文章编号:1671-6833(2022)03-0037-07

基于 CNN 与 BiGRU 融合神经网络的入侵检测模型

张安琳[1], 张启坤[2], 黄道颖[2], 刘江豪[2], 李建春[2], 陈孝文[2]

(1. 郑州轻工业大学 工程训练中心,河南 郑州 450001;2. 郑州轻工业大学 计算机与通信工程学院,河南 郑州 450001)

摘 要: 针对深度学习入侵检测中出现的数据类不平衡及特征学习不全面等问题,提出了一种基于卷积神经网络(CNN)与双向门控循环单元(BiGRU)融合的神经网络入侵检测模型。通过 SMOTE-Tomek 算法完成对数据集的平衡处理,使用基于平均不纯度减少的特征重要性算法实现特征选择,将 CNN 和 BiGRU 模型进行特征融合并引入注意力机制进行特征提取,从而提高模型的总体检测性能。使用入侵检测数据集 CSE-CIC-IDS2018 进行多分类实验,并与经典单一深度学习模型进行对比。实验结果表明:在数据集平衡方面,经 SMOTE-Tomek 算法处理,DoS attacks-Slow HTTP Test 识别准确率从 0 提升至 34.66%,SQL Injection 识别准确率从 0 提升至 100%,DDoS attack-LOIC-UDP、Brute Force-Web 和 Brute Force-XSS 分别提升了 5.22 百分点、6.55 百分点和 35.71 百分点,证明了平衡后的数据集较未经过处理的数据集在少数类的识别精度上提升明显。在模型的总体检测性能方面,在多分类实验对比中,所提模型总的分类精确率、召回率以及 $F1$ 值均高于其他几种单一神经网络模型。其中各攻击流量类别的总评精确率比 LSTM 模型提升了 2.10 百分点;总评召回率比 LSTM 模型提升了 1.50 百分点;总评 $F1$ 值比 GRU 模型提升了 1.97 百分点,从而证明了该模型具有更好的检测效果。

关键词: 入侵检测; 卷积神经网络; 双向门控循环单元; SMOTE 算法; Tomek Links 算法

中图分类号: TP393;TP183 **文献标志码**: A **doi**:10.13705/j.issn.1671-6833.2022.03.003

0 引言

入侵检测系统(intrusion detection system,IDS)中常见的深度学习方法[1-3]主要有多层感知器(multilayer perceptron, MLP)、卷积神经网络(convolutional neural networks, CNN)[4]、循环神经网络(recurrent neural network, RNN)、自编码器(autoencoder, AE)等。其中,循环神经网络又包含长短期记忆网络(long shortterm memory, LSTM)和门控循环单元(gated recurrent unit, GRU)2 种变体。

Roy 等[5]将 MLP 应用于入侵检测,并与支持向量机进行对比,精度有明显的提升,但在特征数量较多、维度较高的环境中,仅 MLP 通常难以得到良好的训练效果。Wang 等[6]将处理过的流量数据转换成像素,利用 CNN 将流量数据以图片的形式进行输入和训练,得到较高的二分类准确率(100%)和多分类准确率(99.17%)。Naseer 等[7]使用 CNN、RNN 和 AE 等不同神经网络架构来构建分类模型,实验结果表明,CNN 和 LSTM 在入侵检测的分类中能够表现出良好的效果。Kim 等[8]使用 LSTM 构建检测系统,在 KDD-Cup99 数据集上取得了 96.93%的准确率,但误报率达到了 10.04%。Putchala[9]提出在物联网领域使用 GRU 进行入侵检测研究,但实验仅在 KDD-Cup99 数据集上进行,未实现应用于物联网相关数据的设想。王伟[10]结合 CNN 和 LSTM 对流量数据进行学习,在 CNN 提取流量数据空间特征的基础上再使用 LSTM 提取数据的时序特征,得到了较高的精度并保持较低的误报率,但该方法在 CNN 特征提取时可能会损失部分数据信息的时间特性。研究表明,这些基于深度学习的入侵检测系统在处理大数据时性能更好[11],但仍存在一些问题。

收稿日期:2021-08-21;**修订日期**:2021-11-16

基金项目:国家自然科学基金资助项目(61772477)

