小型微型计算机系统
Journal of Chinese Computer Systems
DOI:10.20009/j.cnki.21-1106/TP.2023-0449
2025年2月 第2期
Vol.46 No.2 2025

融合改进TCN与DRSN的IoT入侵检测模型

赵 建,姜 伟

(哈尔滨师范大学 计算机科学与信息工程学院,哈尔滨 150025)
E-mail:zzjj123666@163.com

摘 要: 入侵检测系统已逐步成为物联网安全的重要防护手段.然而,现有物联网入侵检测模型的样本数据存在类别不平衡、特征提取不足等问题,这导致了对于小类别攻击的低识别率与较低的精确率.因此,本文提出了一种融合改进时域卷积网络与深度残差收缩网络的物联网入侵检测模型.首先,利用扩张因果卷积与一维卷积充分提取数据的时空特征,形成深层层次的网络结构;然后引入自我注意的软门槛,能够无需专家经验自动地设置门槛,消除冗余特征;最后,使用焦点损失函数来增强对少数类的识别率.实验在TON-IoT数据集上的总体准确率和F1值分别高达99.88%和99.64%,其中小样本类的F1值为100%.实验结果表明,与其他模型相比,所提模型显著提高了对于不平衡入侵数据的检测能力.

关 键 词: 物联网;入侵检测;时域卷积网络;深度残差收缩网络;样本不平衡;焦点损失函数
中图分类号: TP393 **文献标识码:** A **文 章 编 号:** 1000-1220(2025)02-0474-08

IoT Intrusion Detection Model Integrating Improved TCN and DRSN

ZHAO Jian, JIANG Wei
(School of Computer Science and Information Engineering, Harbin Normal University, Harbin 150025, China)

Abstract: Intrusion detection systems have gradually become an important means of protection for IoT security. However, the sample data of existing IoT intrusion detection models have problems such as category imbalance and insufficient feature extraction, which results in low recognition rates and relatively low recognition rates for small category attacks. Therefore, this paper proposes an IoT intrusion detection model that integrates an improved temporal convolutional network and a deep residual shrinkage network. First, dilated causal convolution and one-dimensional convolution are used to fully extract the spatiotemporal features of the data, forming a deep-level network structure; then introduce a soft threshold of self-attention, which can automatically set the threshold without expert experience and eliminate redundant features; finally, use the focus loss function to enhance the recognition rate of minority classes. The experiment is in TON-IoT The overall accuracy and F1 value on the data set are as high as 99.88% and 99.64% respectively, among which the F1 value of the small sample class is 100%. Experimental results show that compared with other models, the proposed model significantly improves the detection ability of imbalanced intrusion data.

Keywords: internet of things; intrusion detection; TCN; DRSN; sample imbalance; focal loss

0 引 言

物联网(IoT)[1]的概念已经提出多年,作为互联网的延伸与拓展其一直以来都饱受关注.随着5G网络的快速发展,无线传感设备数量的激增,物联网的规模也将越来越大.根据最近的分析报告,全球使用的物联网设备超过138亿台,预计到2025年这一数字将增至309亿台[2].然而,因物联网系统具有分布式、对象数量大和开放性等特点,其也成为网络攻击者的理想目标[3].所以,保护物联网系统的安全变得尤为重要.通过密码认证、安全网络拓扑构建等手段,防御者为物联网设备构建了初始的安全防护层.然而,尽管如此,某些攻击者仍然能够通过数据分析等手段对物联网设备发起恶意攻击.为此,入侵检测(IDS)被视为第二道防线,在确保物联网设备安全方面具有关键作用[4].入侵检测技术可以察觉并鉴别物联网设备中潜在的攻击行为,从而能够及时做出正确响应,采取有效的防护措施.

机器学习(ML)和深度学习(DL)技术被广泛用于开发基于分类器的入侵检测系统,可以通过流量数据分析区分正常的网络流量和不同的网络攻击[5].因资源受限和复杂性等因素,传统的入侵检测技术在物联网方面的安全性表现有限.因此,对于物联网设备来说,需要一种更加安全可靠且具有高度弹性的入侵检测系统.深度学习技术具备迅速分析大量数据的能力,并能在发现恶意软件或安全漏洞时自动调整安全系统,同时还能以较低的计算能力实现这些功能[6].在深度神经网络中,随着网络层数的增加,梯度在反向传播时会逐渐变小,导致底层的权重几乎不会更新,导致梯度消失问题.另一方面,网络层数过多还可能导致梯度变得异常大,引发梯度爆炸问题.这些问题都使得训练非常深的网络变得困难.残差

收稿日期:2023-12-08 收修改稿日期:2024-01-17 基金项目:国家自然科学基金面上项目(61872104)资助;黑龙江省教育厅科研项目(11551124)资助;哈尔滨师范大学研究生创新工程项目(HSDSSCX2023-10)资助. 作者简介:赵 建,男,1999年生,硕士研究生,研究方向为网络信息安全;姜 伟(通信作者),男,1974年生,博士研究生,副教授,研究方向为网络信息安全、隐私保护、自律计算.

