文章编号：1003－1251(2024)05－0007－07

基于 PER-PPO2 的入侵检测技术

黄迎春，任国杰

（沈阳理工大学 信息科学与工程学院，沈阳 110159）

摘　要：随着万物信息化与智能化的快速发展，网络攻击范围不断扩大。传统的入侵检测算法，如主成分分析(PCA)结合随机森林和K近邻等，由于网络数据繁多，特征提取能力较差，分类准确率低。针对上述问题，提出一种新的入侵检测技术，称为优先经验采样的近端策略优化裁剪(prioritized experience replay-proximal policy optimization clip，PER-PPO2)算法，基于强化学习实现包裹法特征选择。深度强化学习通过构建以分类器混淆矩阵为基础的奖励函数，使智能体根据奖励反馈选择分类器的较优特征，结合优先经验采样优化算法的训练样本，提高算法的稳定性与收敛性能；使用性能较优的轻量级梯度提升机(LightGBM)作为分类器。使用NSL-KDD数据集对模型进行实验评估，结果表明模型将数据集的41维特征降低为8维时分类F1值达到0.871 3，可以满足入侵检测的要求。

关　键　词：近端策略优化裁剪；优先经验采样；入侵检测；深度强化学习；LightGBM

中图分类号：TP309　　**文献标志码：**A　**DOI：**10.3969/j.issn.1003－1251.2024.05.002

Intrusion Detection Technology Based on PER-PPO2

HUANG Yingchun，REN Guojie

（Shenyang Ligong University，Shenyang 110159，China）

Abstract：With the rapid development of informatization and intelligence of all things，the scope of network attacks continues to expand. Traditional intrusion detection algorithms，such as principal component analysis(PCA)combined with random forests and K-nearest neighbors，have poor feature extraction capabilities and low classification accuracy in the face of the numerous features of current network data. In response to the above problems，a new intrusion detection technology is proposed，called Proximal Policy Optimization Pruning with Prioritized Experience Sampling(PER-PPO2). This algorithm implements wrapping method feature selection based on reinforcement learning. Reinforcement learning constructs a reward function based on the classifier confusion matrix，allowing the agent to select the better features of the classifier based on reward feedback；combined with the training samples of the priority experience sampling optimization algorithm，Improve the stability and convergence performance of the algorithm；use the lightweight gradient boosting machine(LightGBM)with better performance as the classifier. The NSL-KDD data set was used to conduct an experimental evaluation of the model. The results showed that when the model reduced the 41-dimensional features of the data set to 8 dimensions，the classification F1 value reached 0.871 3，which can meet the requirements of intrusion detection.

Key words：proximal policy optimization clip；prioritized experience replay；intrusion detection；deep reinforcement learning；lightweight gradient boosting machine

收稿日期：2023－06－20

基金项目：国家自然科学基金项目(61971291)

作者简介：黄迎春(1976—)，男，副教授，研究方向为网络服务与信息安全。

随着万物互联的不断发展,传统行业逐步信息化[1]。互联网设备的激增导致了网络系统变得复杂,有效防护来自网络中的攻击行为已成为当前亟需解决的问题。入侵检测系统[2]是一种积极主动的安全防护技术,可以通过实时监视网络流量感知网络攻击并提供响应决策,其在军事、医疗、交通、物联网安全、工业控制等领域均有广泛应用。

入侵检测系统一般分为特征提取与分类器训练两个部分。Selvakumar 等[3]利用萤火虫算法和特征选择器提高了分类器的性能,通过 10 个特征即可完成入侵检测。Bala 等[4]使用粒子群优化增强遗传算法进行参数优化并结合随机森林算法进行入侵检测,该模型在 NSL-KDD 数据集[5]上取得了良好的效果。黄迎春等[6]通过萤火虫算法优化加权贝叶斯算法的权值,提高了加权贝叶斯算法的性能。何红艳等[7]利用 K-means 算法提取典型数据,再使用递归特征消除算法结合逻辑回归算法去除重要性低的特征。决策树构建树状模型进行分类,降低了数据维度并提高了分类精度,但模型的评估指标少,数据维度仍然较高。为了提高入侵检测模型的性能,部分研究人员运用了集成学习算法[8]。唐朝飞等[9]使用主成分分析(PCA)进行数据降维,然后采用集成学习算法中的轻量级梯度提升机(LightGBM)进行分类,再使用改良的粒子群算法进行参数优化,但是新特征解释性弱,PCA 也无法对非线性数据进行处理。文献[10]提出了自编码器 – 轻量级梯度提升机(AE-LightGBM)模型,使用自编码器(AE)将高维数据压缩至低维,LightGBM 分类提高了模型的分类性能与实时性,然而 AE 不直接使用与目标变量相关的信息,在特征选择过程中可能无法充分考虑到目标变量的关系。

