研究与开发

基于Powershap和混合采样的动态集成入侵检测模型

黄冬梅[1]，颜昊[2]，张文博[3]，胡安铎[2]，孙锦中[2]，孙园[4]
（1. 上海电力大学电气工程学院，上海 200090；
2. 上海电力大学电子与信息工程学院，上海 201306；
3. 上海海洋大学信息学院，上海 201306；
4. 上海电力大学数理学院，上海 201306）

摘　要：随着互联网技术的迅猛发展，网络安全领域中的入侵检测任务变得更加重要。针对目前入侵检测中存在的特征维度高、数据类别不平衡以及单一分类器检测率低的问题，提出了一种基于Powershap和混合采样的动态集成入侵检测模型。首先，通过Powershap算法对数据集进行特征选择。随后，采用RENN-BorderlineSMOTE混合采样算法，对特定类别数据分别进行欠采样和过采样处理，解决数据集中的类别不平衡问题。最后，基于广义多样性从多个基分类器中筛选出最优组合，并将其集成至动态集成框架KNORAE中以结合多个基分类器的优势。模型在CIC-IDS2017数据集上进行了验证，证实了该模型在网络流量检测中的优越性。

关键词：入侵检测；类别不平衡；集成学习；综合采样
中图分类号：TP393
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2025178

Dynamic integrated intrusion detection model based on powershap and hybrid sampling

HUANG Dongmei[1], YAN Hao[2], ZHANG Wenbo[3], HU Anduo[2], SUN Jinzhong[2], SUN Yuan[4]
1. College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China
2. College of Electronic and Information Engineering, Shanghai University of Electric Power, Shanghai 201306, China
3. School of Information, Shanghai Ocean University, Shanghai 201306, China
4. College of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 201306, China

Abstract: With the rapid development of Internet technology, the task of intrusion detection of the field of network security has become more important. Aiming at the problems of high feature dimension, imbalance of data categories and low model detection rate of single classifiers in current intrusion detection, a dynamic integrated intrusion detection model based on Powershap and hybrid sampling was proposed. Firstly, the Powershap algorithm was used for fea-

收稿日期：2025−03−06；修回日期：2025−05−18
通信作者：孙园，combmathe@shiep.edu.cn
基金项目：国家自然科学基金青年科学基金资助项目（No.62102243）
Foundation Item: Young Scientists Fund of the National Natural Science Foundation of China (No.62102243)

ture selection of the dataset. Subsequently, the hybrid RENN-BorderlineSMOTE sampling algorithm was applied to address the category imbalance in the dataset by under-sampling and over-sampling specific categories of data. Finally, the optimal combination was filtered from multiple base classifiers based on Generalization Diversity and integrated into the dynamic integration framework KNORAE to combine the advantages of multiple base classifiers. The model was validated on the CIC-IDS2017 dataset, which confirmed the superiority of the model in network traffic detection.

Key words: intrusion detection, class imbalanced, ensemble learning, mixed sampling

0　引言

随着时代的不断发展和进步，互联网逐渐和人们的生活变得息息相关，但是网络中入侵行为的频率不断提高、规模不断扩大。这些入侵不仅危害人们的日常生活，还对国家安全和社会稳定有着很大的威胁。因此，发现网络流量中的异常数据对于网络空间的安全与稳定有着重要意义。

入侵检测的概念由 James Anderson 在 1980 年首次提出[1]。根据检测入侵的方法可以将入侵检测系统区分为基于误用的入侵检测系统和基于异常的入侵检测系统。基于误用的入侵检测系统基于已有知识库进行检测，对不在知识库中的异常行为检测效果较差。基于异常的入侵检测系统能通过对比正常行为和异常行为来检测新的和独特的攻击，故而基于异常的入侵检测系统更具优势。网络中的异常流量数量庞大并且种类多样，使得异常流量检测在此背景下不断地被关注，同时也在不断的发展，本文研究基于异常的入侵检测系统。