作者简介:张安琳(1971—),女,四川绵阳人,郑州轻工业大学高级实验师,主要从事计算机网络、人工智能研究,E-mail:alzhang@zzuli.edu.cn。

通信作者:黄道颖(1967—),男,河南信阳人,郑州轻工业大学教授,博士,主要从事计算机网络、智能网络研究,E-mail:dyhuang@zzuli.edu.cn。

(1) 数据集过时。以往的入侵检测研究大多基于 KDD-Cup99 和 NSL-KDD 数据集,距今已有近 20 年历史,不能很好地反映当下网络状况。

(2) 数据样本不平衡。分类研究通常更注重提升模型总体的评价指标,如准确率、精确率等,忽视了对少数类样本的分类问题。但在真实的网络环境中,这些少数类攻击会比多数类攻击产生更大的破坏和影响。然而,目前基于 KDD-Cup99 等数据集的研究通常都直接使用官方提供的训练和测试样本,较少有研究工作涉及入侵检测问题下的数据不平衡问题以及相关的解决方案[12]。

(3) 特征学习不全面。以往的研究大多基于单种神经网络:CNN 可以学习数据中的空间特征,精确地提取局部特征,但是学习不到时序特征;RNN 可以对数据中的时序特征进行提取,分析信息的长期依赖关系,但是不能有效提取空间特征[13],且 RNN 只能学习数据单一方向的时序特征,没有充分考虑到流量数据前后信息对当前状态的共同影响。

1 基于 CNN 与 BiGRU 融合神经网络入侵检测模型

为了解决入侵检测研究中出现的上述问题,本文提出一种基于 CNN 与 BiGRU(bidirectional gated recurrent unit)[14]融合神经网络入侵检测模型,结构如图 1 所示。本文模型主要由数据预处理模块和融合神经网络模块两部分组成,其特点有:

(1) 使用 2018 年最新的入侵检测数据集 CSE-CIC-IDS2018 进行模型的训练和测试[15],该数据集包含最新的网络攻击,且满足现实世界攻击的所有标准[16]。

(2) 对数据集进行清洗,并通过基于平均不纯度减少(mean decrease impurity,MDI)的特征选择算法进行特征优化降维,降低了计算机的资源消耗,提高了整体的计算效率。

(3) 使用 SMOTE-Tomek 算法对数据集进行了数据层面的平衡处理,避免了使用一般过采样方法产生的过拟合、样本“入侵”以及模糊边界等问题。

(4) 将一维数据转化为二维矩阵,使用 CNN 进行二维卷积,充分提取数据的空间特征;使用双向门控循环单元(BiGRU)进行数据时序特征的双向提取,充分学习数据的时间特征。

(5) 引入注意力机制(attention mechanism,AM)进行特征加权,提高模型的整体性能。

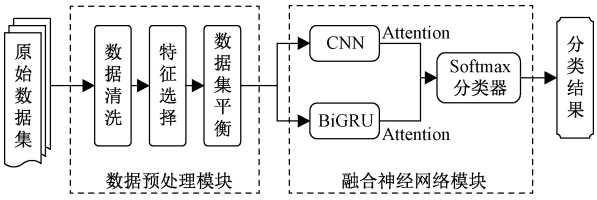

图 1 基于 CNN 与 BiGRU 融合神经网络入侵检测模型

Figure 1 Neural network intrusion detection model based on the fusion of CNN and BiGRU model

1.1 数据预处理模块

数据预处理部分主要包括原始数据集的清洗、特征选择、数据集平衡等操作。数据集的清洗采用的是常规方法,限于篇幅,不再介绍。

1.1.1 特征选择

(1) 无用特征删除。入侵检测系统应该根据流量信息的行为特征进行分类,避免偏向于特定标识所附带的信息,因此需要将涉及特定网络标识的特征删除;通过分析数据,将同一特征下数值相同的特征列删除。

(2) 特征重要性。特征重要性用来评估数据特征对入侵检测系统分类的影响程度,用其 Gini 重要性来衡量,通过计算特征重要性评分 *VIM* 可以筛选出数据的重要特征,从而进行特征优化。

假设有 m 个特征 $X_1, X_2, \ldots, X_m$,要计算出每个特征 X_j 的 Gini 指数评分 VIM_j,即第 j 个特征在随机森林所有决策树中节点分裂“不纯度”的平均改变量,Gini 指数用 GI 表示:

$$GI_m = \sum_{k=1}^{|K|}\sum_{k'\neq k} p_{mk}p_{mk'} = 1 - \sum_{k=1}^{|K|} p_{mk}^2。 \tag{1}$$