随着硬件设备的不断更新,研究人员将深度学习与入侵检测相结合。Zhao 等[11]利用一维卷积神经网络实现数据特征的提取,使模型的性能得到改善,然而新特征的解释性依旧较差。舒豪等[12]使用长短期记忆网络(BiLSTM)和深度神经网络,考虑特征前后信息的影响后提取显著性特征,并使用激活函数 Softmax 输入数据的概率分布进行分类,模型的性能有所提升。Kushwah 等[13]使用自适应差分进化算法结合极限学习机提取特征,在不同数据集中取得了良好的效果,但测试时间略长,实时性能较差。Yang 等[14]改进密度峰值聚类算法,结合深度置信网络降维分类,与传统随机森林等算法相比,虽然模型的性能明显提高,但仍有进一步改进的空间。Ieracitano 等[15]结合 AE 将数据的特征转换到浅层空间并用 Softmax 输出类别概率分布得到结果,然而传统深度学习模型通常较大,模型的解释性较差,对资源的要求也较多。

强化学习[16]作为一种新兴技术,在不断发展的过程中通过与神经网络结合形成深度强化学习,增强了强化学习的普适性。文献[17]采用符号变换的特征构造和深度 Q 学习的特征提取方法,相比原始分类器模型,各项指标都有明显提升,但未使用入侵检测数据集进行实验。文献[18]使用深度强化学习中的深度 Q 网络算法(DQN)对入侵检测数据进行分类,并验证了其性能。强化学习自身特性契合包裹法特征选择,且有众多的改进算法,适用于入侵检测领域。

综合上述文献,本文提出基于包裹法的特征提取方法,旨在提高入侵检测中的特征提取能力。包裹法具有以下优点:根据分类性能进行特征选择、能够捕捉特征之间的非线性关系、更好地探索特征子集;特征子集与分类指标联系直观。考虑到集成学习的性能更优,提出一种利用深度强化学习实现包裹法特征选择和集成学习分类的入侵检测方法,称为优先经验采样的近端策略优化裁剪(prioritized experience replay-proximal policy optimization clip, PER-PPO2)。在增强特征提取能力的同时维持分类性能。由于该强化学习环境简单,使用浅层的深度神经网络(deep neural network, DNN),可减少参数计算量。

1 数据集及预处理

本文使用 NSL-KDD 数据集进行实验分析。NSL-KDD 数据集的训练集与测试集设置合理,可用于入侵检测领域不同研究的比较和评估。数据由 41 个特征组成,其中包括 3 个字符特征。

针对原始特征的差异性,首先采用标签编码对字符特征进行数值化;为了方便判断模型分类的性能,将数据的标签设为 1 和 0,0 表示正常数据,1 表示异常数据;为使模型更易收敛,需要对数据进行归一化处理,将所有特征数值映射至0 ~ 1 之间。

表 1 和表 2 展示了 NSL-KDD 数据集中异常攻击类型与各类别数据的分布情况。

表 1　NSL-KDD 数据集异常攻击类型

Table 1　NSL-KDD dataset exception attack types

攻击大类	攻击子类
Dos	**Apache2**, Back, Land, Neptune, **Mailbomb**, Pod, **Processtable**, Smurf, Teardrop, **Udpstorm**, **Worm**
Probe	Ipsweep, **Mscan**, Nmap, Portsweep, **Saint**, Satan
U2R	Buffer _ overflow, Loadmodule, Perl, **Ps**, Rootki, **Sqlattack**, **Xterm**
R2L	Ftp_write, Guess_passwd, **Httptunnel**, Imap, Multihop, **Named**, Phf, **Sendmail**, **Snmpgetattack**, Spy, **Snmpguess**, Warezclient, Warezmaster, **Xlock**, **Xsnoop**

如表 1 所示,该数据集包含了 4 大类 39 种攻击,其中用粗体显示的 17 种攻击只在测试集中出现,更加考验模型的泛化性。

表 2 中 NSL-KDD 训练集与测试集的类别分布不平衡,能更好模拟现实环境。

表 2　NSL-KDD 各类型数据分布

Table 2　NSL-KDD data distribution

类型	训练集	测试集
正常	67 343	9 711
Dos	45 927	7 458
Probe	11 656	2 421
U2R	52	200
R2L	995	2 754

2　入侵检测模型

本文使用 PER-PPO2 构建自主选择较优特征子集的入侵检测模型,其中数据分类由 LightGBM 进行,基于 PER-PPO2 的入侵检测模型框架如图 1 所示。