近年来，机器学习快速发展并且可以快速进行流量类型检测，故而近年来越来越多的机器学习算法应用于入侵检测领域。例如，何红艳等[2]结合了极限树特征递归消除和 LightGBM，以应对数据维度高、样本不平衡和数据集的高分散性问题，该方法特别针对样本量较小的攻击类型，能有效提升其检测效率；陈俊彦等[3]利用 5 个不同的特征选择算法进行集成，并筛选出最优的特征子集，最后加权集成多个异质基分类器进行分类，使模型的泛化能力及精确率均得到提升；蹇诗婕等[4]提出了一种基于稀疏异常样本数据场景中的新型深度神经网络入侵检测方法，并在 CIC-IDS2017 数据集和 UNSW-NB15 数据集上进行了验证；魏明军等[5]提出了一种用来筛选车辆网络数据中可能的入侵攻击的 RGSNet 模型，并在 CIC-IDS2017 数据集上进行了验证；孙敬等[6]利用堆叠降噪稀疏自编码器进行特征降维，并利用卷积注意力机制对残差网络进行改进以实现对网络流量的分类预测，在 UNSW-NB15 数据集和 CICIDS 2017 数据集上进行了验证。徐会斌等[7]利用自适应合成采样处理类别不平衡的问题，然后利用 GBDT 评估特征的重要性以筛选出更重要的特征，最后利用 k 折交叉验证的 stacking 方法进行分类预测，在 CICIDS 2017 数据集和 NSL-KDD 数据集上证实了模型的优越性。

上述研究表明了现有的入侵检测机器学习模型取得了一定的效果，但仍存在一些问题。首先是数据的特征维度高，其中有许多冗余特征或者不相干的特征，使得模型的计算成本骤增，同时也影响模型的检测性能。其次是现有真实世界的网络流量数据往往存在着数据类别高度不平衡的问题，异常流量的数量要远远少于正常流量的数量，而这会导致模型在检测过程中更加倾向多数类，致使在实际检测中对少数类网络流量的检测效果较差，但是对于入侵检测系统来说，对少数类的检测更为重要一些。最后是仅使用单一分类器或者静态集成策略进行预测，无法更有效结合多个分类器的优势以更好地提升分类器的性能。

针对以上几个问题，本文提出了一个入侵检测模型，首先利用Powershap算法[8]去除冗余或不相关的特征，以实现对数据的维度的降低，避免了维度灾难，从而使得模型运行更加高效、检测性能更好。然后采用重复编辑最近邻（repeated edited nearest neighbours，RENN）[9]和边界合成少数类过采样技术（border line synthetic minority oversampling technique，BorderlineSMOTE）[10]进行混合采样，从而有效地消除数据集的类别高度不平衡问题。最后，将筛选出的基分类器集成到基于K近邻预测和Oracle的方法（k-nearest-oracles-eliminate，KNORAE）[11]中来区分流量类型，动态集成模型相较于传统静态集成模型能更好地结合多个模型的优势来提升集成模型的性能以及降低过拟合风险。在CIC-IDS2017数据集上进行了实验以验证所提模型的性能，实验表明，本文所提模型有较好的检测性能。

1 基于Powershap和混合采样的动态集成入侵检测模型

1.1 Powershap特征选择算法

Powershap算法是一种新颖的包装器特征选择方法，它结合了统计学原理和Shapley值来快速识别和选择对模型预测有显著影响的特征。该算法基于的核心假设是，与随机生成的特征相比，含有有效信息的特征对预测结果的影响更大。Powershap算法由2个主要组件构成：解释组件和核心组件。

在解释组件中，算法通过不同的随机种子在多个数据子集上训练多个模型，每个子集包含一个额外的随机特征以及所有原始特征，这个随机特征将作为一个基准用来与实际特征进行比较。这一步骤的目的是利用Shapley值来解释样本外数据集中每个特征的平均影响，从而评估真正的无偏影响。每个特征的Shapley值的绝对值被取平均，得到该特征的总平均影响。这一过程在不同的迭代中重复，每次使用不同的随机特征和数据子集。

在核心组件中，给定每个特征对每次迭代的平均影响，就可以将其与核心Powershap组件中随机特征的影响进行比较，通过这种方法，Powershap算法能够有效地从大量特征中筛选出对模型预测有实质性影响的关键特征，从而提高模型的解释性和预测性能。这种比较通过式（1）所示的百分位数计算式进行量化，其中考虑了特征的平均Shapley值数组与随机特征影响的比较。

$$P(s,x)=\sum_{i=1}^{n}\frac{M(x>s_i)}{n} \tag{1}$$

其中，s表示某个特征在多次迭代中的平均Shapley值数组，i表示迭代次数，x表示单个值，M表示指示函数。此计算式计算x高于该迭代的平均Shap值的迭代分数。

Powershap算法需要设定2个超参数，分别是P值阈值α和迭代次数I。其中I控制了算法的迭代深度，而α用于确定特征影响的显著性水平。Powershap还存在自动模式，自动模式能自动确定迭代次数I，从而减轻了用户手动选择、调整超参数的压力。