式中:K 表示有 K 个特征样本类别;p_{mk} 表示节点 m(将特征 m 逐个对节点计算 Gini 值变化量)中类别 k 所占的比例。

特征 X_j 在节点 m 的重要性,即节点 m 分枝前后的 Gini 指数变化量为

$$VIM_{jm} = GI_m - GI_1 - GI_r。 \tag{2}$$

式中:GI_1 和 GI_r 分别为节点 m 左、右分支的新节点的 Gini 指数。

如果特征 X_j 在决策树 i 中出现的节点在集合 M 中,则第 i 棵树中特征 X_j 的重要性为

$$VIM_{ij} = \sum_{m\in M} VIM_{jm}。 \tag{3}$$

在有 n 棵树的随机森林中,特征 X_j 的重要性为

$$VIM_j = \sum_{i=1}^{n} VIM_{ij}。 \tag{4}$$

将特征重要性进行归一化处理,得到特征重要性评分:

$$VIM_j = \frac{VIM_j}{\sum_{i=1}^{m} VIM_i}。\tag{5}$$

所有特征的特征重要性评分之和为 1,其特征重要性评分代表了该特征对数据分类的贡献大小。

1.1.2　数据集平衡

(1) SMOTE (synthetic minority over-sampling technique)算法。SMOTE 算法是数据层面应用最广泛的过采样算法[17]。SMOTE 算法有效解决了随机过采样方法产生的过拟合问题,SMOTE 算法为解决数据不平衡问题提供了新的思路,但存在一定的局限性:SMOTE 算法中根样本与辅助样本的选择决定了新样本的合成,然而在合成少数类样本时,忽视了多数类样本的分布情况。若根样本与辅助样本均处于少数类所在区域,则合成的新样本是符合要求的。但在实际运算中,如果根样本与辅助样本中有一个是噪声样本,那么合成的新样本将极有可能出现在多数类样本区域,出现合成的少数类样本入侵多数类样本空间的问题。此外,若 SMOTE 算法选定位于类边界的少数类样本合成新样本,且其 k 近邻样本也处于类的边界,那么经插值合成的少数类样本也同样会出现在两类的重叠区域,从而模糊两类的边界。

(2) Tomek Links 算法。Tomek Links 算法是一种数据欠采样算法[18],其基本原理:给定一些样本对 (x_i, x_j),其中 $x_i \in S_{\max}$,$x_j \in S_{\min}$,x_i 和 x_j 的距离记为 $d(x_i, x_j)$。若不存在任一样本 x_k 使得 $d(x_i, x_k) < d(x_i, x_j)$,则样本对 (x_i, x_j) 可称为 Tomek Links 对。通常情况下 Tomek Links 对中存在噪声样本或两样本位于两类的边界上。Tomek Links 算法应用于采样时,将剔除掉样本对中属于多数类的样本;应用于数据清洗时,样本对中噪声样本和边界重叠样本均被剔除。

为解决 SMOTE 算法导致的样本入侵问题以及模糊边界问题,将 Tomek Links 算法和 SMOTE 算法相结合,称之为 SMOTE-Tomek 算法,来解决数据的不平衡问题。首先使用 SMOTE 算法对原始数据进行过采样操作来合成少数类样本,然后使用 Tomek Links 算法对经过 SMOTE 算法处理的数据进行清洗,移除数据集中存在的 Tomek Links 对,从而过滤掉两类之间的噪声数据和重叠样本。

1.2　融合神经网络模块

融合神经网络模块主要由 CNN、BiGRU、注意力机制以及分类器构成。在入侵检测的数据分析及特征提取中,既需要分析空间层面的特征联系,也应该考虑到时间层面上特征之间的变化规律。融合神经网络可以提取局部平行特征,同时能够分析各特征点前、后信息对该特征点的影响。CNN 在图像处理上表现出更加优异的性能,因此,为了充分学习流量数据的空间特征,在使用 CNN 进行学习之前,将一维流量数据转化为二维矩阵进行输入,以便 CNN 将流量数据当作图像进行卷积处理。使用 BiGRU 代替普通的 GRU,以一维文本数据格式进行特征提取,充分获取流量数据的双向时间特征。CNN 和 BiGRU 在本模型中的实现见图 2 中对应部分,限于篇幅,本文不再展开介绍。