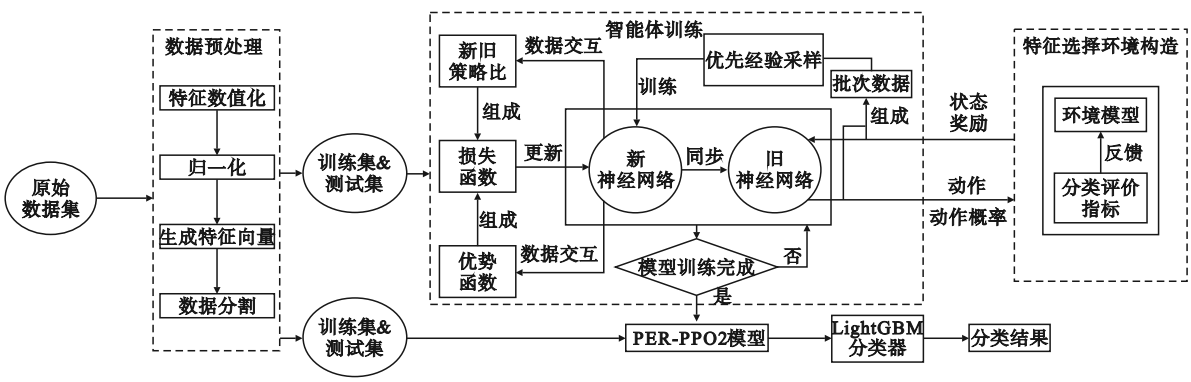

图 1　入侵检测模型框架

Fig. 1　Framework of intrusion detection mode

2.1　PPO2 算法

PPO2[19]是一种基于策略的强化学习算法,通过构建神经网络使智能体与环境进行交互,经环境的反馈修改智能体动作,使智能体找到解决目标问题的较优方法。智能体与环境的一次交互过程如下:设单次迭代任意 t 步时,状态为 s_t,智能体根据状态 s_t 输出动作 a_t 并与环境交互;环境返回状态 s_{t+1},并向智能体反馈奖励 r_t 以评估智能体行为;智能体重复上述过程,在多轮迭代中不断调整其输出以改进其行为,直至训练结束。

本文中,PPO2 方法将状态空间 $\boldsymbol{S}$ 表示为特征子集的所有可能组合,动作空间 $\boldsymbol{N}$ 为特征的选择过程,智能体的动作受到一定限制,即被选择过的特征应该被排除,将分类器的评估指标作为奖励,采用的奖励公式为

$$r = \mathrm{F1} - \omega_1 \frac{2 \times \mathrm{FPR} \times \mathrm{FNR}}{\mathrm{FPR} + \mathrm{FNR}} \tag{1}$$

式中:F1 值为模型精确率和召回率的调和平均数,可更好地评估模型性能;FPR、FNR 是模型误报率与漏报率,与模型的精确率、召回率线性相关,为避免奖励反馈时误报率、漏报率的任意上升,使用 FPR 和 FNR 的调和平均值作为奖励的惩罚项;ω_1 为惩罚项的权重。算法根据奖励对智能体进行反馈,使智能体选择的特征可获得更大的奖励,最终选出优良的特征。

为了确定最优特征子集,PPO2 算法结合了策略网络与价值网络共同决策,两个网络在训练过程中协同工作,从而产生用于特征选择的关键训练数据。神经网络通过生成的训练数据更新网络模型,实现最优特征子集的选择。

图 2 为单次迭代的 M 步交互中神经网络训练数据的生成过程。具体说明为:随机初始化策略网络参数 $\boldsymbol{\theta}^*$、价值网络参数 $\boldsymbol{\theta}$;策略神经网络与环境交互,得到初始状态(空特征子集 $\boldsymbol{s}_1$);基于 $\boldsymbol{s}_1$

输出特征动作的选择概率 p_1 和具体特征动作 a_1；策略神经网络再次将 $\boldsymbol{s}_1$、a_1 与环境交互，得到选择特征动作 a_1 后的不完全特征子集 $\boldsymbol{s}_2$、奖励 r_1、特征子集选择完成的标志。由上述步骤构建的原始训练数据为 $(\boldsymbol{s}_1, p_1, a_1, r_1, Done)$，*Done* 表示特征子集选择是否完成的标志。基于 $\boldsymbol{s}_2$ 循环上述操作直至 M 步交互完成。

