基于 SE-PCA-CNN 的网络入侵检测研究

徐海洋[1] 黄迎春[1]
XU Haiyang HUANG Yingchun

摘 要 随着互联网领域飞速发展，随之而来的恶意网络攻击的复杂性与隐蔽性也日益加剧，维护网络安全成为互联网稳定发展之路的必然焦点。针对传统的入侵检测系统无法有效的处理网络入侵时产生的高维、冗余数据问题，文章提出了一种基于合成少数类过采样（synthetic minority over-sampling technique，SMOTE）结合编辑最近邻欠采样（edited nearest neighbors，ENN）、主成分分析（PCA）和卷积神经网络（CNN）的入侵检测模型。首先使用 Smote+ENN 混合采样技术有效处理数据不平衡问题，再使用主成分分析法对高维数据进行特征降维，最后使用卷积神经网络对异常和正常流量数据进行分类，并在 NSL-KDD 入侵检测数据集上进行实验评估。结果表明，与传统 CNN 方法相比，基于 SE-PCA-CNN 的入侵检测表现出色，有着较高的准确率和更快的检测速度。

关键词 网络安全；入侵检测；混合采样；卷积神经网络

doi：10.3969/j.issn.1672-9528.2024.12.021

0 引言

近年来，互联网领域技术呈迅猛发展之势，网络科技的普及给人们的生活带来了诸多便利。但与此同时，如影随形的还有日益严峻的网络安全问题。如何在高效便捷地使用互联网技术的同时保障网络系统的安全与稳定，是如今网络环境所面临的巨大挑战。截至 2023 年 6 月，从互联网使用者遭遇各类网络安全问题的实际情况来看，其中遭遇个人信息泄露问题的网民占比达到 23.2%[1]。例如，黑客或不法分子通过各种手段在网络中散布病毒程序，进而非法盗取个人信息，严重危害到当下的互联网环境。

入侵检测系统（IDS）作为维护网络秩序的强有力手段也需要随科技的发展逐步更新升级。传统的入侵检测系统通过分析传输的网络流量，可以针对拒绝服务攻击（DoS）等攻击进行有效的防御[2]，但如今面对网络入侵所产生的海量高维数据时难以达到预期的防御效果。随着人工智能技术的发展，机器学习和深度学习方法被成熟且广泛地应用于维护网络安全中。例如，基于决策树和随机森林分类器的网络入侵检测系统，最早被 Nirupama 等人[3]提出，得益于决策树对最佳特征的划分以及随机森林集成多个决策树模型，该模型在 NSL-KDD 数据集上评估结果表现优异，对比其他基于传统分类器的模型具有更低的误报率和高检测率。

深度学习作为机器学习的一个重要分支，对比传统机器学习模型，表现出了更高的准确性、更强的拟合能力，如今已经成为众多学者的研究热点。Li 等人[4]将卷积神经网络首次引入到入侵检测系统中，其团队对 GoogLeNet 网络模型进行大量训练，实验结果准确率为 77.14%。虽然效果一般，但从一定程度上拓宽了神经网络的应用思路，具有很好的启蒙基奠意义。Yu 等人[5]提出了一种基于多尺度卷积神经网络（MSCNN）的高精度入侵检测系统，解决了入侵检测系统与神经网络的兼容问题，减少了处理参数，提高了收敛速度。

1 相关概念

1.1 SMOTE 过采样

数据类别不平衡是机器学习中的一个常见问题[6]，如果训练数据集中存在类分布不均衡的情况而不进行适当处理，在模型训练结束后则会出现严重偏差。为解决此问题，Chawla 等人[7]在实验研究中提出了 SMOTE（synthetic minority over-sampling technique）合成少数类过采样算法，目的在于对数据集中有着重要实际意义的少数类样本进行合理扩展，通过分析少数类样本周围的近邻样本信息，以插值的方式在其周围生成特征相似的新样本，防止在随机过采样时产生过拟合的现象，如图 1。SMOTE 算法的具体步骤如下：

（1）对于每一个少数类样本 x_i，以欧式距离为标准计算其本身到其周围所有少数类样本之间的距离。得到 x_i 的 k 个近邻值。

（2）从 x_i 的 k 个近邻值中任取一个样本，设为近邻 $\hat{x}_i$。

（3）对选取的随机近邻 $\hat{x}_i$，在 (0, 1) 的范围内生成一个

1. 沈阳理工的大学计算机科学与信息工程学院 辽宁沈阳 110159

随机数，最后重复上述操 N 次直至生成新的样本。其公式为：

$$x_{new} = x_i + \text{rand}(0, 1) \times (\hat{x}_i - x_i) \quad (1)$$

式中：$i = 1, 2, \cdots, N$；x_{new} 表示算法生成的少数类样本；x_i 表示原始的少数类样本；rand(0, 1) 表示在 (0, 1) 之间生成的随机数；$\hat{x}_i$ 表示样本特征为 3 维的 k 近邻值。