注意力机制[19]是对输入信息的加权求和,根据信息的重要程度确定权重。通过注意力机制可以充分挖掘数据的特征信息,提高分类的准确率[20]。本文 CNN 和 BiGRU 中分别添加相应的注意力机制进行特征加权,以提高模型的整体性能。注意力机制的运算过程:使用 tanh 函数对经过神经网络处理的隐藏状态序列 $\boldsymbol{h}_t$ 进行非线性变换,得到 $\boldsymbol{h}_t$ 的隐式表示 $\boldsymbol{u}_t$。对 $\boldsymbol{u}_t$ 进行加权处理,得到注意力权值 α_t。依据注意力权值对序列的隐含向量进行加权,得到最终流量的新特征向量 $\boldsymbol{v}$。

$$\boldsymbol{u}_t = \tanh(\boldsymbol{W}_t \boldsymbol{h}_t + b_t);\tag{6}$$

$$\alpha_t = \frac{\exp(\boldsymbol{u}_t \boldsymbol{u}_w)}{\sum_{k=1} \exp \boldsymbol{u}_k};\tag{7}$$

$$\boldsymbol{v} = \sum_t \alpha_t \boldsymbol{h}_t。\tag{8}$$

式中:$\boldsymbol{W}_t$、$\boldsymbol{u}_w$ 均为权重矩阵;b_t 为偏置量。

在 CNN 和 BiGRU 中分别引入注意力机制赋予特征权重,随后将两神经网络提取的特征信息进行融合,通过全连接网络进行处理,并通过 Softmax 分类器输出分类结果。本文构建的融合神经网络模型结构如图 2 所示。

这样一来,在入侵检测的特征提取中,既考虑了空间层面的特征联系,也考虑了时间层面上特征之间的变化规律。该融合神经网络利用 CNN 提取局部平行特征,同时利用 BiGRU 获取流量数据的双向时间特征,对长距离时间关联依赖特征进行特征提取,从而分析各特征点前、后信息对该特征点的影响。由于 CNN 与 LSTM 为并联处理关系,而非串联处理关系[10],这样就避免了文献[10]方法在最初的 CNN 特征提取时可能会损失部分数据信息时间特性的问题,并且可以得到较高的精度和较低的误报率,与文献[10]结果一致,从而较好地解决了单一神经网络特征学习不全面的问题。

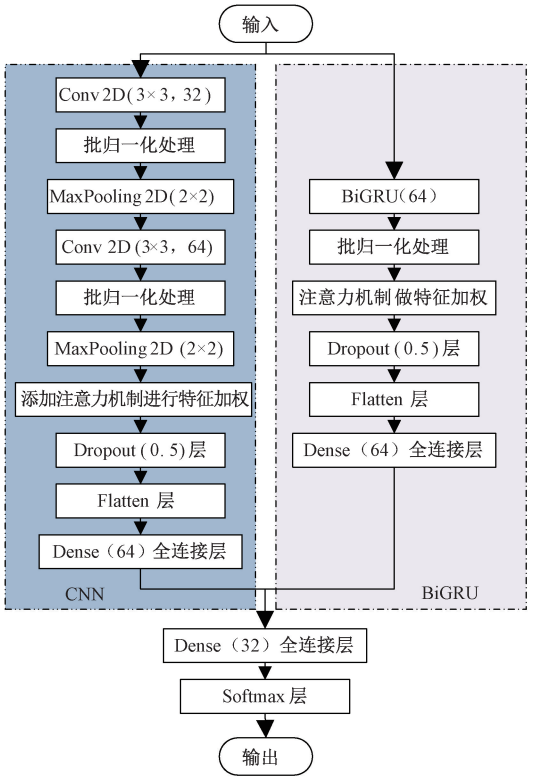

图 2 本文融合神经网络模型

Figure 2 Fusion neural network model in this paper

2 实验部分

2.1 实验环境

实验采用的微型机处理器为 Intel Core i5,内存为 16.0 GB,显卡为 GeForce GTX 1650,操作系统为 Windows 10,编程语言为 Python 3.5.4,框架为 Keras 2.4.3,后端采用 Tensorflow-GPU 2.2.0,实验利用 CUDA(compute unified device architecture)进行 GPU 加速,加快模型训练。