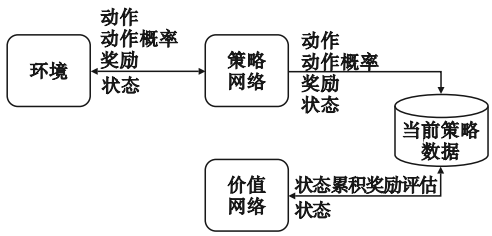

图 2 神经网络训练数据生成过程

Fig. 2 Generation of neural network training data

在单次迭代中，考虑每步动作对其后各步奖励的影响，在每步基于奖励值计算累积奖励。任意 t 步时，累积奖励公式如式(2)所示。

$$G_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots + \gamma^M r_M \tag{2}$$

式中：G_t 为 t 步时的累积奖励；γ 为折扣因子；r_t 为 t 时刻获得的即时奖励；r_M 为 M 步(最后一步)的累积奖励。

经过 h 次特征选择迭代，积累批次数据对策略和价值神经网络进行训练更新，以实现最优特征子集的选择，即根据式(3)、式(4)更新参数 $\boldsymbol{\theta}^*$、$\boldsymbol{\theta}$，使神经网络输出的特征子集最优。

$$\boldsymbol{\theta}^* \rightarrow \operatorname{argmax}\left[\sum_{s_t \in S} p(s_t)\boldsymbol{V}(\boldsymbol{s}_t)\right] \tag{3}$$

式中：$\boldsymbol{s}_t$ 为单次迭代的任意 t 步时智能体基于概率选择的特征子集；$V(\boldsymbol{s}_t)$ 表示 $\boldsymbol{s}_t$ 的累积奖励值期望，由价值网络估计得出；$p(\boldsymbol{s}_t)$ 为 t 步时出现 $\boldsymbol{s}_t$ 的概率值。使用梯度上升的方式更新参数 $\boldsymbol{\theta}^*$，累积奖励值较大的特征子集的概率值变大，求出较优特征子集。由于在使用梯度上升更新参数 $\boldsymbol{\theta}^*$ 时需要用到 $\boldsymbol{s}_t$ 选择具体动作 a_t 后的实际累积奖励值 Q，为了减小梯度上升的计算方差，算法引入了一个偏置量 b 与 Q 做差，当 b 值与 $V(\boldsymbol{s}_t)$ 相等时方差达到最小。b 与 Q 的差值命名为 t 步优势函数值 A_t，其不仅减小了梯度的方差，而且可以衡量当前实际动作的好坏，构成最终训练数据 $(\boldsymbol{s}_t, p_{t'}, a_t, A_t, Done)$。

$$\boldsymbol{\theta} \rightarrow \operatorname{argmin}\left[(\boldsymbol{V}_{\text{target}}(\boldsymbol{s}_t) - \boldsymbol{V}(\boldsymbol{s}_t))^2\right] \tag{4}$$

式中 $\boldsymbol{V}_{\text{target}}(\boldsymbol{s}_t)$ 为状态 $\boldsymbol{s}_t$ 的实际累积奖励值期望，通常使用近似方式获得，本文中使用 G_t 近似。通过最小化实际累积值与估计值的均方误差来更新价值网络的参数 $\boldsymbol{\theta}$。由于使用了大量多样数据训练，价值网络估计的累积奖励值期望逐渐接近实际值。

为了充分利用训练数据，对模型进行多次训练，PPO2 初始化了两对相同的策略神经网络 $\boldsymbol{\theta}^*$、$\boldsymbol{\theta}^{*'}$。通过重要性采样技术使得 $\boldsymbol{\theta}^{*'}$ 的训练数据可以用于 $\boldsymbol{\theta}^*$。同时，PPO2 算法将价值网络与策略网络的损失函数统一，以更好地学习到有效的策略和更准确地进行价值估计。通过计算损失函数的梯度求取 $\boldsymbol{\theta}^*$ 与 $\boldsymbol{\theta}$ 的目标值，其损失函数如式(5)所示。

$$(\boldsymbol{\theta}^*, \boldsymbol{\theta})_{\text{loss}} = -\min\left(\frac{\boldsymbol{p}_{\theta^*}}{\boldsymbol{p}_{\boldsymbol{\theta}^{*'}}}A, \operatorname{clip}\left(\frac{\boldsymbol{p}_{\theta^*}}{\boldsymbol{p}_{\theta^{*'}}}, 1-\epsilon, 1+\epsilon\right)A\right) + C_1(\boldsymbol{V}_{\text{target}} - \boldsymbol{V})^2 \tag{5}$$