收缩网络(DRSN)[7]就很好地解决了上面的问题,同时还加强了神经网络在含噪声的数据中提取有用特征的能力. 时域卷积网络(TCN)[8]由多个残差模块组成,每个残差块包含两层的膨胀因果卷积和非线性映射,在每层中还加入了权重归一化(WeightNorm)和 Dropout 层来正则化网络. 本文将 TCN 与 DRSN 相结合,使网络模型能够提取数据样本的时空维度特征,自适应地剔除冗余信息,进而提高模型分类的准确度与鲁棒性.

许多不平衡数据集已被用于研究入侵检测. 数据集中的多数类是具有大量样本的类,而少数类是具有少量样本的类. 大多数研究人员要么忽略了这个问题,要么试图通过过采样(即通过随机复制少数类中的样本)或欠采样(即通过随机删除多数类样本)来平衡数据. 然而,过采样可能会导致过拟合,因为新样本是原始样本的精确复制品. 另一方面,欠采样可能会导致数据集过于简单,因为数据样本太少而无法构建有效的模型,从而导致欠拟合问题. 无论是过采样还是欠采样,他们都可能会一定程度上破坏原始数据集. 此外,传统的损失函数(例如交叉熵损失)在训练深度学习模型时执行平均梯度更新,因此没有适当关注少数类实例. 为了解决这个问题,本文使用改进的焦点损失函数,在不改变原始数据集分布的情况下,提高对小样本类的检测率.

提出的物联网入侵检测模型,由于融合了 TCN 与 DRSN,因此将所提模型命名为 TDF-IDS. 本文的创新点如下:

1)在入侵检测领域,首次以并联的方式应用融合一维卷积与扩张因果卷积的残差收缩网络,使模型能够提取数据样本的时空特征,自适应剔除冗余信息.

2)提出了一种新的融合改进 TCN 与 DRSN 的 IoT 入侵检测模型,并在物联网数据集 TON-IoT 上进行了测试,实验表明,本文模型与基线模型相比准确率高达 99.88%.

3)引入焦点损失函数来替代交叉熵损失函数以解决数据样本分布严重不平衡的问题. 对比实验表明,改进后的模型对小样本类别的检测率有很大的提高.

1　相关工作

深度学习被广泛的应用于入侵检测领域并取得了良好的效果. Halbouni 等人[9]利用 CNN 提取空间特征的能力和 LSTM 提取时间特征的能力来创建混合入侵检测系统模型,在多个数据集上基于二元与多分类都显示出高检测率、高精度和较低的 FAR. 印传龙等人[10]提出一种了一种使用 RNN 的入侵检测模型,该模型在二元与多分类下性能均优于传统的机器学习分类方法. 胡志全等人[11]提出一种混合自适应合成采样(ADASYN)与基于 CNN 改进分割卷积模块(SPC-CNN)的 AS-CNN 模型入侵检测模型,通过平衡样本分布、增加特征的多样性,消除通道间信息冗余对模型训练的影响,在 NSL-KDD 数据集上准确率、误识率和检测率都有不错的效果. 李艳苗等人[12]提出了一种使用多卷积神经网络(multi-CNN)的融合模型,根据数据特征相关性,将其分为 4 部分,然后将一维特征数据转换为灰度图,经过 CNN 融合模型得出 4 种结果中最好的结果,在工业物联网数据集 NSL-KDD 上获得了优异的分类准确率. 李艳等人[13]提出一种基于 CBAM-ResNet 和 Self-attention 的僵尸网络检测模型,利用 CBAM-ResNet 学习网络流量的空间特征,Self-attention 学习网络流量的时间特征,该模型提高了检测性能和泛化能力,在未知僵尸网络流量的检测中也起到了良好的作用. Bhavsar 等人[14]开发了一种基于称为皮尔逊相关系数-卷积神经网络(PCC-CNN)的 IoT 入侵检测模型. 该模型结合了从基于线性的提取和卷积神经网络获得的重要特征,在 NSL-KDD、CIC-IDS2017 和 IOTID20 3 个数据集上二元和多类入侵检测分类器的误报率分别为 0.02、0.02 和 0.00.

上述入侵检测模型在异常数据的检测上面均表现出了不错的性能,但它们所使用的数据集均存在不平衡的问题,这一点往往被研究人员所忽略. 为了解决现有的入侵检测数据样本不平衡的问题,燕昺昊[15]等人提出了一种利用深度递归神经网络(DRNN)和区域自适应复合过采样算法(RA-SMOTE)的组合 NIDS 模型,提高了模型描述数据的能力和检测性能. 然而,在一些长期依赖问题中,传统 RNN 由于其结构特点,在训练过程中可能会遇到梯度消失、梯度爆炸等问题. 周小康等人[16]提出了一种分布偏差感知协作 GAN(DB-CGAN)模型,用于工业物联网中的不平衡深度学习,相比基线模型具有较高的分类精度. Khan 等人[17]提出了一种成本敏感(CoSen)深度神经网络,它可以自动学习多数类和少数类的鲁棒特征表示,所提出的方法显著优于基线算法. 尽管上述包括数据生成方法在内的入侵检测方法取得了令人满意的性能,但它们仍然存在检测率较低、FPR 较高以及低频、少数和未知攻击类别的检测性能较低的问题.