2.2 数据集预处理

2.2.1 数据集的数据清洗

CSE-CIC-IDS2018 入侵检测数据集共有 83 个数据特征字段(又称为标签)和 1 个类别标签,共计 84 个标签。由正常流量(标签为 Benign)和 14 种攻击流量构成,包含 7 种攻击场景。经过数据清洗的数据集流量类别分布如表 1 所示。

2.2.2 特征选择

CSE-CIC-IDS2018 入侵检测数据集中每个数据包括 84 个特征字段[15]。在实际入侵检测系统运用中,数据集中包含冗余特征,使用时会增加计算工作量、降低检测速度、影响入侵检测的泛化性。入侵检测系统应该根据网络流量的行为特征进行分类,不应偏向于 IP 地址等具有特定网络标识的信息,因此预处理过程中删掉了 Flow ID、Src IP、Src Port、Dst IP、Timestamp 这 5 个涉及特定网络标识的特征字段;信息相同的数据字段不会对分类提供有效信息,因此删掉了 Bwd PSH Flags、Fwd URG Flags、Bwd URG Flags、CWE Flag Count、Fwd Byts/b Avg、Fwd Pkts/b Avg、Fwd Blk Rate Avg、Bwd Byts/b Avg、Bwd Pkts/b Avg、Bwd Blk Rate Avg 这 10 个数据值全为 0 的特征字段,数据集剩余 68 个特征字段。依据前述特征重要性评分公式(5),计算得到剩余 68 个特征字段的特征重要性评分。为了构建 8×8 的卷积神经网络数据矩阵,再按照特征重要性评分大小对特征字段进行排序,将特征重要性最低的 Fwd PSH Flags、FIN Flag Cnt、SYN Flag Cnt、Active Mean 这 4 个特征字段删除,数据集剩余 64 个特征字段。

表 1 数据集流量类别分布

Table 1 Dataset flow label distribution

流量类别	数量	占比/%
Benign	13 461 587	85.121 51
DDoS attack-HOIC	668 461	4.226 87
DDoS attacks-LOIC-HTTP	576 191	3.643 42
DoS attacks-Hulk	434 873	2.749 83
Bot	282 310	1.785 13
Infiltration	160 604	1.015 55
SSH-BruteForce	117 322	0.741 86
DoS attacks-GoldenEye	41 455	0.262 13
FTP-BruteForce	39 346	0.248 80
DoS attacks-Slow HTTP Test	19 462	0.123 06
DoS attacks-Slowloris	10 285	0.065 04
DDoS attack-LOIC-UDP	1 730	0.010 94
Brute Force-Web	611	0.003 86
Brute Force-XSS	230	0.001 45
SQL Injection	87	0.000 55

2.3 数据集拆分与平衡

考虑到计算机性能限制,本次实验从正常流量中随机抽取 2×10^6 条流量作为正常流量的总样本,攻击类流量样本不变。经过其他处理的数据集数据按照 8:1:1的比例拆分为训练集、验证集和测试集。

在多分类中,使用随机采样对训练集中部分多数类样本进行下采样操作,同时使用 SMOTE-Tomek 算法对部分少数类样本进行合成和清洗,完成训练集的数据平衡。

2.4 评价指标

为了对模型的性能进行充分评估,实验采用准确率、精确率、召回率以及 $F1$ 值作为模型评价指标。

对于多分类问题,对某种流量类别进行分类评估时,将该流量类别视为正样本,其余流量类别均视为负样本。

2.5　超参数设置

模型超参数的选取会直接影响模型的效果。本文模型采用 Nadam 优化算法代替 Adam 加强对学习率的约束[21],使用 $F1$ 值评价指标作为模型超参数优化指标,得到实验最佳超参数设置:BiGRU 隐藏层节点数为 64;卷积核大小为 3×3;丢弃率为 0.5;批大小为 256;迭代次数为 100;学习率为 0.002。其中,学习率为动态变化的,初始值为 0.002,当参数状态越来越逼近全局最优点时,学习率按一定比例降低。