式中：$(\boldsymbol{\theta}^*, \boldsymbol{\theta})_{\text{loss}}$ 代表策略神经网络与价值神经网络的共同损失函数，将策略网络取负实现两者共同梯度下降；C_1 代表价值神经网络的损失函数权重，价值网络不断更新使预估值逼近实际值；A 表示通过神经网络计算出的优势函数值；$\boldsymbol{p}_{\theta^*}/\boldsymbol{p}_{\theta^{*'}}$ 为两个策略网络生成的训练数据概率比值项，使 $\boldsymbol{\theta}^{*'}$ 生成的训练数据可用于 $\boldsymbol{\theta}^*$，该比值也作为更新步长的一部分，使两个策略网络同步参数且限制策略更新的幅度，确保训练的稳定性；使用 clip 函数剪切比值；ϵ 为 clip 剪切的范围，选择策略比值与剪切值两者的最小值来保证训练的稳定性。在训练完毕后将 $\boldsymbol{\theta}^*$ 的参数赋予 $\boldsymbol{\theta}^{*'}$，以完成下一批次的数据生成与策略更新。

2.2 PER 方法

PER 是一种增强学习中的经验回放方法，传统的基于价值函数的强化学习方法，如 Q-learning、DQN 等算法在训练过程中随机从经验池中选择样本进行训练，导致一些重要的经验较少被选择到，从而影响训练效果。为增强重要样本的训练，引入一个优先级机制，即根据样本的重要性进行选择。

虽然 PPO2 算法在更新过程中使用当前神经网络生成的所有数据进行更新，但为使重要的样本可以得到更多次数的训练，提高网络的训练效果与效率，拟采用 PER 方法对训练数据进行非均匀采样。本文采用 PPO2 算法的优势函数值 A 作为训练数据的优先级，优先级计算公式为

$$\beta_j = \mathrm{e}^{A_j} \tag{6}$$

式中 A_j 表示训练数据 j 的优势函数值；β_j 表示数据 j 的优先级。优势函数值大表示当前动作具有积极效果，应该增加采取该行动的概率。使用

Sumtree 数据结构[20]结合优先级进行具体的采样,通过具有优先级的数据节点对数据进行随机采样,采样概率为

$$P_j = \frac{\beta_j}{\sum_{j=1}^{k}\beta_j} \tag{7}$$

式中:k 为所有数据的数量;P_j 为数据 j 的采样概率,优先级高的数据采样概率更高,而优先级较低的数据也有被采样的机会。

在 PER-PPO2 网络更新的过程中,首先对生成的所有训练数据计算优势函数得出数据优先级,然后根据优先级将数据存入 Sumtree 中,最后通过多次采样数据训练神经网络,并更新数据的优先级。网络更新过程如图 3 所示。

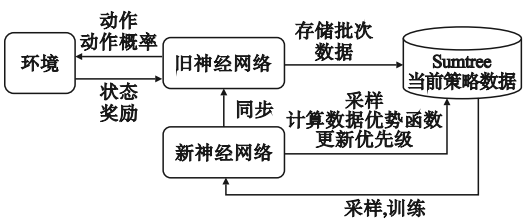

图 3　PER-PPO2 网络更新过程

Fig. 3　PER-PPO2 network update process

3　实验结果与分析

3.1　实验环境

本文算法模型的训练与测试实验环境如表 3 所示。

表 3　实验环境

Table 3　Experimental environment

软硬件环境	版本型号
Python	3.9
PyTorch	1.11.0 + cu113
Scikit-learn	1.2.1
CPU	Intel(R)Core(TM)i7 – 11800H
GPU	NVIDIA GeForce RTX 3060 Laptop
RAM	16 GB

3.2　实验结果

实验使用准确率、精确率、召回率、F1 值作为模型性能的评估指标,分类结果混淆矩阵如表 4 所示。

表 4　分类结果混淆矩阵

Table 4　Classification result confusion matrix

	预测攻击	预测正常
真实攻击	真阳例(TP)	假阴例(FN)
真实正常	假阳例(FP)	真阴例(TN)

为了检验模型的性能,限制特征选择数量为 7 ~ 10,不同数量特征得出的准确率如图 4 所示,根据准确率将特征数目设置为 8。

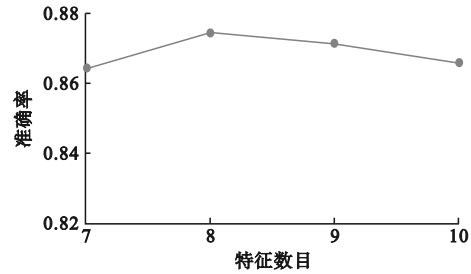

图 4　NSL-KDD 测试集中模型不同特征数量准确率对比

Fig. 4　Comparison of the accuracy of models with different number of features in the NSL-KDD test set