上述研究主要研究如何有效更丰富的提取网络数据的特征,同时解决数据集中类的不平衡的问题,以提高对异常数据的检测率. 本文引入融合时间卷积网络与残差收缩网络,学习网络数据集的时空特征的同时自适应的剔除冗余信息以提高入侵检测模型的泛化能力;使用焦点损失函数增加小数量样本类别与被错误分类样本的权重,从而解决样本数量不平衡的问题.

2　相关理论

2.1　时域卷积

2.1.1　因果卷积

TCN 有两个主要特性:每一层都具有相同的输入输出长度,而且当前的输出仅与当前和过去时间输入有关. 为了满足上述两个特性,TCN 使用一维全卷积网络,有些层会采用零填充,以保证所有的卷积层具有相同的长度;此外,TCN 的卷积是因果的,即仅对当前时刻及以前的输入进行卷积,而不使用未来的数据. 上述表述可定义如式(1),TCN_i 其中表示时域因果卷积网络的第层,表示一维全卷积网络层,$Ccon$ 表示因果卷积层:

$$TCN_i = 1D\ FCN + Ccon \tag{1}$$

TCN 使用的因果卷积(Causal convolution)如图 1(a)所示,其卷积核为 2,但这个结构有一个缺点:要实现较长的有效历史时间序列大小,需要使用很多的卷积层,然而这会导致计算量过大、梯度消失弥散等问题. 所以,TCN 使用扩张卷积

来解决这一问题.

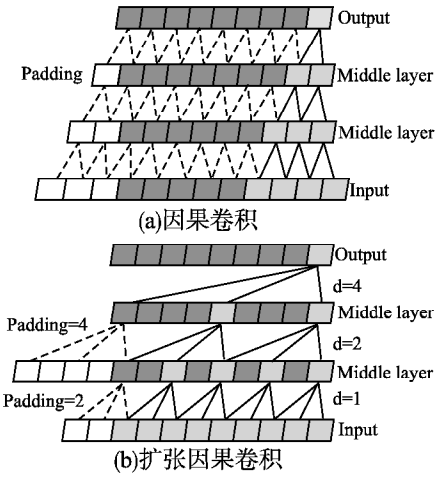

图 1 因果卷积与扩张因果卷积

Fig. 1 Causal convolution and dilated causal convolution

2.1.2 扩张卷积

TCN 通过扩张卷积来获得更大的感受野,一般通过增加扩张因子和卷积核大小来扩大感受野. 如图 1(b)所示,扩张卷积的卷积核大小为 3,扩张因子 $d=1,2,4$,以 2 的倍数增长,并保持输入输出序列大小不变,从而在不改变计算量的情况下使网络观测到了更长的时间序列.

2.1.3 残差块

残差块(Residual Block)主要用来解决神经网络在训练过程中梯度爆炸与弥散问题,能有效的加快网络的训练速度的 训练速度. 通过引入“跳跃”连接,即使某一层的输入与输

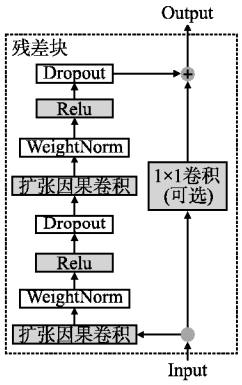

图 2 时域卷积网络的残差块

Fig. 2 Residual block of temporal convolutional network

出进行线性叠加,从而能够将原始信息传入较深的网络层中. TCN 的残差块包含两个扩张因果卷积层,每层还加入了权重归一化与随机失活(Dropout)来优化网络,如图 2 所示.

2.2 深度残差收缩网络(DRSN)

2.2.1 软阈值化

软阈值化(Soft Thresholding)是用于信号降噪与压缩的一种技术,其主要是对信号取绝对值并与一个预设的阈值进行比较,当绝对值低于这个阈值时将其置零,反之则将被设为原始值减去阈值. 软阈值公式为式(2),其中$[-\eta,\eta]$为阈值区间,x 为输入特征,软阈值函数(a)及其导数(b)如图 3 所示.

$$y=\begin{cases}x-\eta, & x>\eta\\ 0, & -\eta\leqslant x\leqslant\eta\\ x+\eta, & x<-\eta\end{cases}\tag{2}$$

对于这个预设的阈值往往需要专家知识的介入,一般很难选择一个合适的取值,而 DRSN 通过一种注意力机制很好的解决了这个问题,即压缩与激励网络(SENet),它能自适应的学习一组权重,以达到去噪的效果.

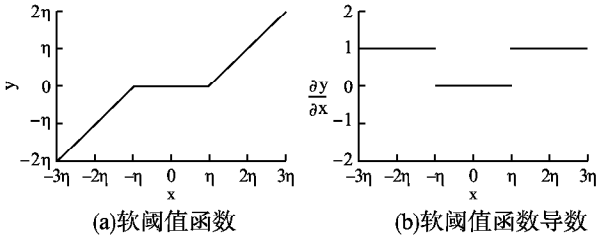

图 3 软阈值函数及其导数

Fig. 3 Soft threshold function and its derivative

2.2.2 注意力机制

SENet 是一种非常典型的注意力机制方法,它通过扫描全局信息并从中发现有用的局部信息,自适应的得到一组权重,使用这组权重对特征图的每个通道进行加权操作,相当于多关注有用的特征通道,而不重要的通道则减少关注度. SENet 由全局池化层、全连接层、ReLU 激活函数层与 Sigmoid 排列组合而成.