统计功效表示为$1-\beta$，用来评估避免假阴性（即错误地将信息特征标记为非信息特征）的能力，其中β是假阴性（false negative，FN）的概率。由式（2）计算的测试样本的统计检验输出P值α，表示特征被偶然标记为有显著影响的特征概率。如果统计测试中的数据很少，则可能会有较低的α和较大的β同时出现的情况，导致输出结果不可信。统计检验的功效利用式（3）的基础检验分布H_1的累积分布函数F来计算，其中H_0表示随机特征影响分布，H_1表示经过测试的特征影响分布。

$$\alpha(x)=F_{H_0}(x) \tag{2}$$

$$\text{power}(\alpha)=F_{H_1}(F_{H_0}^{-1}(\alpha)) \tag{3}$$

1.2　RENN-BorderlineSMOTE 混合采样算法

解决数据不平衡的方法主要有减少多数类样本数量，即对多数类数据进行降采样和增加少数类样本数量，即对少数类数据进行过采样这两种重采样方法。本文结合 2 种采样方法各自的优势提出 RENN-BorderlineSMOTE 混合采样方法，既能避免单独使用过采样方法可能会增加冗余数据从而增加过拟合风险，以及导致数据分布不合理甚至会脱离真实情况，也避免了单独使用降采样方法可能导致数据集中有价值的信息被舍弃的风险。

Wilson 等[12]提出了编辑最近邻（edited nearest neighbours，ENN）算法。ENN 算法利用一个样本的近邻样本的类别分布情况来对数据进行处理，识别并移除那些与多数邻近样本类别不一致的样本。例如，设定近邻样本数量为 5，那么那些与至少 3 个近邻样本类别不同的样本将被移除，即当选定样本与其近邻样本类别相同的个数小于 3 时，选定样本会被移除，不对其他样本进行操作。ENN 算法示意图如图 1 所示。通过多次应用 ENN 算法，可以进一步消除噪声样本，这种方法被称为重复编辑最近邻（RENN）欠采样算法。

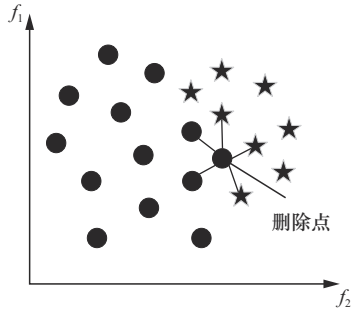

图 1　ENN 算法示意图

BorderlineSMOTE 算法旨在通过创建合成样本来增加少数类样本的数量，它对 SMOTE 算法进行了优化。SMOTE 算法的不足是它忽略了少数类样本之间的个体差异，直接在少数类样本和其近邻样本之间进行随机线性插值，插入的样本不一定位于类别边缘，而实际检测中位于边界附近的样本更容易被分类错误。

BorderlineSMOTE 算法将边界样本的分类纳入了考虑范围，综合了原始 SMOTE 算法以及边界样本的信息，在类别边缘生成新的样本，避免了 SMOTE 算法过多生成无益于模型学习的样本。在 BorderlineSMOTE 算法中，设 T_{maj} 和 T_{min} 分别为多数类样本集和少数类样本集，对 T_{min} 中的第 i 个样本 x_i 寻找数据集中的 k 个近邻样本，并记其中不同类别样本的数量为 k_i。若 $k_i=k$，则视 x_i 为噪声点，不做任何操作。若 $\frac{k}{2}<k_i<k$，则视 x_i 为危险点，将其加入集合 D 并对 D 中的样本使用 SMOTE 算法进行少数类样本合成。若 $0<k_i<\frac{k}{2}$，则视 x_i 为安全点，不做任何操作。BorderlineSMOTE 中的 3 类样本示意图如图 2 所示，图 2 中 $k=5$。

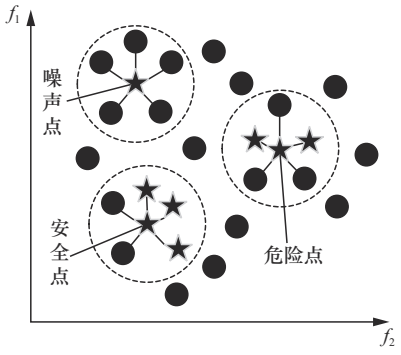

图 2　BorderlineSMOTE 中 3 类样本示意图

提出的 RENN-BorderlineSMOTE 混合采样方法首先利用 RENN 算法去除除最少数类之外的类别的特定的样本，然后利用 BorderlineSMOTE 算法生成除最多数类之外的类别的样本，以此实现优化决策边界、清除噪声以及解决数据类别不平衡的问题。