3　结果分析

3.1　SMOTE-Tomek 数据平衡算法的有效性

为了验证 SMOTE-Tomek 算法对数据平衡的有效性,将经过 SMOTE-Tomek 算法处理和未经过 SMOTE-Tomek 算法处理的数据集在本文模型上进行对比检测,得到数据集平衡前后各类别识别准确率如图 3 所示。

经 SMOTE-Tomek 算法处理的数据集较未经过处理的数据集在少数类的识别精度上提升明显。经 SMOTE-Tomek 算法处理,DoS attacks-Slow HTTP Test 识别准确率从 0 提升至 34.66%,SQL Injection 识别准确率从 0 提升至 100%,DDoS attack-LOIC-UDP、Brute Force-Web 和 Brute Force-XSS 分别提升了 5.22 百分点、6.55 百分点和 35.71 百分点,FTP-BruteForce 识别准确率降低了 10.62 百分点。通过分析混淆矩阵发现,FTP-BruteForce

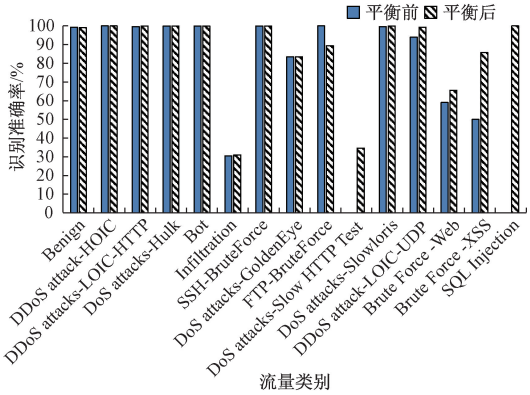

图 3　数据集平衡前后各流量类别识别准确率

Figure 3　Recognition accuracy of each data class before and after data set balance

和 DoS attacks-Slow HTTP Test 极易发生混淆,平衡处理前,系统将 DoS attacks-Slow HTTP Test 全部误判为 FTP-BruteForce,经过平衡处理后,虽然 FTP-BruteForce 识别准确率有所下降,但可以更好地将两类区分。

实验结果证明了 SMOTE-Tomek 数据集平衡算法对于提升模型检测效果的有效性。

3.2　融合神经网络模型性能分析

选取经典的单一深度学习模型与本文模型进行性能比较,对比实验为多分类方式。

实验中,使用 CNN、LSTM 和 GRU 单一模型与本文模型进行对比。所有模型均使用经过 SMOTE-Tomek 算法处理过的训练集对模型进行训练,分别比较这几种模型对各类别流量分类的性能,以精确率、召回率和 $F1$ 值为评价指标。多分类结果如表 2 所示。

表 2　CNN、LSTM、GRU 模型与本文模型的多分类性能对比

Table 2　Comparison of multi classification performance of CNN, LSTM and GRU models with proposed model %

流量类别	CNN 模型			LSTM 模型			GRU 模型			本文模型		
	精确率	召回率	$F1$ 值	精确率	召回率	$F1$ 值	精确率	召回率	$F1$ 值	精确率	召回率	$F1$ 值
Benign	93.99	98.53	96.21	94.02	98.73	96.32	94.11	98.59	96.30	96.52	99.13	97.81
DDoS attack-HOIC	99.98	100.00	99.99	99.94	100.00	99.97	100.00	100.00	100.00	100.00	100.00	100.00
DDoS attacks-LOIC-HTTP	99.54	99.58	99.56	99.80	99.58	99.69	99.96	99.58	99.77	99.93	99.90	99.91
DoS attacks-Hulk	98.41	99.97	99.18	98.45	99.97	99.21	98.45	99.97	99.20	98.46	99.97	99.21
Bot	99.64	99.90	99.77	99.57	99.89	99.73	99.90	99.98	99.94	99.90	99.99	99.95
Infiltration	62.68	21.62	32.15	63.70	22.29	33.02	59.60	23.71	33.92	68.16	31.00	42.61
SSH-BruteForce	99.74	99.97	99.86	99.98	99.97	99.98	99.95	99.96	99.95	99.98	99.97	99.98
DoS attacks-GoldenEye	98.92	83.46	90.54	99.38	83.29	90.63	99.44	83.46	90.76	99.50	83.46	90.78
FTP-BruteForce	66.55	100.00	79.92	66.59	100.00	79.94	66.61	100.00	79.96	73.19	89.38	80.48
DoS attacks-Slow HTTP Test	0	0	0	0	0	0	0	0	0	62.05	34.66	44.47
DoS attacks-Slowloris	93.23	99.81	96.41	97.25	99.91	98.56	97.88	99.81	99.84	98.42	99.91	99.16
DDoS attack-LOIC-UDP	64.78	100.00	78.62	65.83	98.75	79.00	64.78	100.00	78.62	71.12	99.25	82.87
Brute Force-Web	18.75	51.43	27.48	21.43	51.43	30.25	25.28	64.29	36.29	24.10	65.57	35.24
Brute Force-XSS	10.17	70.59	17.78	29.27	70.59	41.38	29.79	82.35	43.75	31.58	85.71	46.15
SQL Injection	1.81	100.00	3.56	3.48	100.00	6.72	14.29	100.00	25.00	15.38	100.00	26.67
总评分	94.81	95.77	94.86	94.92	95.89	94.97	94.86	95.89	95.03	97.02	97.39	97.00

