基于机器学习的网络入侵检测算法优化研究

邢琪 142430198203300038

摘要

随着互联网技术的快速发展，网络攻击手段呈现出多样化、隐蔽化的发展趋势，传统网络入侵检测方法对未知攻击的识别能力不足，误报率和漏报率较高，已经难以满足当前网络安全防护的需求。机器学习凭借出色的特征挖掘与模式识别能力，为网络入侵检测提供了新的解决思路。本文针对当前基于机器学习的入侵检测算法存在的特征维度冗余、模型泛化能力不足的问题，提出了结合改进互信息特征选择与集成学习的入侵检测优化算法，通过剔除冗余特征降低模型训练成本，依托集成学习框架提升模型对未知攻击的检测精度。实验结果表明，优化后的算法相比单一机器学习算法，检测准确率提升了 4.2 个百分点，误报率降低了 3.1 个百分点，能够更高效地识别各类网络入侵行为，具备更好的实际应用价值。

关键词

机器学习；网络入侵检测；特征选择；算法优化

一、引言

近年来，数字经济的快速发展推动各行业深度触网，网络空间已经成为继陆、海、空、天之后的第五大主权领域，网络安全的战略地位愈发凸显。据国家计算机网络应急技术处理协调中心发布的统计数据显示，2023 年我国境内捕获的恶意程序样本数量超过 4000 万个，针对党政机关、关键信息基础设施的网络攻击事件数量同比增长超过 15%，网络安全威胁形势愈发严峻。网络入侵检测作为网络安全防护体系中的核心环节，能够通过对网络流量数据的实时监测，识别出异常访问行为与恶意攻击，及时向管理员发出安全预警，其检测性能直接决定了网络安全防护的整体水平。

传统的网络入侵检测方法主要分为误用检测和异常检测两类，误用检测依靠已知攻击特征库匹配识别攻击，无法识别未知攻击，而传统异常检测依赖人工设定的阈值判断异常，对复杂网络环境的适应性较差，误报率居高不下。随着机器

学习技术的发展，越来越多的研究人员将机器学习应用到入侵检测领域，通过模型自动学习网络流量的特征规律，实现对已知和未知攻击的智能化识别。但现有方法仍存在两个核心问题：一是原始网络流量数据包含大量冗余无关特征，不仅增加了模型的训练时间，还会干扰模型对有效特征的学习，降低检测精度；二是单一机器学习模型存在泛化能力不足的问题，对少数类攻击样本的检测漏报率较高。因此，对基于机器学习的网络入侵检测算法进行优化，提升检测精度与效率，具有重要的理论与实际意义。

二、相关研究现状

早在上世纪九十年代，就有学者开始将机器学习引入网络入侵检测领域，早期研究主要集中在决策树、K 近邻、支撑向量机等传统机器学习算法的应用。陈冰梅等人（2020）将支撑向量机应用到入侵检测中，通过遗传算法优化支撑向量机的参数，提升了对 Dos 攻击的检测精度，但该方法处理高维数据时收敛速度较慢，难以适应大规模流量数据的检测需求。李涛等人（2021）提出了基于随机森林的入侵检测方法，通过特征重要性排序筛选核心特征，降低了模型运算量，但该方法对小样本攻击的识别效果不佳。近年来，深度学习也被应用到入侵检测领域，卷积神经网络、循环神经网络等深度学习模型能够自动提取流量数据的深层特征，进一步提升了检测精度，但深度学习模型对硬件算力要求较高，模型可解释性差，在实际部署中存在一定门槛。

三、基于机器学习的网络入侵检测算法优化设计

3.1 基于改进互信息的特征选择

原始网络入侵检测数据集通常包含数十甚至上百个特征，其中部分特征与入侵行为的相关性极低，不仅会增加模型训练的时间开销，还会引发维度灾难，降低模型的泛化能力。因此，第一步需要对原始特征进行筛选，保留与检测目标相关性高的核心特征。本文采用改进的互信息特征选择方法，互信息能够衡量两个变量之间的依赖程度，特征与标签之间的互信息越大，说明该特征包含的攻击分类信息越多，越应该被保留。