2.2.3 网络结构

深度残差收缩网络一共包含两个版本,分别为通道阈值共享的残差收缩网络(DRSN-CS)和每个通道独享一个阈值的残差收缩网络(DRSN-CW). 本文选择 DRSN-CW 版本,因为在实际应用中理论上不可能每个特征通道都有相同的冗余信息,如果每个通道都采用相同的阈值,可能会导致一些通道的冗余特征不能被剔除,最终很有可能会一定程度上影响网络的性能. DRSN-CW 整体结构如图 4 所示,包含输入层、卷积层(Conv)、多个串接的通道级阈值残差收缩块(Residual Shrinkage Block-CW,RSB-CW)、批量归一化(BN)、激活函数(Relu)、全局平均池化层(GAP)和全连接层(FC),其中 C、W 分别表示特征图的通道与宽度.

2.3 焦点损失函数

焦点损失函数最初被用于解决图像领域中样本不平衡问题,本文将其引入以解决物联网入侵检测中数据样本不平衡造成少样本检测率低的问题. 焦点损失函数是基于交叉熵损失函数的改进,与传统的交叉熵函数相比,焦点损失函数能够动态的缩放梯度更新来优化模型,从而降低易分类样本的权重,使模型更加关注难分类的样本示例,其公式见式(3):

$$FL(p_t)=-\alpha_t(1-p_t)^{\gamma}\log(p_t)\tag{3}$$

上式中,γ 相当于一个调节因子,能平滑的调节难易分类样本的关注度,α_t 则用来平衡样本类别不均衡的问题,给不同样本类别分配不同的权重. γ 只能解决样本难易分类的问题,但不能平衡不同样本类别的重要性,所以添加 α_t 就变得很有必要. 本文将 γ 设为 2,α_t 公式见式(4):

$$\alpha_t=1-\frac{n_t}{sum_n}\tag{4}$$

上式中,n_t 代表 t 类型样本的数量,sum_n 代表用于训练的样本数量 p_t 定义式见式(5):

$$p_t = \begin{cases} p & ,y=1 \\ 1-p & ,y=0 \end{cases} \tag{5}$$

上式中,y 表示真实的标签(1 表示属于该类别,0 表示不属于该类别),p 表示模型预测样本属于该类别的概率,p_t 代表分类的置信度. 如果模型对某个类别的分类置信度 p 越高,那么 p_t 也越高,反之亦然.

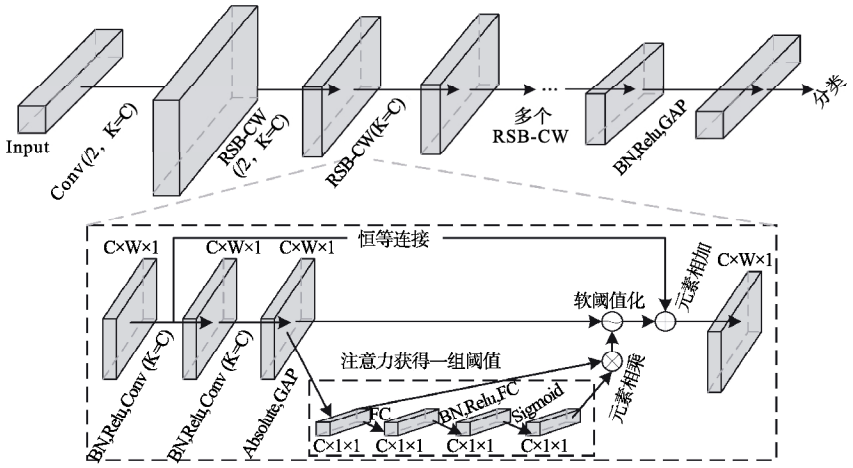

图 4　通道级残差收缩网络(DRSN-CW)

Fig. 4　Channel-level residual shrinkage network(DRSN-CW)

3　TDF-IDS 物联网入侵检测模型设计

TDF-IDS 模型的结构如图 6 所示,使用一维卷积对输入数据进行一次特征提取,其中设置卷积核的个数(K)等于通道数(C),所提特征图经过批量归一化与激活函数后再通过 3 个时域残差收缩块(TRSB-WC),将在后面对其进行介绍. 随后再进行一次 BN 与 Relu 操作,以加快模型的训练速度,然后再通过全局平局池化层对高维度的特征图进行降维,减少模型训练的参数,以增加模型的准确性与稳定性,防止发生过拟合形象;接着通过 3 层全连接层将高维特征映射到低维. 本文模型选择焦点损失函数来计算模型的损失,从难易分类与类别不平衡两个维度减少数据样本不平衡对模型预测结果的影响,增强模型的鲁棒性;最后采用 softmax 函数对处理后的样本进行分类.