1.3　动态集成模型——KNORAE

传统静态集成在进行分类预测时，仅选出一个固定的基分类器组合，然后将其集成并适用

于所有的测试样本进行分类预测，不能根据不同测试样本之间的差异性“量身定制”适合参与每个测试样本预测的基分类器。而动态集成技术在进行预测时，需要将大量机器学习模型拟合到训练数据集，针对不同的测试样本从基分类器池中动态地选择不同的基分类器参与预测，所以动态集成相较于传统的静态集成能更好提高集成模型的分类性能。基分类器池的选择机制同时考虑了池中模型的准确性和多样性，选择一组相互补充的基分类器能更好提高模型的整体性能。

KNORAE是一种动态集成模型，其主要目标是为每个测试样本动态地选择一组最能准确预测其类别的基分类器用于集成。首先，对于每个测试样本，找到动态选择集中距离最近的k个样本作为该测试样本的胜任域。然后利用每个基分类器对该测试样本的胜任域中的每个样本进行预测，以准确度为指标对每个基分类器进行评估。在KNORAE中，一个基分类器只有对胜任域内的所有样本都分类正确时才被认为是胜任的，即对于每个基分类器，检查其在胜任域内的所有预测是否与真实标签一致，如果一致，那么该分类器在该胜任域内的胜任度被认为是高的，将该分类器加入对应测试样本的优分类器组中，KNORAE将使用这些分类器利用多数投票法对测试样本进行预测。如果没有任何分类器在给定胜任域内完全分类正确，那么会先缩小k值以减小胜任域，并对每个基分类器的胜任度重新进行评估。如果在最小的胜任域内仍然没有找到任何能完全分类正确的分类器，KNORAE将集成基分类器池中的所有基分类器运用多数投票法对测试样本进行预测，或者采用其他策略。

1.4 多样性指标

研究表明，集成多个基分类器得到的集成系统的性能比任何参与集成的单一分类器更好，而这要求每个参与集成的基分类器有较高的精确率并且与参与集成的其他分类器互补[13]。这种性能的提升源于集成学习能够整合各个基分类器的预测结果，从而提高模型的泛化能力。

选择性集成的核心目标在于减少模型中的分类器数量的同时保持或提升模型的预测精度，从而降低模型的存储需求和计算复杂度。

广义多样性（generalization diversity，GD）由Partridge等[14]提出。在随机选取的两个基分类器中，当一个基分类器能够正确分类而另一个基分类器错误分类时，这两个基分类器之间的多样性被认为是最高的。相反，如果两个基分类器在对任意样本的分类中总是同时出错，那么它们的多样性程度被认为是最低的。令p_i表示恰好有i个分类器错误分类样本的概率，L表示参与计算GD值的基分类器的数量，则广义多样性定义为：

$$\mathrm{GD}=1-\frac{\sum_{i=1}^{L}\frac{i}{L}p_i}{\sum_{i=1}^{L}\frac{i(i-1)}{L(L-1)}p_i} \tag{4}$$

GD值的范围为从0到1。GD值为0则表示参与对比的基分类器在所有样本上的分类错误完全相同，即参与对比的基分类器之间没有差异；而GD值达到1时，表示参与对比的基分类器对任意样本的预测结果不可能同时出错，即参与对比的基分类器之间的差异性达到最大。

2 实验结果

实验的硬件环境：Windows 11操作系统，超威半导体（advanced micro devices，AMD）Ryzen 7 5800 H with Radeon Graphics中天处理器（central processing unit，CPU），16 GB内存。软件环境：Python3.9。

2.1 数据集描述

Sharafaldin等[15]创建的CIC-IDS2017数据集是入侵检测领域研究的通用数据集。数据集中的流量类型包括良性（BENIGN）、Web攻击（Web Attack）、DoS攻击、端口扫描（Sniffing）攻击、暴力（Brute-Force）攻击、渗入威胁（Infiltra-

tion）攻击、僵尸网络（Botnets）攻击共7种流量类型，本文对每个类别的数据均按照6:22的比例进行随机划分后进行合并成为混合的训练集、验证集和测试集，CIC-IDS2017数据集分布见表1。