超参数的选择直接影响模型的性能,本文通过多次实验选取 PER-PPO2 算法的超参数如表 5 所示。

表 5　PER-PPO2 超参数列表

Table 5　PER-PPO2 hyperparameter list

参数	含义	数值
ω_1	奖励函数中惩罚项权重系数	0.1
lr	梯度下降学习率	0.007
h	生成批次数据需要迭代的次数	64
ϵ	clip 函数裁剪截断范围	0.2
episode	PER-PPO2 总迭代次数	1 025
k_epoch	单批次数据更新模型的次数	10
γ	奖励折扣因子	0
M	单次迭代交互步数(特征选择数)	8

本文模型中通过分类指标构建奖励函数,因此奖励值大小可反映模型的性能。图 5 为训练过程中 PER-PPO2 算法与 PPO2 算法在 NSL-KDD 测试集中奖励值的对比结果。

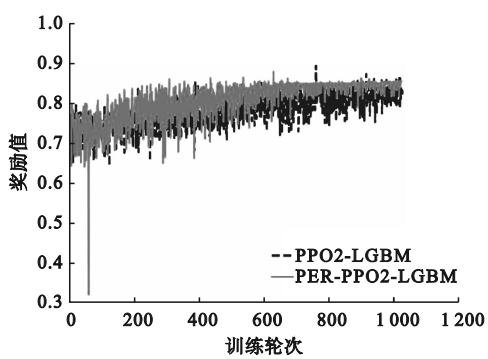

图 5　PER-PPO2 与 PPO2 奖励值对比

Fig. 5　Comparison of PER-PPO2 and PPO2 reward values

由图 5 可以看出,本文 PER-PPO2 的奖励值波动小于 PPO2,且最终奖励高于 PPO2,说明了 PER 的有效性。

为了测试实验模型的性能,以 NSL-KDD 数据集为基准,将实验模型与传统的特征提取分类算法 PCA + 随机森林、PCA + K 近邻、未使用优先经验采样的 PPO2 + LightGBM 在 NSL-KDD 测试集上进行对比。将各个模型的特征选择数目均设置为 8,实验结果如表 6 所示。

表 6 NSL-KDD 测试集上不同模型性能对比(1)

Table 6 Comparison of the performance of different models on the NSL-KDD test set(1)

模型	准确率	精确率	召回率	F1 值
PCA + 随机森林	0.763 2	0.969 9	0.602 7	0.743 4
PCA + K 近邻	0.753 6	0.966 0	0.587 7	0.730 8
PPO2 + LightGBM	0.837 0	0.845 8	0.741 0	0.838 0
本文模型	**0.874 3**	**0.879 9**	**0.801 8**	**0.878 9**

由表 6 可见,本文提出的模型在 NSL-KDD 测试集中的精确率低于 PCA + 随机森林、PCA + K 近邻,但其他指标都远高于对比模型。

为进一步验证模型性能,将本文模型与其他文献中基于 NSL-KDD 数据集的模型进行性能对比,实验结果如表 7 所示。可见,本文提出的模型在 NSL-KDD 测试集中的准确率、精确率与 F1 值均高于对比模型,仅召回率较低,由于精确率与召回率为负相关关系,F1 值为两者调和平均值,综合评估四项指标,本文模型在入侵检测中整体性能较好。

表 7 NSL-KDD 测试集上不同模型性能对比(2)

Table 7 Comparison of the performance of different models on the NSL-KDD test set(2)

模型	准确率	精确率	召回率	F1 值
文献[9]	0.864 0	0.873 4	0.864 0	0.866 1
文献[12]	0.790 3	0.785 7	**0.868 4**	0.825
文献[14]	0.820 8	N/A	0.705 1	N/A
文献[15]	0.842 1	0.87	0.803 7	0.819 8
本文模型	**0.874 3**	**0.879 9**	0.801 8	**0.878 9**

4 结论

由于冗余特征对分类造成干扰,使用 PPO2 算法结合优先经验采样与入侵检测环境交互选取特征子集,降低特征数量,并采用 LightGBM 算法进行分类。该方法基于特征子集与分类结果的直观联系,适用于各种分类算法。实验表明,本文算法在将特征维度降低至 8 时,仍有良好的分类性能,但算法的训练时间较长。

参考文献(References):

[1] 马标,金映言,那幸仪,等. 工业控制系统入侵检测技术研究综述[J]. 计算机应用与软件,2023,40(5):10 – 18,43.
MA B,JIN Y Y,NA X Y,et al. A summary of research on industrial control system intrusion detection technology[J]. Computer Applications and Software,2023,40(5):10 – 18,43. (in Chinese)

[2] DWIVEDI S,VARDHAN M,TRIPATHI S,et al. Implementation of adaptive scheme in evolutionary technique for anomaly-based intrusion detection[J]. Evolutionary Intelligence,2020,13(1):103 – 117.