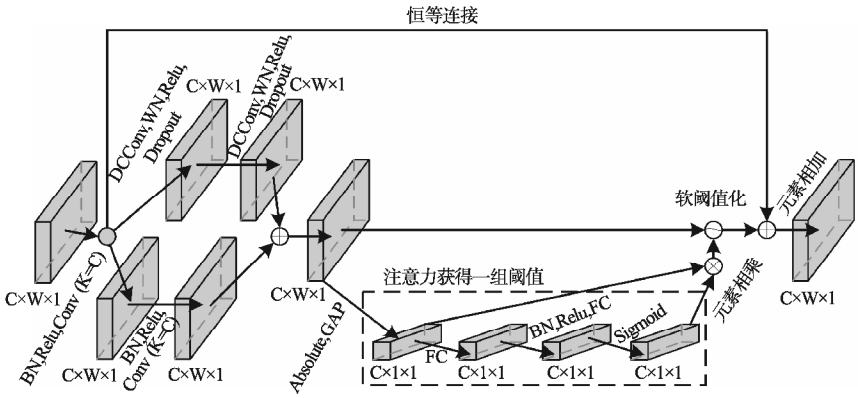

图 5　时域残差收缩块

Fig. 5　Temporal residual shrinkage block

时域残差收缩块的结构如图 5 所示,将两层扩张膨胀卷积以并联的方式与两层一维卷积相连接,然后将各自提取到的特征进行相加,避免简单与直接的串接可能造成一些基本特征丢失的问题,使模型从时间与空间两个维度上提取更加丰富的特征. 这里对一维卷积层采用 padding 零填充的方式,保证卷积先后特征的宽度 W 不会改变,以匹配扩张因果卷积的输出维度,使得能够进行简单的特征融合. 然后通过注意力网络自适应的获得一组阈值,去除融合特征中的冗余信息,最后再进行一次恒等连接获得一组新的特征图. 扩张因果卷积的卷积核大小设为 $3*1$,步长 stride 设为 1,扩张因子 $d=2^{i-1}$(i 为残差块的层数),使用权重归一化与 Relu 激活函数加快模型的训练与收敛速度,在每层中加入 dropout 层防止过拟合. 一维卷积的卷积核

大小设为 3 * 1,每块步长分别为,每层卷积操作都使用 BN 与 Relu 激活函数,以提高模型的稳定性与收敛速度.

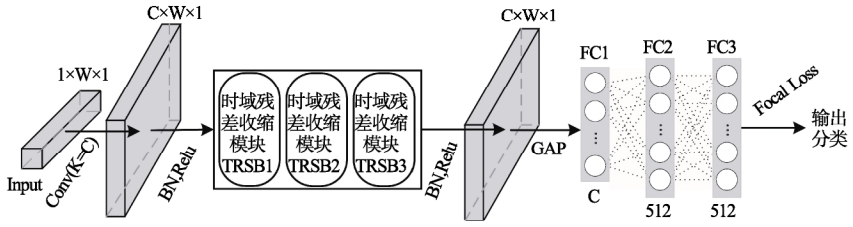

图 6 TDF-IDS 总体结构

Fig. 6 TDF-IDS overall structure

4 实验与结果分析

4.1 实验环境

该实验是使用 Python 语言设计实现的,表 1 描述了模型的实验环境配置.

表 1 实验环境配置

Table 1 Experimental environment configuration

环境	参 数
操作系统	Windows10
CPU	Intel i5-8300H CPU @ 2.30GHz 2.30GHz
内存	DDR4 16GB
Anaconda3	2022.10
Python	3.7.16
Pytorch	1.7.1

4.2 数据来源与预处理

4.2.1 数据来源

本研究使用的 TON-IoT 数据集[18]是由新南威尔士大学堪培拉物联网实验室与澳大利亚国防军学院网络靶场共同收集的. 它是新一代工业 4.0 物联网数据集,其中包含从整个 IIoT 系统收集的各种数据源,包括来自连接设备的遥测数据、Linux 与 Windows 操作系统记录以及 IIoT 系统的网络流量,用以评估基于机器学习与深度学习算法的不同网络安全应用的性能与保真率.

4.2.2 数据预处理

本文所使用的 TON-IoT 数据集共有 8 种标签类型分别为“dos”、“ddos”、“injection”、“normal”、“xss”、“password”、“scanning”、“mitm”,将数据集按照 8:2 的比例以随机打乱的方式划分为训练集与测试集. 训练集与测试集的类别与数量分布如表 2 所示,可以发现数据集类别严重不平衡.

原始数据集中包含许多空缺值,这些空特征并不能参与模型的训练,本文选择将包含空缺值的行进行删除;删除特征维度中全为零的列,如“Processor_pct_C3_Time”、“rocessor_pct_C2_Tim”、“Processor_C2_ransitions_sec”、“Processor_C3_ransitions_sec”等;将上述 8 种类别标签采用标签编码的方式转化为数字特征,以便参与模型的运算.