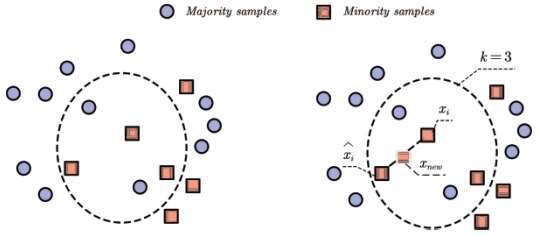

图 1 SMOTE 算法原理

1.2 ENN 欠采样

随机欠采样是处理不平衡数据集的常用方法之一，其技术简单而有效，核心思想是从多数类样本中随机选取一部分样本进行剔除，使得多数类样本与少数类样本的数量相近从而达到数据类别平衡。Alejo 等人 [8] 提出了改进神经网络分类的编辑最近邻规则，将 ENN 的思想实践应用在了当时流行的人工神经网络结构中，为人工智能的发展起到了重要的推动作用，如图 2。

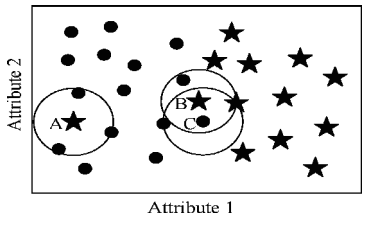

图 2 ENN 算法原理

ENN 算法描述为对于数据集中的每个样本 S_i 都要找到其 3 个最近的邻居。如果 S_i 属于多数类，并且至少有两个最近的邻居属于少数类，那么 S_i 将被删除。同时，如果 S_i 属于少数类，并且至少有两个最近的邻居属于多数类，那么 S_i 将被删除。

由于本实验所使用的 NSL-KDD 数据存在较为严重的数据类别不平衡问题，故采用 SMOTE+ENN 混合采样技术来处理此数据集，尽可能避免因类别不平衡而导致的模型训练精度差的问题。

1.3 主成分分析 PCA

主成分分析法（principal component analysis，PCA）[9] 是数据处理中一种常用的降维方法，其基本思想是将高维的 n 维特征映射为低维的 k 维特征，具体方法如下：

（1）通过协方差矩阵找到方差最大的变量，以确定其变化程度，将相关变量转换为互不相关的主成分，减少维度并保留信息。

（2）主成分的信息量由方差衡量。第 1 主成分是最大方差的线性组合，第 2 主成分与第 1 主成分正交，并具有第二大方差。后续的主成分依次是与之前所有主成分正交，并且在剩余方向中方差最大的线性组合。方差越大，主成分包含的信息越多。

通过这种方式，PCA 能够将原始数据中的相关变量转换为一组新的、互不相关的变量（主成分），并且这些主成分按包含的信息量排序。其主要步骤如下：

（1） 数据中心化：将数据矩阵 $\boldsymbol{X}$ 每列减去均值，得到中心化数据 $\boldsymbol{X}_i$。

$$\boldsymbol{X}_i = \boldsymbol{X} - \mu \quad (2)$$

（2）计算协方差矩阵：计算中心化数据的协方差矩阵 $\boldsymbol{\Sigma}$。

（3）特征值分解：对协方差矩阵 $\boldsymbol{\Sigma}$ 进行特征值分解，得到特征值 λ_i 和特征向量 $\boldsymbol{v}_i$。

$$\boldsymbol{\Sigma v}_i = \lambda_i \boldsymbol{v}_i \quad (3)$$

（4）选择主成分：选取最大特征值对应的 k 个特征向量组成矩阵 $\boldsymbol{V}_k$。

（5）投影数据：将原数据投影到新坐标系（主成分空间），得到降维后的数据 $\boldsymbol{Z}$。

$$\boldsymbol{Z} = \boldsymbol{X}_i \boldsymbol{V}_k \quad (4)$$

这样，PCA 通过分析主成分实现降维的同时尽可能保留数据的主要信息。

1.4 卷积神经网络

卷积神经网络（ convolutional neural network，CNN）是神经网络的一种特殊类型，最早由 Fukushima 等人 [10] 提出的一种神经网络模型，是一种专门用于处理图像数据的深度学习模型，擅长捕捉图像中的空间结构和模式。CNN 的核心思想是通过卷积操作提取局部特征，并逐层构建更复杂的表示。

2 SE-PCA-CNN 算法研究

入侵检测数据集天然存在数据类别不平衡的问题，如果不加以适当的处理，可能会导致模型无法充分学习，在进行多分类任务时无法准确识别少数类样本，极端情况下可能出现 0% 的准确率，显然用这样的数据训练模型是不可行的。故本文提出使用 SMOTE+ENN 混合采样，利用 PCA 进行特征降维，将流量数据转化为二维灰度图像，最后使用 CNN 将处理后的数据进行特征提取。通过在 NSL-KDD 入侵检测数据集上进行实验评估，结果表明该模型在少数类样本识别中准确率较高。