由表2可知,本文模型应用于多分类问题时的总体性能良好,除FTP-BruteForce外,模型各类别分类评分均高于CNN、LSTM和GRU模型。因FTP-BruteForce和DoS attacks-Slow HTTP Test极易发生混淆,故DoS attacks-Slow HTTP Test召回率得到明显提升后,FTP-BruteForce召回率有所下降,但本文模型中FTP-BruteForce分类取得了更高的精确率和$F1$值。Infiltration类分类性能较其他几种神经网络模型提升明显,召回率较CNN模型提升了9.38百分点。

本文模型总的分类精确率、召回率以及$F1$值均高于其他几种单一神经网络模型,其中各攻击流量类别的总评精确率比LSTM模型提升了2.10百分点;总评召回率比LSTM模型提升了1.50百分点;总评$F1$值比GRU模型提升了1.97百分点。

综上实验结果,之所以取得较好的性能,是由于本文模型利用CNN对局部平行特征进行特征提取,避免特征丢失,能够提取深层次信息;同时,利用BiGRU对长距离时间关联依赖特征进行特征提取,充分考虑到数据集平衡前后信息对当前的影响,有效降低检测误报率,较好地解决了单一神经网络特征学习不全面的问题。

4 结论

针对当前深度学习入侵检测中存在的数据集不平衡及特征信息学习不全面等问题,提出了一种基于CNN与BiGRU融合神经网络入侵检测模型。通过SMOTE-Tomek算法完成对数据集的平衡处理,使用基于平均不纯度减少的特征重要性算法实现特征选择,将CNN和BiGRU模型进行特征融合并引入注意力机制进行特征提取,从而提高模型的总体检测性能。实验结果表明,SMOTE-Tomek算法能够有效对数据进行平衡,加强了各检测模型对少数类样本的学习。同时,通过与经典单一深度学习模型的对比实验表明,该模型能够在更全面地提取数据特征的同时,还能在各项评价指标中获得更好的效果。

参考文献:

[1] FERNÁNDEZ G C, XU S H. A case study on using deep learning for network intrusion detection[C]// MILCOM 2019 IEEE Military Communications Conference (MILCOM). Piscataway: IEEE, 2019: 1-6.

[2] 张蕾,崔勇,刘静,等. 机器学习在网络空间安全研究中的应用[J]. 计算机学报,2018,41(9):1943-1975.

[3] 张玉清,董颖,柳彩云,等. 深度学习应用于网络空间安全的现状、趋势与展望[J]. 计算机研究与发展,2018,55(6):1117-1142.

[4] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4): 541-551.

[5] ROY S S, MALLIK A, GULATI R, et al. A deep learning based artificial neural network approach for intrusion detection[J]. Mathematics and computing, 2017, 655: 44-53.

[6] WANG W, ZHU M, ZENG X W, et al. Malware traffic classification using convolutional neural network for representation learning[C]//2017 International Conference on Information Networking (ICOIN). Piscataway: IEEE, 2017: 712-717.

[7] NASEER S, SALEEM Y, KHALID S, et al. Enhanced network anomaly detection based on deep neural networks[J]. IEEE access, 2018, 6: 48231-48246.