原始数据集进行上述处理后某些特征的值可能远大于其他特征的值,在模型训练的过程中会凸显数值较高的指标的作用,而一定程度上削弱数值较低的指标的作用,从而影响模型训练的结果. 因此为了保证数据的可靠性与可用性,有必要

表 2 训练集与测试集的类别与数量分布

Table 2 Category and number distribution of training set and testing set

样本类型	训练集	Number	测试集	Number
ddos	3642		897	
Password	2879		715	
xss	987		253	
Injection	483	8749	123	2186
dos	394		111	
scanning	352		82	
mitm	8		5	
normal	7791	7791	1949	2186
总样本数	16536	16536	4135	4135

对数据进行标准化,使数据的所有特征都在同一范围内. 本文使用 Z-score 归一化,使数据集特征落入[-1,1]范围内,如式(6)所示:

$$z = \frac{x - \mu}{\sigma} \tag{6}$$

其中,x 表示原始数据集,μ 表示数据集的均值,σ 表示数据集的标准差.

4.3 评价指标

本文采用以下 4 个指标来评估模型的性能表现,分别是准确率(ACC)、精确率(Pre)、召回率(Recall)、F1 分数(F1)和 AUC(ROC 曲线下面积). 这些指标一般在样本类别不平衡时使用,因为准确率在这种情况下常常会产生误导. 比如在进行二分类时,如果正类别的占有率很高,此时如果模型只预测负类别也可以获得较高的准确率,然而它的精确率与召回率可能会很低. AUC 对于不平衡的正负例分布更具敏感性,一般不会受到数据不平衡的干扰,但当样本类别严重不平衡时,模型可能很难正确分类少数类别的样本. 同时它对不同阈值下的分类器性能进行了综合评估,提供了一个清晰的性能度量,无论正负例分布如何,都具有一定的稳健性. 因此,这些指标能够更全面的评估模型的性能. 这些指标的描述如下:

1)准确率(ACC)表示正确分类的样本数量与总样本数量之间的比率,如式(7)所示:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

2)精确率(Pre)表示正确分类为正类别的样本数量与所有被分类为正类别的样本数量之间的比率,如式(8)所示:

$$Pre = \frac{TP}{TP + FP} \tag{8}$$

3)召回率(*Recall*)表示正确分类为正类别的样本数量与所有实际正类别的样本数量之间的比率,如式(9)所示:

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

4)F1 值(*F1*)是精确度与召回率的调和平均值,用于综合考虑模型的精度与召回率,如式(10)所示:

$$F1 = \frac{2 \times Pre \times Recall}{Pre + Recall} \tag{10}$$

5)AUC 表示为 ROC 曲线下面积,这里用 *Area* 表示面积大小,*Area* 值越大,则模型的性能越好($Area \in [0,1]$),如式(11)所示:

$$Area = \int_0^1 \frac{TP}{TP + FN} d \frac{FP}{TN + FP} \tag{11}$$

上式中,*TP* 代表模型正确预测为正类别的样本数量;*TN* 代表模型正确预测为负类别的样本数量;*FP* 代表模型错误地将负类别样本预测为正类别的样本数量;*FN* 代表模型错误地将正类别样本预测为负类别的样本数量.

4.4　实验参数设置

模型超参数设置是深度学习中至关重要的一部分,它直接影响了模型的性能、训练速度和资源消耗.在设置模型参数时,比较关键因素有:迭代次数(Epoch)、学习率(Learning Rate)、批量大小(Batch Size)、优化器(Optimizer)等.

迭代次数能够影响训练的时长和性能,本文为了使所有参与对比实验的模型都能够收敛,对每个模型都进行 50 次迭代训练.

学习率决定了每次参数更新的步长.学习率太小会导致收敛缓慢,学习率太大可能会导致不稳定的训练过程.在控制其他参数不变的情况下,设置所有模型在前 30 个 Epoch 的学习率为 0.001,后 20 个 Epoch 的学习率为 0.0001,以使模型快速的到达最优解附近,然后再进行微调.

较大的批量可以加速训练,但可能需要更多的内存.考虑到模型运行环境的内存,本文统一设置 Batch Size 为 64.

本文使用了具有自适应学习率机制和动量性质的 Adam 优化函数,它在训练神经网络时能够更快地收敛到局部最小值,并且相对容易配置.

表 3　TRSB-WC 参数配置

Table 3　TRSB-WC parameter configuration

Input channel	d	Output channel	Stride		Kernel size	
			1Dconv	DCconv	Dconv	DCconv
8	1	8	2	1	3	5
8	2	8	1	1	3	5
8	4	8	2	1	3	5

本文所提出模型的 TRSB-WC 详细参数如表 3 所示.其中 d 为扩张因果卷积的扩张因子;1Dconv 为一维卷积网络;DCconv 为扩张因果卷积.

4.5　实验结果与分析

4.5.1　消融实验分析

为了验证所改进模型的有效性,设计了 4 组消融实验,它们分别是 TCN、DRSN、DRSN-FL、TDRSN.对照试验均训练 50 个 Epoch,每组选择 50 次训练中的最优结果进行比较.基准模型与本文模型的比对结果如表 4 所示.