2.1 模型总体框架

针对传统入侵检测模型无法准确提取少数类特征的问题，提出一种基于 SE-PCA-CNN 的入侵检测模型，该模型总体框架如图 3 所示。

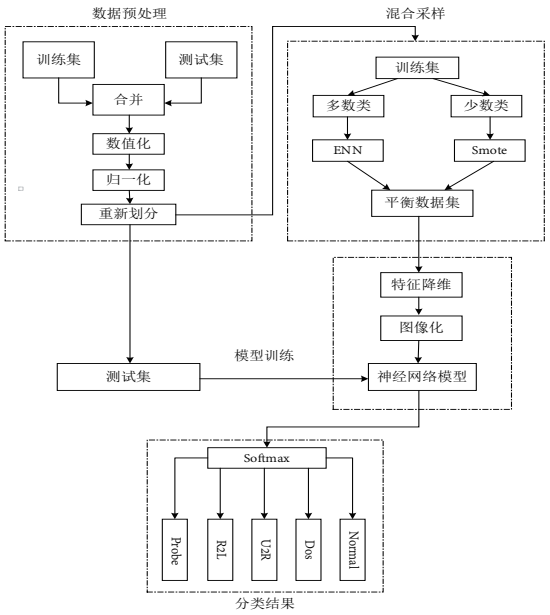

图 3 模型总体框架

该模型主要包含三个部分：数据预处理、混合采样模块和网络模型特征分类处理部分。

2.2 数据预处理

针对 NSL-KDD 入侵检测数据集中数据特征类型不一致和取值范围有差异等问题，在模型训练前做以下预处理操作：分别创建 KDDTrain 数据集和 KDDTest 数据集作为本实验的训练集和测试集。其数据类别及数量分布如表 1 所示。

表 1 实验所用数据类别与数量

数据类型	训练集	测试集
Normal	84 179	12 139
Dos	57 409	9323
U2R	65	250
Probe	14 570	3027
R2L	1244	3443

（1）数值化

该数据集第 2、3、4 维三个特征均属于字符型特征，故对此类特征采用独热编码（One-hot code）将其转换为数值型，进而输入到深度学习模型中进行训练。例如，协议类型的 3 个属性（TCP、UDP 和 ICMP）分别用 [1,0,0]、[0,1,0]、[0,0,1] 来表示，从而将一维的字符型特征转化为多维的数值特征。

（2）归一化

为提升网络模型的收敛速度，本文使用最大 - 最小（Max-Min）方法将特征缩放到 [0,1] 范围中，其公式为：

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{5}$$

式中：X 是原始数据；X' 是归一化后的数据；$X_{\min}$ 和 $X_{\max}$ 分别是数据的最小值和最大值。

2.3 混合采样与降维

使用 SMOTE 对少数类样本进行过采样，生成新的少数类样本，使少数类的数量增加。通过 ENN 清理数据，删除多数类样本中不一致或噪声样本，保持数据质量并减少多数类的样本数量，使得数据集在保持少数类信息的同时，去除了多数类中的噪声样本，增强了模型的泛化能力。

经过混合采样平衡后，其样本特征为 119 维，再使用 PCA 算法分析其主成分，当贡献率达到 99% 时，其特征已经降低到 64 维，最后将这些特征转为能够被 CNN 模型识别的图像，其像素均为 8×8。

2.4 卷积神经网络

图 4 从左到右依次为输入层、卷积层、批量标准化层（batch normalization，BN），最大池化层（Max-pooling）、激活函数（Relu）、全连接层（Fully Connected）和输出层。卷积层的卷积核大小为 3×3，数量为 16，步长为 1 并进行填充处理，激活函数设置为 Relu。