2.2　检测流程

基于Powershap和混合采样的动态集成入侵检测模型包括以下步骤，模型总体结构如图3所示。

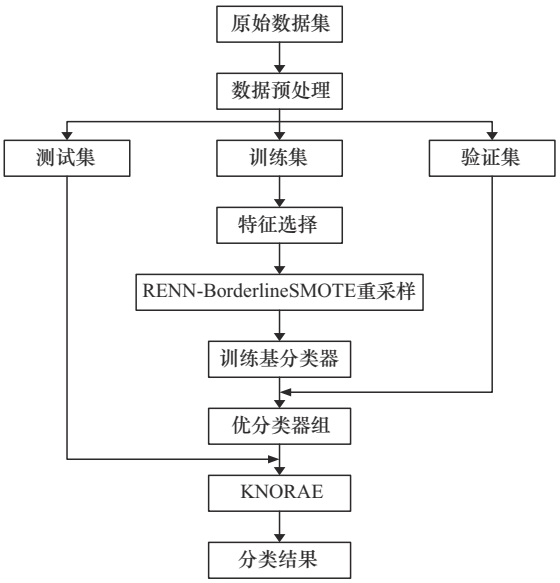

图3　模型总体结构

（1）进行了数据清洗、标签编码和数据归一化操作。将含有缺失值、无穷值、Nan值的样本去除。将标签信息转化为整数值，以便于分类算法的训练和预测。将数值型数据进行归一化，提升模型的训练速度。本文使用min-max归一化将数据集中的数值型特征放缩到[0,1]的范围内。其转换计算式为：

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

其中，x'表示对x进行归一化处理后得到的结果，x表示某一特征值，$x_{\max}$表示该特征属性值中的最大值，$x_{\min}$表示该特征属性值中的最小值。

（2）对数据集进行预处理后，将数据集如表1所示划分为训练集、验证集、测试集，将训练集引入Powershap算法进行特征筛选。

（3）经过特征选择后，将特征子集利用RENN-BorderlineSMOTE混合采样算法处理训练数据类别不平衡的问题。先利用RENN算法对除Infiltration类之外类别的样本进行降采样处理，随后再利用BorderlineSMOTE算法对除BENIGN类之外类别的样本进行过采样处理，异常流量样本重采样后各类别之间的比例与重采样前相同，重采样后所有异常流量样本总数等于正常流量样本的数量，得到的流量类别分布，训练集重采样前后数据对比见表2，重采样后的训练集用于最终模型训练。

（4）利用训练集训练10个基分类器，基于它们的检测性能和多样性指标筛选出优分类器组。

（5）模型构建，将筛选出的优分类器组集成到动态集成框架KNORAE中进行分类预测，最后输出测试样本的分类结果。

表1　CIC-IDS2017 数据集分布

类型	训练集		验证集		测试集	
	数量	比例	数量	比例	数量	比例
BENIGN	1 362 791	80.32%	454 264	80.32%	454 264	80.32%
Web Attack	1 308	0.07%	436	0.08%	436	0.08%
DoS	227 849	13.43%	75 949	13.43%	75 949	13.43%
Sniffing	95 282	5.62%	31 761	5.62%	31 761	5.62%
Brute-Force	8 299	0.49%	2 767	0.49%	2 766	0.49%
Infiltration	22	0.001%	7	0.001%	7	0.001%
Botnets	1 174	0.07%	391	0.07%	391	0.07%

表2 训练集重采样前后数据对比

类型	重采样前		重采样后	
	数量	比例	数量	比例
BENIGN	1 362 791	80.32%	1 355 801	50.00%
Web Attack	1 308	0.07%	5 311	0.20%
DoS	227 849	13.43%	925 087	34.12%
Sniffing	95 282	5.62%	386 853	14.27%
Brute-Force	8 299	0.49%	33 695	1.24%
Infiltration	22	0.001%	89	0.003%
Botnets	1 174	0.07%	4 767	0.18%

2.3 评价指标

本文采用4个评价指标来验证所提模型的性能，分别是准确率（Accuracy）、召回率（Recall）、精确率（Precision）、F1值（F1-score）。本文所选评价指标计算如式（6）～式（9）所示。

准确率：分类器分类正确的样本数量与总样本数量之比，是一个用于评估入侵检测模型整体性能的指标。

$$\text{Accuracy}=\frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}} \tag{6}$$

召回率：分类器正确识别某一类别的数量占该类别数量的比例。

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{7}$$

精确率：分类器正确识别某一类别的数量占被分类器识别为该类别的数量的比例。

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{8}$$

F1值：召回率和精确率的加权平均值。

$$\text{F1}-\text{score}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{9}$$

其中，真阳性（true positive，TP）表示某个特定类别，模型正确地将属于该类别的样本预测为该类别的样本数量；真阴性（true negative，TN）表示对于某个特定类别，模型正确地将不属于该类别的样本预测为不属于该类别的样本数量；假阳性（false positive，FP）表示预测样本属于某个类别，但实际上样本并不属于那个类别的样本数量；假阴性（false negative，FN）预测一个样本不属于某个类别，但实际上样本属于那个类别的样本数量。

对于入侵检测模型而言，准确率、精准率、召回率和F1值越高，则所提出的入侵检测模型的检测性能越好。

2.4 选择性集成

参与集成的基分类器的选择对于集成模型的性能起到至关重要的作用，集成模型要求参与集成的每个单一基分类器均有着较好的检测性能并且与其他参与集成的多数不同基分类器之间具有较大的差异性。根据多种不同的训练原理共选出10个基分类器用于筛选，各基分类器在CIC-IDS2017数据集上的性能见表3。