表 4　一些基准模型与本文模型的性能比较(%)

Table 4　Performance comparison of some benchmark models and the model in this article(%)

模型	准确率	精确率	召回率	F1 值	AUC
TCN	94.75	74.14	70.07	71.65	72.17
DRSN	99.47	86.17	86.27	86.21	94.65
DRSN-FL	99.18	89.51	99.34	91.61	97.05
TDRSN	99.83	99.46	89.63	91.21	96.80
TDF-IDS	**99.88**	**99.66**	**99.63**	**99.64**	**98.65**

1)DRSN-FL 与 DRSN 的对比分析

通过观察表 4 可以发现,DRSN-FL 与 DRSN 相比,精确率提升了 3.34%,召回率提升了 13.07%,F1 值提升了5.4%,AUC 提高了 2.4%.在准确率基本相同的情况下,引入 FL 的 DRSN 其他各项指标均有所提高.其原因在于焦点损失函数通过引入 α 与 γ 两个调节因子提高了模型对样本小类别与难分类类别的检测率;α 通过平衡多类别与小类别的权重,使模型更加关注样本少数类,γ 则用于调整易分类样本的权重,γ 值越大,易分类样本的权重越小,这使得模型更集中在难分类的样本上.同时,焦点损失还减少了模型对噪声的敏感性,由于它更关注难分类的样本,使模型对噪声数据的影响较小,更容易适应嘈杂的训练数据.

2)TDRSN 分别与 TCN 和 DRSN 的对比分析

表 4 中的 TDRSN 与 TCN、DRSN 相比,准确率分别提升了 5.08%、0.36%,精确率分别提升了 25.32%、13.29%,召回率分别提升了 19.56%、3.36%,F1 值分别提升了 19.56%、5%,AUC 分别提高了 24.63%、2.15%.各项指标的提升原因在于融合了 TCN 与 DRSN 的模型能够提取到更加丰富的数据特征.TCN 通过具有大的感受野(Receptive Field)和扩展的滑动窗口等机制,可以更好地捕获长期依赖关系,以提取数据的时序特征,使其在处理长序列时更加有效;DRSN 通过多个卷积层使模型能够提取到数据的空间特征,同时引入自我注意力机制的软门槛,使模型在没有专家知识的情况下自动设置门槛,消除冗余特征,减少不重要特征的影响,两者结合能够一定程度上提升模型分类性能.

3)TDF-IDS 与 TDRSN 对比分析

观察表 4 发现,TDF-IDS 与 TDRSN 相比,准确率提升了 0.05%,精确率提升了 0.2%,召回率提升了 10%,F1 值提升了 8.43%,AUC 提高了 1.85%.再一次证明了引入焦点损失函数对少数类样本检测的有效性,且各项检测指标都获得了不错的效果,由此可以看出本文所提模型的有效性.

4)模型训练与测试准确率曲线分析

图 7 显示了 TCN、DRSN、DRSN-FL、TDRSN 与 TDF-IDS 模型的训练准确率随迭代训练次数变化的曲线.可以看出,所有模型在 40 个 Epoch 后基本趋于一致了,其中本文所提出的模型大概在 25 个 Epoch 后准确率就基本保持不变,收敛速度相对较快.图 8 描述了这些模型的测试准确率随迭代训练次数变化的曲线

图 8 中 TDF-IDS 模型的测试准确率在收敛后基本高于

其他几个模型. 其中 TDF-IDS 与 DRSN-FL 的测试准确率在训练期间相对其他几个模型会有所波动,这是因为焦点损失函数会提高少数类样本与难分类样本的权重,从而降低网络对多数类样本与易分类样本的关注度,导致一定程度上会降低模型的准确率,但当模型趋于收敛后也能获得较高的准确率.

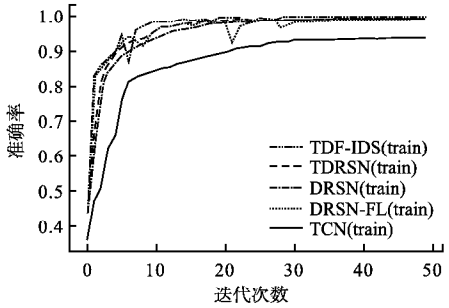

图 7 TCN、DRSN、DRSN-FL、TDRSN 与 TDF-IDS 模型训练准确率曲线

Fig. 7 TCN, DRSN, DRSN-FL, TDRSN and TDF-IDS model training accuracy curves

5)各个类别的多指标分析

消融实验模型在每个类别上的精确率、召回率、F1 值和 AUC如表5所示. 可以发现模型DRSN、TCN在ddos、normal、password这些大样本的类别上都有不低于95%的召回率,而在dos、scanning、mitm样本数量较少的情况下召回率相对低一些,特别是在极少数类别 mitm 上召回率为0. 也就表明了,这两个模型很少或不能学习到少数样本的特征,以致于模型对小样本的检测率比较低. DRSN-FL 相较于 DRSN 在少数类别上的各个指标均有所提高,特别是在 mitm 上提高最为明显,分别为25%、100%、40%、17.09%,这说明引入的焦点损失函数能够使模型学习到小样本数据的特征,提高对小样本的检测能力.