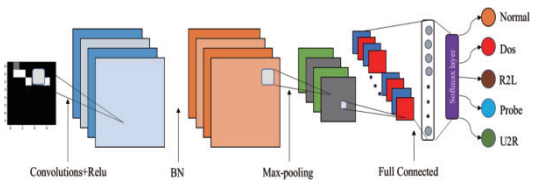

图 4 卷积神经网络结构图

模型训练过程如下：

（1）将得到的二维矩阵作为 CNN 的输入层输入。

（2）进入卷积层提取流量数据的特征，其中激活函数设置为 Relu 函数。

（3）BN 层可以加速网络训练，缓解梯度消失，帮助提升卷积神经网络的整体性能。

（4）该层采用最大池化，减少计算量的同时降低过拟合，保留重要特征。

（5）输出层采用 Softmax 函数求出分类结果。

3 实验结果与分析

3.1 实验环境

本文所用的硬件配置：Windows 10 操作系统，Intel i5 12500H CPU，16 GB 内存。

软件配置： Anaconda 3，Python3. 6，TensorFlow2.6。

3.2 评价指标

本文使用的性能评价指标为准确率（Acc）和召回率（Rec）。

准确率（Acc）：衡量模型整体的正确率，即分类任务中模型预测正确的频率。其公式为：

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{6}$$

召回率（Rec）：表示在所有实际为正类的样本中，模型正确预测为正类的比例，如表 2。反映了模型对正类样本的识别能力。其公式为：

$$Rec = \frac{TP}{TP + FN} \quad (7)$$

表 2 评估矩阵

真实情况	预测情况	
	正例	负例
正例	TP	FN
负例	FP	TN

3.3 结果分析

如图 5 所示，随着模型训练迭代次数的增加，其准确率逐渐上升并趋于稳定，迭代次数在 50 次左右时准确率达到最高值 99.4%。

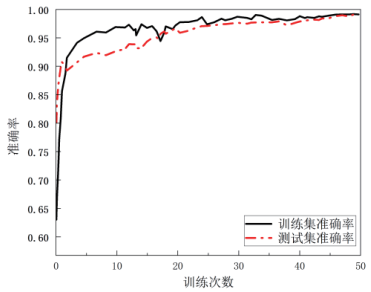

图 5 数据集训练与测试的准确率

采用消融实验评估模型的性能分别记录误报率、准确率和检测时间，如表 3 所示。

表 3 消融实验结果

模型	误报率 /%	准确率 /%	检测时间 /s
CNN	4.29	97.31	51.70
SE-CNN	1.43	99.35	46.83
SE-PCA-CNN	1.35	99.42	39.55

通过对比表 3 中实验结果的数据，可以看出结合了 SE-PCA 的模型相比于其他模型更具优势，在保证准确率高达 99.42% 的同时，其误报率相比于传统 CNN 减少到了 1.35%。在检测时间上，SE-PCA-CNN 模型也略优于其他模型，在面对海量冗余的真实入侵检测场景中，更能体现检测效率，由此可以证明本模型在入侵检测中的有效性和优越性。

4 结语

入侵检测系统作为保护网络安全的一道重要措施，在日益复杂的网络环境中起着关键的作用。本文提出了一种融合 SMOTE+ENN 混合采样、PCA 特征降维的卷积神经网络入侵检测模型。通过对入侵检测数据集进行预处理和采样，有效的剔除了无用的噪声样本并补充了少数类样本的数量，在模型训练中起到了的优化作用，可以有效提升入侵检测的性能。

参考文献：

[1] 中国互联网络信息中心 . 第 51 次《中国互联网络发展状况统计报告》[J]. 国家图书馆学刊 ,2023,32(2):39.

[2] MITCHELL R, CHEN I R. A survey of intrusion detection techniques for cyber-physical systems [J]. ACM computing surveys, 2014, 46(4):1-29.

[3] NIRUPAMA B K,NIRANJANAMURTHY M. Network intrusion detection using decision tree and random forest[C]//2022 International Conference on Advances in Computing,Communication and Applied Informatics(ACCAI). Piscataway: IEEE, 2022:1-9.

[4] LI Z P, QIN Z, HUANG K, et al. Intrusion detection using convolutional neural networks for representation learning[C]// International Conference on Neural Information Processing. Berlin:Springer,2017:858-866.

[5] YU J, YE X J, LI H B.A high precision intrusion detection system for network security communication based on multi-scale convolutional neural network[J].Future generation computer systems, 2022, 129(4):399-406.

[6] 李艳霞 , 柴毅 . 不平衡数据分类方法综述 [J]. 控制与决策 , 2019, 34(4): 673-688.

[7] CHAWLA N V, BOWYER K W, HALL L O,et al. SMOTE: synthetic minority over-sampling technique[J]. Journal of artificial intelligence research, 2002, 16(1):321-357.

[8] ALEJO R, SOTOCA J M, VALDOVINOS R M , et al.Edited nearest neighbor rule for improving neural networks classifications[C]//International Symposium on Neural Networks. Berlin: Springer, 2010: 303-310.

[9] WANG W, BATTITI R. Identifying intrusions in computer networks with principal component analysis[C]//First International Conference on Availability, Reliability and Security (ARES'06). Piscataway:IEEE, 2006: 279.

[10] FUKUSHIMA K. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36: 193-202.

【作者简介】

徐海洋（1996—），男，黑龙江齐齐哈尔人，硕士研究生，研究方向：网络安全。

黄迎春（1976—），通信作者（email: 1365153370@qq.com），男，辽宁瓦房店人，硕士，副教授、硕士生导师，研究方向：网络服务与信息安全。

（收稿日期：2024-09-13）