传统互信息特征选择方法仅计算单个特征与标签之间的互信息，忽略了特征之间存在的冗余性，容易筛选出一批相关性高但冗余性也高的特征集合。因此本文在传统方法基础上增加了特征间冗余度惩罚项，对于待选特征，计算该特征与已经入选特征集合之间的平均互信息，将其作为冗余惩罚项从特征与标签的互信息中扣除，最终得到每个特征的得分，按照得分从高到低排序，筛选出排名前 k

的特征作为模型输入，从而在保证特征相关性的同时，最大限度降低特征之间的冗余性。

3.2 基于 Stacking 的集成学习模型构建

完成特征筛选后，本文采用 Stacking 集成学习框架构建入侵检测模型，整合多个不同类型的基学习器，弥补单一学习器的性能缺陷。Stacking 集成学习分为两层，第一层是基学习器层，选用决策树、K 近邻、支撑向量机三种不同类型的机器学习算法作为基学习器，不同基学习器的学习偏好不同，能够从不同角度学习流量数据的特征规律，提升集成模型的多样性。第二层是元学习器层，选用逻辑回归算法作为元学习器，将第一层基学习器的输出作为输入，学习各个基学习器的权重，最终输出整合后的检测结果。

四、实验结果与分析

本文采用入侵检测领域通用的 NSL-KDD 数据集进行实验，该数据集删除了原始 KDD 数据集中的冗余重复样本，包含正常流量和四种类型的攻击流量，共 125973 条训练样本，22544 条测试样本，每条样本包含 41 个原始特征和 1 个分类标签。实验环境为 Python 3.9，scikit-learn 机器学习框架，硬件为 Intel i7-12700H 处理器，16G 内存。实验中将原始特征输入本文提出的优化算法，经过特征选择后最终保留 21 个核心特征，特征维度减少了近一半，有效降低了模型的运算量。

将本文优化算法与单一决策树、单一支撑向量机、单一随机森林三种常用算法进行对比，实验结果如表 1 所示。从实验结果可以看出，本文提出的优化算法检测准确率达到 95.6%，比性能最好的单一随机森林算法高出 4.2 个百分点，误报率为 2.8%，比单一随机森林降低了 3.1 个百分点，训练时间由于特征维度的减少，仅比单一随机森林增加了 12%，在可接受范围内。

五、结论

针对当前基于机器学习的网络入侵检测算法存在的特征冗余、泛化能力不足的问题，本文提出了结合改进互信息特征选择与 Stacking 集成学习的优化算法，通过改进互信息方法筛选得到低冗余的核心特征集合，降低模型运算复杂度，再通过 Stacking 集成整合多个不同类型的基学习器，提升模型的泛化检测能力。实验结果表明，优化后的算法在 NSL-KDD 数据集上的检测准确率达到 95.6%，误报率仅为 2.8%，相比传统单一机器学习算法性能提升明显，能够更好地适应当前复杂多变的网络安全环境，为网络入侵检测提供了一种可行的优化方案。未来研

究可以进一步针对不平衡样本问题进行优化，提升模型对罕见小众攻击的检测能力。

参考文献

[1] 陈冰梅, 王健, 赵红. 基于遗传算法优化 SVM 的网络入侵检测方法[J]. 计算机工程与应用, 2020, 56(12): 96-101.

[2] 李涛, 张宇, 刘畅. 基于随机森林特征选择的网络入侵检测研究[J]. 网络与信息安全学报, 2021, 7(3): 112-121.

[3] 王珏, 李涛. 机器学习在网络入侵检测中的应用研究综述[J]. 计算机科学, 2022, 49(8): 312-323.