表3 各基分类器在CIC-IDS2017数据集上的性能

基分类器	准确率	精确率	召回率	F1值
Adaboost	78.61%	73.64%	78.61%	75.97%
DT	99.89%	99.89%	99.89%	99.89%
ET	99.88%	99.87%	99.88%	99.87%
GNB	38.21%	89.73%	38.21%	46.03%
GBDT	99.36%	99.36%	99.36%	99.33%
LR	87.95%	91.20%	87.95%	88.56%
MLP	96.76%	97.12%	96.76%	96.82%
RF	99.78%	99.79%	99.78%	99.78%
LGBM	98.41%	98.41%	98.41%	98.40%
KNN	99.23%	99.25%	99.23%	99.24%

由表3可以看出，Adaboost、GNB、LR这3个基分类器的准确率远低于其他7个分类器，并且其余指标的表现也较其余7个分类器差，故先排除Adaboost、GNB、LR这3个基分类器参与集成。而DT的性能表现最佳，故优先选取DT作为参与集成的1个基分类器。因此，保留DT、ET、GBDT、MLP、RF、LGBM、KNN这7个基分类器作为下文对比分析各不同基分类器之间差异性所用的个体分类器。

在集成学习中，不仅要保证参与集成的基分类器的性能较好，还需要所选参与集成的几个基分类器之间的差异度较大。不同的基分类器采取不

同的方式对数据进行预测，各有优劣，而集成学习正是利用这一点，用1个基分类器的优势去补充其他基分类器的不足，因此由差异度较大的数个基分类器进行集成，更有利于提升集成模型的检测性能。各基分类器之间的GD值见表4，展示了7个基分类器之间的广义多样性。

表4　各基分类器之间的GD值

基分类器	ET	GBDT	MLP	RF	LGBM	KNN
DT	0.266	0.790	0.954	0.547	0.912	0.798
ET	—	0.794	0.949	0.547	0.898	0.752
GBDT	—	—	0.739	0.513	0.728	0.861
MLP	—	—	—	0.876	0.743	0.738
RF	—	—	—	—	0.878	0.839
LGBM	—	—	—	—	—	0.908

由表4可以看出，其中MLP、LGBM与DT之间的广义多样性较大，即它们的差异性较大，能够较好补充DT的不足，故选取MLP、LGBM作为集成模型中的基分类器。ET与DT之间的广义多样性较小，即它们之间的差异很小，故不选取ET参与集成。KNN与LGBM差异性较大，但与MLP差异性较小，故不选取。RF与MLP、LGBM之间的广义多样性也较大，虽然RF与DT之间的广义多样性较小，但仍可一定程度上扩大基分类器的差异度，补充其余基分类器的不足，故选取RF作为集成模型中的基分类器。

综上所述，最终选取DT、MLP、RF、LGBM作为集成模型的基分类器。

2.5　实验结果与分析

2.5.1　类平衡方法对比

为了量化RENN-BorderlineSMOTE综合采样算法对不平衡样本的影响，本文设计了一组BorderlineSMOTE、自适应合成采样（adaptive synthetic sampling，ADASYN）、重复编辑最近邻共3种采样算法与本文所提综合采样算法的对比实验。实验中RENN算法中设置sampling_strategy为not minority，max_iter为200，其余参数为默认值，BorderlineSMOTE算法中random_state为0，sampling_strategy设置为表2中重采样后训练集各类别样本比例，其余参数为默认值。

采样方法的数据对比见表5，直观展示了不同重采样算法在CIC-IDS2017数据集中进行数据平衡后模型的检测性能，可以发现在准确率、精确率、召回率、F1值上RENN-BorderlineSMOTE的表现均为最好，这一现象表明RENN-BorderlineSMOTE在重采样时，能够更有效提高分类器的识别能力。

表5　采样方法的数据对比

采样算法	准确率	精确率	召回率	F1值
B-SMOTE	99.842%	99.841%	99.842%	99.818%
ADASYN	99.832%	99.833%	99.832%	99.810%
RENN	99.892%	99.890%	99.892%	99.889%
本文模型	99.900%	99.899%	99.900%	99.899%

2.5.2　特征选择的比较

本文进行了有无特征选择的对比实验以验证所用特征选择方法对模型性能的影响，有无特征选择的数据比较见表6，有无特征选择模型训练时间对比如图4所示。原始数据集特征数为78，经特征选择后，特征维数降至13，经特征选择后特征数量显著减少。