[3] SELVAKUMAR B,MUNEESWARAN K. Firefly algorithm based feature selection for network intrusion detection[J]. Computers & Security,2019,81:148 – 155.

[4] BALA R T,INDIA R S G H. A review on KDD Cup99 and NSL-KDD dataset[J]. International Journal of Advanced Research in Computer Science,2019,10(2):64 – 67.

[5] BALYAN A K,AHUJA S,LILHORE U K,et al. A hybrid intrusion detection model using EGA-PSO and improved random forest method[J]. Sensors,2022,22(16):5986.

[6] 黄迎春,张雨薇. 一种改进的加权贝叶斯恶意软件识别方法[J]. 沈阳理工大学学报,2019,38(1):43 – 47.
HUANG Y C,ZHANG Y W. An improved weighted Bayesian malware recognition method[J]. Journal of Shenyang Ligong University,2019,38(1):43 – 47. (in Chinese)

[7] 何红艳,黄国言,张炳,等. 基于多种特征选择策略的入侵检测模型研究[J]. 信息安全研究,2021,7(3):225 – 232.
HE H Y,HUANG G Y,ZHANG B,et al. Research on intrusion detection model based on multiple feature selection strategies[J]. Journal of Information Security Research,2021,7(3):225 – 232. (in Chinese)

[8] DONG X B,YU Z W,CAO W M,et al. A survey on ensemble learning[J]. Frontiers of Computer Science,2020,14(2):241 – 258.

[9] 唐朝飞,努尔布力,艾壮. 基于 LightGBM 的网络入侵检测研究[J]. 计算机应用与软件,2022,39(8):298 – 303,311.
TANG C F,NURBOL,AI Z. Research on network intrusion detection based on LightGBM[J]. Computer Applications and Software,2022,39(8):298 – 303,311. (in Chinese)

[10] YAO R Z,WANG N,LIU Z H,et al. Intrusion detection system in the smart distribution network:a feature engineering based AE-LightGBM approach[J]. Energy Reports,2021,7:353 – 361.

[11] ZHAO G S,WANG Y,WANG J A. Intrusion detection model of internet of things based on LightGBM[J]. IEICE Transactions on Communications,2023,E106. B(8):622 – 634.

[12] 舒豪,王晨,史峄. 基于 BiLSTM 和注意力机制的入侵检测[J]. 计算机工程与设计,2020,41(11):3042 – 3046.
SHU H,WANG C,SHI Y. Intrusion detection based on BiLSTM and attention mechanism[J]. Computer Engineering and Design,2020,41(11):3042 – 3046. (in Chinese)

[13] KUSHWAH G S,RANGA V. Optimized extreme learning machine for detecting DDoS attacks in cloud computing[J]. Computers & Security,2021,105:102260.

[14] YANG Y Q,ZHENG K F,WU C H,et al. Building an effective intrusion detection system using the modified density peak clustering algorithm and deep belief networks[J]. Applied Sciences,2019,9(2):238.
[15] IERACITANO C,ADEEL A,MORABITO F C,et al. A novel statistical analysis and autoencoder driven intelligent intrusion detection approach[J]. Neurocomputing,2020,387:51 - 62.
[16] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015,518(7540):529 - 533.
[17] 张鹏,张瑞. 基于强化学习的特征选择方法及材料学应用[J]. 上海大学学报(自然科学版),2022,28(3):463 - 475.
ZHANG P,ZHANG R. Feature selection based on reinforcement learning and its application in material informatics[J]. Journal of Shanghai University (Natural Science Edition), 2022,28(3):463 - 475. (in Chinese)
[18] ALAVIZADEH H, ALAVIZADEH H, JANG-JACCARD J. Deep Q-learning based reinforcement learning approach for network intrusion detection[J]. Computers,2022,11(3):41.
[19] MAYER S, CLASSEN T, ENDISCH C. Modular production control using deep reinforcement learning:proximal policy optimization[J]. Journal of Intelligent Manufacturing,2021,32(8):2335 - 2351.
[20] 黄浩,胡智群,王鲁晗,等. 基于 Sumtree DDPG 的智能交通信号控制算法[J]. 北京邮电大学学报,2021,44(1):97 - 103.
HUANG H,HU Z Q,WANG L H,et al. Intelligent traffic signal control algorithm based on sumtree DDPG[J]. Journal of Beijing University of Posts and Telecommunications,2021,44(1):97 - 103. (in Chinese)

(责任编辑:和晓军)

(上接第 6 页)