[8] KIM J, KIM J, LE T T H, et al. Long short term memory recurrent neural network classifier for intrusion detection[C]//2016 International Conference on Platform Technology and Service (PlatCon). Piscataway: IEEE, 2016: 1-5.

[9] PUTCHALA M K. Deep learning approach for intrusion detection system (IDS) in the internet of things (IoT) network using gated recurrent neural networks (GRU)[D]. Dayton: Wright State University, 2017.

[10] 王伟. 基于深度学习的网络流量分类及异常检测方法研究[D]. 合肥:中国科学技术大学,2018.

[11] YIN C L, ZHU Y F, FEI J L, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. IEEE access, 2017, 5: 21954-21961.

[12] 张勇东,陈思洋,彭雨荷,等. 基于深度学习的网络入侵检测研究综述[J]. 广州大学学报(自然科学版),2019,18(3):17-26.

[13] 陈洁,邵志清,张欢欢,等. 基于并行混合神经网络模型的短文本情感分析[J]. 计算机应用,2019,39(8):2192-2197.

[14] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE transactions on signal processing, 1997, 45(11): 2673-2681.

[15] SHARAFALDIN I, LASHKARI A H, 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. Funchal: ICISSP, 2018: 108-116.

[16] PANIGRAHI R, BORAH S. A detailed analysis of CI-

CIDS2017 dataset for designing Intrusion Detection Systems [J]. International journal of engineering & technology, 2018, 7: 479-482.

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

[18] BATISTA G E A P A, PRATI R C, MONARD M C. A study of the behavior of several methods for balancing machine learning training data[J]. ACM SIGKDD explorations newsletter, 2004, 6(1):20-29.

[19] 李勇,金庆雨,张青川.融合位置注意力机制和改进 BLSTM 的食品评论情感分析[J].郑州大学学报(工学版),2020,41(1):58-62.

[20] 朱张莉,饶元,吴渊,等.注意力机制在深度学习中的研究进展[J].中文信息学报,2019,33(6):1-11.

[21] DOGO E M, AFOLABI O J, NWULU N I, et al. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks [C]// 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS). Piscataway: IEEE, 2018:92-99.

Intrusion Detection Model Based on CNN and BiGRU Fused Neural Network

ZHANG Anlin[1], ZHANG Qikun[2], HUANG Daoying[2], LIU Jianghao[2], LI Jianchun[2], CHEN Xiaowen[2]

(1. Engineering Training Center, Zhengzhou University of Light Industry, Zhengzhou 450001, China; 2. College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, China)

Abstract: Aiming at the problems of unbalanced data types and incomplete feature learning in deep learning intrusion detection, a neural network intrusion detection model based on the fusion of convolutional neural networks (CNN) and bidirectional gated recurrent unit (BiGRU) was proposed. The SMOTE-Tomek algorithm was used to balance the data set, the feature importance algorithm based on mean decrease impurity was used to realize feature selection; the CNN and BiGRU models used for feature fusion and attention mechanism was introduced for feature extraction, so as to improve the overall detection performance of the model. The intrusion detection data set CSE-CIC-IDS2018 was used for multi classification experiments, the model was compared with the classical single deep learning models. The experimental results showed that, firstly, in terms of data set balance, after being processed by SMOTE-Tomek algorithm, the recognition accuracy of DoS attacks-Slow HTTP Test class was improved from 0 to 34.66%, that of SQL Injection class was improved from 0 to 100%, and DDoS attack-LOIC-UDP, Brute Force-Web and Brute Force-XSS classes were improved by 5.22 percentage points, 6.55 percentage points and 35.71 percentage points respectively. It was proved that the balanced data set improved the recognition accuracy of a few classes significantly compared with the unprocessed data set. Secondly, in terms of the overall detection performance of the model, in the comparison of multi classification experiments, the overall classification accuracy, recall and $F1$ value of the model in this study were higher than those of several other single neural network models. The overall evaluation accuracy of each attack traffic category was about 2.10 percentage points higher than that of the highest LSTM model. The recall rate of the overall evaluation was about 1.50 percentage points higher than that of the highest LSTM model. Compared with the highest GRU model, the overall $F1$ value increased by about 1.97 percentage points. It was proved that the model had better detection effect.

Keywords: intrusion detection; convolutional neural networks; bidirectional gated recurrent unit; synthetic minority over-sampling technique algorithm; Tomek Links algorithm