191-196.
[29] HSU Y C, SMITH J, SHEN Y L, et al. A closer look at knowledge distillation with features, logits, and gradients[EB/OL]. (2022-05-18)[2025-01-01]. https://arxiv.org/abs/2203.10163v1.
[30] SHARAFALDIN I, HABIBI LASHKARI A, GHORBANI A A. Toward generating a new intrusion detection dataset and intrusion traffic characterization[C]//Proceedings of the 4th International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and Technology Publications. Oxford: ICISS, 2018: 108-116.
[31] SARHAN M, LAYEGHY S, PORTMANN M. Towards a standard feature set for network intrusion detection system datasets[J]. Mobile Networks and Applications, 2022, 27(1): 357-370.
[32] HABIBI LASHKARI A, DRAPER GIL G, MAMUN M S I, et al. Characterization of tor traffic using time based features[C]//Proceedings of the 3rd International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and Technology Publications. Oxford: ICISS, 2017: 253-262.
[33] MOUSTAFA N, SLAY J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)[C]//2015 Military Communications and Information Systems Conference (MilCIS). Piscataway: IEEE, 2015: 1-6.
[34] KORONIOTIS N, MOUSTAFA N, SCHILIRO F, et al. A holistic review of cybersecurity and reliability perspectives in smart airports[J]. IEEE Access, 2020, 8: 209802-209834.
[35] MOUSTAFA N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets[J]. Sustainable Cities and Society, 2021, 72: 102994.
[36] OTOVIĆ E, NJIRJAK M, JOZINOVIĆ D, et al. Intra-domain and cross-domain transfer learning for time series data: How transferable are the features?[J]. Knowledge-Based Systems, 2022, 239: 107976.
[37] GALLO M, FINAMORE A, SIMON G, et al. Fenxi: Deeplearning traffic analytics at the edge[C]//in 2021 IEEE/ACM Symposium on Edge Computing (SEC). Piscataway: IEEE, 2021: 202-213.
[38] CHEN Z H, CHENG G, WEI Z J, et al. Classify traffic rather than flow: Versatile multi-flow encrypted traffic classification with flow clustering[J]. IEEE Transactions on Network and Service Management, 2024, 21(2): 1446-1466.

作者简介

陆浩天　男,1998年9月出生于江苏省南通市. 现为南京邮电大学通信与信息工程学院博士研究生. 主要研究方向为多媒体通信、网络流识别和联邦学习.
E-mail: lhtnjupt@163.com

董育宁　男,1955年出生于江苏省南京市. 现为南京邮电大学通信与信息工程学院教授、博士生导师. 主要研究方向为无线网络、多媒体通信和网络流识别.
E-mail: 19900011@njupt.edu.cn

全宇轩　男,1998年6月出生于山东省枣庄市. 2024年毕业于南京邮电大学,获工学硕士学位. 现为天翼安全科技有限公司研发工程师.
E-mail: quanyx1233@163.com