[4] 张频捷,张立军,孟德建,等. 汽车车内噪声主动控制系统扬声器与麦克风布放优化方法[J]. 振动与冲击,2017,36(5):169 - 175.
ZHANG P J, ZHANG L J, MENG D J, et al. Vehicle ANC hardware optimal placement using Multi-objective genetic algorithm[J]. Journal of Vibration and Shock, 2017, 36(5): 169 - 175. (in Chinese)
[5] 张翔,李传光. 自适应有源噪声控制算法的研究与实现[J]. 北京理工大学学报,2002,22(1):53 - 55.
ZHANG X, LI C G. Study and implementation of the algorithm for adaptive active noise control[J]. Journal of Beijing Institute of Technology,2002,22(1):53 - 55. (in Chinese)
[6] 李维松,许伟杰,张涛. 基于小波变换阈值去噪算法的改进[J]. 计算机仿真,2021,38(6):348 - 351,356.
LI W S, XU W J, ZHANG T. Improvement of threshold denoising method based on wavelet transform[J]. Computer Simulation,2021,38(6):348 - 351,356. (in Chinese)
[7] 陆苗霞. 小波变换在信号去噪方面的仿真研究[J]. 科技创新与应用,2022,12(5):48 - 50.
LU M X. Simulation research on wavelet transform in signal denoising[J]. Technology Innovation and Application,2022, 12(5):48 - 50. (in Chinese)
[8] 祖丽楠,刘志远,生宁. 非平稳声信号下的小波变换去噪方法研究[J]. 现代电子技术,2022,45(11):35 - 40.
ZU L N,LIU Z Y,SHENG N. Research on wavelet transform denoising method for non-stationary acoustic signal[J]. Modern Electronics Technique,2022,45(11):35 - 40. (in Chinese)
[9] 陆真,裴东兴. 基于改进小波阈值法的语音去噪算法[J]. 电声技术,2016,40(4):39 - 44.
LU Z,PEI D X. Algorithm of image denoising based on the optimized method of wavelet thresholding[J]. Audio Engineering,2016,40(4):39 - 44. (in Chinese)
[10] 陈景良,李东新. 基于 LMS 的语音信号去噪算法[J]. 国外电子测量技术,2017,36(6):22 - 25,30.
CHEN J L,LI D X. Voice signal removal noise based on LMS algorithm[J]. Foreign Electronic Measurement Technology, 2017,36(6):22 - 25,30. (in Chinese)
[11] SHEN B B,LV X F,ZHANG S. An improved LMS adaptive filtering algorithm and its analysis[C]//2019 International Conference on Intelligent Computing, Automation and Systems(ICICAS). Chongqing,China:IEEE,2020:549 - 551.
[12] 颜昉. 基于小波变换的语音增强算法设计[D]. 湘潭:湘潭大学,2021.
[13] 吴瑶,张海霞. 一种变步长 LMS 自适应滤波的改进算法[J]. 通信技术,2021,54(2):307 - 311.
WU Y,ZHANG H X. An improved LMS adaptive filtering algorithm with variable step size[J]. Communications Technology,2021,54(2):307 - 311. (in Chinese)
[14] HE D H, WANG M J, HAN Y F, et al. Variable step size LMS adaptive algorithm based on exponential function[C]// 2019 IEEE 2nd International Conference on Information Communication and Signal Processing(ICICSP). Weihai, China: IEEE,2020:473 - 477.
[15] 仝喜峰,陈卫松,钱隆彦,等. 一种非线性变步长 LMS 自适应滤波算法[J]. 无线电通信技术,2019,45(4):391 - 396.
TONG X F,CHEN W S,QIAN L Y,et al. A nonlinear variable step size LMS adaptive filtering algorithm[J]. Radio Communications Technology, 2019, 45(4): 391 - 396. (in Chinese)
[16] 伍彩云,翁晶晶. 一种改进的自适应噪声抵消系统算法研究[J]. 沈阳理工大学学报,2022,41(3):1 - 7.
WU C Y,WENG J J. Algorithm study on an improved adaptive noise cancellation system[J]. Journal of Shenyang Ligong University,2022,41(3):1 - 7. (in Chinese)
[17] WU C,ZHANG W,WANG Y,et al. Study on the performance of the variable step-size LMS algorithms[C]//2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence(ICIBA). Chongqing,China:IEEE,2020:159 - 162.
[18] 于新颖. 自适应噪声抵消系统的 MATLAB 仿真与分析[J]. 山西电子技术,2020(3):14 - 16.
YU X Y. MATLAB simulation and analysis of adaptive noise cancellation system[J]. Shanxi Electronic Technology,2020 (3):14 - 16. (in Chinese)

(责任编辑:和晓军)