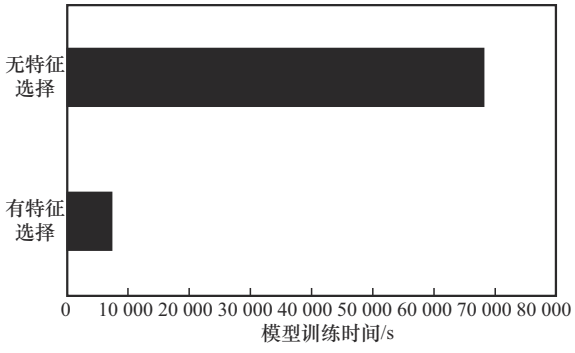

图4　有无特征选择模型训练时间对比

从表6中可以看出，经过特征选择后的模型的性能较无特征选择更好，说明了特征选择对整体模型的性能有着较为重要的贡献。

表6　有无特征选择的数据比较

	准确率	精确率	召回率	F1值
无特征选择	99.894%	99.893%	99.894%	99.893%
有特征选择	99.900%	99.899%	99.900%	99.899%

从图4中可以看出，经过特征选择后的模型的训练时间较无特征选择显著缩短，说明进行特征选择对整体模型的运行效率有很大提升。

2.5.3　与传统算法对比

与传统算法的数据对比见表7，这是关于基于Powershap和混合采样的动态集成入侵检测模型、参与其集成的单一基分类器算法及多数投票法的检测性能对比。

表7　与传统算法的数据对比

算法	准确率	精确率	召回率	F1值
DT	99.886%	99.886%	99.886%	99.886%
MLP	96.736%	97.104%	96.736%	96.796%
RF	99.771%	99.780%	99.771%	99.774%
LGBM	98.421%	98.433%	98.421%	98.413%
多数投票	99.784%	99.777%	99.784%	99.777%
本文模型	99.900%	99.899%	99.900%	99.899%

从表7可以看到，本文模型相比于参与集成的单一基分类器的准确率上升至99.900%，同时其各性能指标相较于各单一基分类器均有提升，并且本文模型性能也更优于多数投票算法。

2.5.4　与其他算法对比

本文模型与其他模型在各种性能指标方面进行比较，与其他模型的数据对比见表8，与其他模型训练时间对比如图5所示。从结果可以看出，提出的模型在准确率、精确率、召回率和F1值上均取得了较好的效果，且在同一环境、同一数据量下本文模型训练时间上较对比算法快，与对比模型进行运行时间对比的环境配置如下：Intel十六核处理器（i7-14650HX），64位操作系统，Windows 11。

算法1[4]使用k均值综合少数过采样技术来处理数据集中不平衡的正常流量数据和异常流量数据，再结合自动编码器进行分类预测，能够减少数据维数，去除噪声数据，并捕获更有效的特征信息从而能够有效地从大规模流量数据中提取非线性结构信息。文中数据集采用CIC-IDS2017数据集中周一、周五2天的部分进行实验，数据总量与类别数均少于本文实验采用的数据集。

表8　与其他模型的数据对比

算法	准确率	精确率	召回率	F1值
算法1	99.16%	99.16%	98.28%	98.22%
算法2	99.84%	99.88%	99.85%	99.86%
本文模型	99.90%	99.90%	99.90%	99.90%

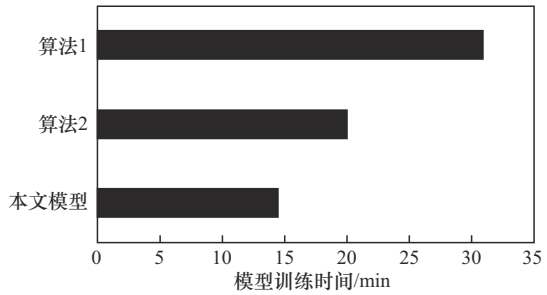

图5　与其他模型训练时间对比

算法2[5]利用提出的对抗量化变分自编码器来解决数据不平衡问题，使用ResNet与改进的分段残差神经网络对输入的样本数据进行联合学习并预测其攻击类型。文中数据集采用CIC-IDS2017数据集中删除Infiltration类别之后的子集，数据总量与类别数均少于本文实验采用的数据集。

3　结束语

随着网络流量规模的不断扩大以及网络异常流量数量和种类的增长，对性能更好的入侵检测模型的需求更加迫切。本文研究将Powershap、RENN-BorderlineSMOTE混合采样算法与动态集成模型相结合，提出了一种用于处理特征维度高的不平衡数据集的模型。采用Powershap作为特征选择算法减少了数据集的特征维度，提高了模型的检测效率，使用RENN-BorderlineSMOTE混

合采样算法处理数据集存在的类不平衡问题，利用基分类器之间的广义多样性筛选出相互补足并能提高整体性能的优分类器组进行集成。通过多组对比实验表明，本文所提模型的效果表现较好。

参考文献：

[1] ANDERSON J P, Computer security threat monitoring and surveillance[R]. 1980.

[2] 何红艳, 黄国言, 张炳, 等. 基于极限树特征递归消除和LightGBM的异常检测模型[J]. 信息网络安全, 2022, 22(1): 64-71.
HE H Y, HUANG G Y, ZHANG B, et al. Intrusion detection model based on extra trees-recursive feature elimination and LightGBM[J]. Netinfo Security, 2022, 22(1): 64-71.

[3] 陈俊彦, 卢贤涛, 黄雪锋, 等. 基于Double-Bagging特征降维异质集成入侵检测[J]. 计算机工程与科学, 2023, 45(6): 1011-1019.
CHEN J Y, LU X T, HUANG X F, et al. Double-Bagging based feature dimension reduction heterogenous integrated intrusion detection[J]. Computer Engineering & Science, 2023, 45(6): 1011-1019.

[4] 蹇诗婕, 刘岳, 姜波, 等. 基于聚类过采样和自动编码器的网络入侵检测方法[J]. 信息安全学报, 2023, 8(6): 121-134.
JIAN S J, LIU Y, JIANG B, et al. Network intrusion detection using cluster oversampling and auto-encoder[J]. Journal of Cyber Security, 2023, 8(6): 121-134.

[5] 魏明军, 李凤, 刘亚志, 等. 基于改进WGAN-GP和ResNet的车联网入侵检测方法[J]. 郑州大学学报(工学版), 2024, 45(4): 30-37.
WEI M J, LI F, LIU Y Z, et al. An intrusion detection method for Internet of vehicles based on improved WGAN-GP and ResNet[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(4): 30-37.

[6] 孙敬, 丁嘉伟, 冯光辉. 一种基于自编码器降维的神经卷积网络入侵检测模型[J]. 电信科学, 2025, 41(2): 129-138.
SUN J, DING J W, FENG G H. A neural convolutional network intrusion detection model based on autoencoder dimension reduction[J]. Telecommunications Science, 2025, 41(2): 129-138.

[7] 徐会彬, 方龙, 张莎. 车联网中基于stacking集成学习的攻击检测模型[J]. 电信科学, 2024, 40(12): 38-50.
XU H B, FANG L, ZHANG S. Attack detection model based on stacking ensemble learning for Internet of vehicles[J]. Telecommunications Science, 2024, 40(12): 38-50.

[8] VERHAEGHE J, VAN DER DONCKT J, ONGENAE F, et al. Powershap: a power-full shapley feature selection method[M]// Machine Learning and Knowledge Discovery in Databases. Cham: Springer International Publishing, 2023: 71-87.

[9] TMOEK I. An experiment with the edited nearest-neighbor rule[J]. IEEE Transactions on Systems, Man, and Cybernetics, 2007, SMC-6(6): 448-452.

[10] HAN H, WANG W Y, MAO B H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning[M]//Advances in Intelligent Computing. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005: 878-887.

[11] KO A H R, SABOURIN R, BRITTO J. From dynamic classifier selection to dynamic ensemble selection[J]. Pattern Recognition, 2008, 41(5): 1718-1731.

[12] WILSON D L. Asymptotic properties of nearest neighbor rules using edited data[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1972, SMC-2(3): 408-421.

[13] 孙博, 王建东, 陈海燕, 等. 集成学习中的多样性度量[J]. 控制与决策, 2014, 29(3): 385-395.
SUN B, WANG J D, CHEN H Y, et al. Diversity measures in ensemble learning[J]. Control and Decision, 2014, 29(3): 385-395.

[14] PARTRIDGE D, KRZANOWSKI W. Software diversity: practical statistics for its measurement and exploitation[J]. Information and Software Technology, 1997, 39(10): 707-717.

[15] 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, 2018: 108-116.

[作者简介]

黄冬梅（1964− ），女，上海电力大学电气工程学院教授，主要研究方向为电力与海洋时空信息技术。

颜昊（2000− ），男，上海电力大学电子与信息工程学院硕士生，主要研究方向为网络入侵检测。

张文博（1992- ），男，上海海洋大学信息学院讲师，主要研究方向为形式化验证、理论计算机科学。

孙锦中（1980- ），男，上海电力大学电子与信息工程学院讲师，主要研究方向为电力时空信息技术。

胡安铎（1983- ），男，上海电力大学电子与信息工程学院讲师，主要研究方向为电力时空信息技术。

孙园（1980- ），男，上海电力大学数理学院副教授，主要研究方向为数据分析挖掘与建模。