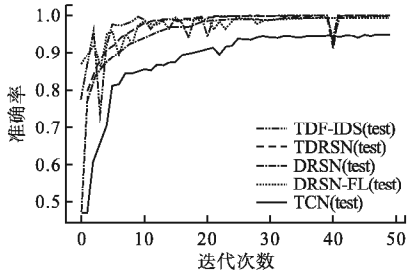

图 8 TCN、DRSN、DRSN-FL、TDRSN 与 TDF-IDS 模型测试准确率曲线

Fig. 8 TCN, DRSN, DRSN-FL, TDRSN and TDF-IDS model testing accuracy curves

表5 8种类别各项指标上的比较(%)

Table 5 Comparison of various indicators among 8 categories(%)

模型	指标	类别							
		dos	ddos	injection	normal	xss	password	scanning	mitm
DRSN	精确率	97.22	99.78	99.19	99.95	99.21	98.75	95.24	0
	召回率	94.59	99.67	99.19	99.95	99.60	99.58	97.56	0
	F1 值	95.89	99.72	99.19	99.95	99.41	99.16	96.39	0
	AUC	93.79	100	99.49	96.98	97.98	99.97	97.32	70.51
DRSN-FL	精确率	96.49	99.22	100	99.90	99.22	99.86	96.43	25
	召回率	99.10	99.67	100	99.28	100	97.90	98.78	100
	F1 值	97.78	99.44	100	99.59	99.61	98.87	97.59	40
	AUC	95.30	99.22	99.37	97.98	99.24	98.55	98.87	87.60
TDRSN	精确率	99.10	99.89	100	100	99.22	99.86	97.59	100
	召回率	99.10	100	99.19	100	100	100	98.78	20
	F1 值	99.10	99.94	99.59	100	99.61	99.93	98.18	33.33
	AUC	99.27	100	99.56	96.56	98.96	100	93.36	86.49
TCN	精确率	76.42	96.39	79.27	99.59	77.37	92.86	68.18	0
	召回率	72.97	98.22	52.85	99.69	83.79	96.36	54.88	0
	F1 值	74.65	97.29	63.41	99.64	80.46	94.58	60.81	0
	AUC	41.20	99.55	75.55	93.90	76.90	96.63	53.25	40.80
TDF-IDS	精确率	100	100	100	99.95	100	99.72	97.62	**100**
	召回率	97.30	99.78	100	100	100	100	100	**100**
	F1 值	98.63	99.89	100	99.97	100	99.86	98.80	**100**
	AUC	94.44	99.99	99.58	96.76	99.64	99.32	99.37	**99.95**
样本数		111	897	123	1949	253	715	82	5

使用了时域卷积网络与残差收缩网络的 TDRSN 模型相对于 TCN 模型在 mitm 攻击类型的各项指标上表现出明显提升,分别提高了 100%、20%、33.33% 和 45.69%. 同时,对于 dos、injection、xss、scanning 这 4 种攻击类型,TDRSN 模型也表现出了很大程度的性能提升,分别为 58.07%、24.01%、22.06% 和 40.11%. 与 DRSN 模型相比,TDRSN 模型在 mitm 攻击类型的各项指标上同样取得了显著提升,从原来的 0% 提高到了 100%、20%、33.33% 和 15.98%. 此外,对于 dos、injection、xss 和 scanning 等小样本攻击类型,TDRSN 模型的各项指标也有着明显的改善,提高了 5% 左右. 这也说明了融合时域卷积网络与残差收缩网络,能够提取小样本数据特征,一定程度上提高对小样本攻击类别的检测能力. 本文所提出的

TDF-IDS 模型在每个攻击类别上的各个指标上相比于其他模型基本都有所提高,特别是在极少数类别 mitm 上提高最大,且各个指标均大于 99.95%.尽管 TDF-IDS 在 dos 攻击类别上 AUC 仅为94.44%,相比于 TDRSN 降低了4.83%,这可能是因为引入焦点损失函数后,mitm 极少数类别被赋予更高的权重,从而降低了模型对该攻击类别的关注度,但总体 AUC 还是提高了 1.85%.

4.5.2　与其他模型比较研究

最后,在 TON-IoT 数据集上将本文提出的 TDF-IDS 模型与一些基准模型[19-22]及当前比较流行的模型进行对比.可以发现,本文所提出的方法各项指标均优于表中的方法,证明了 TDF-IDS 模型的有效性与优越性.对比数据如表 6 所示.

表 6　现有方法与本文方法的性能比较(%)

Table 6　Performance comparison between existing methods and this method(%)

方法	准确率	精确率	召回率	F1 值	AUC
文献[19]	99.59	–	–	99.59	–
文献[20]	96.35	94.80	98.45	97.03	
文献[21]	99.57	99.66	99.59	99.62	–
文献[22]	98.00	97.80	97.80	97.80	–
GRU	98.96	84.67	84.84	84.75	96.91
Resnet	99.06	85.06	84.56	84.76	84.79
本文方法	**99.88**	**99.66**	**99.63**	**99.64**	**98.65**

5　结　论

物联网的迅速发展使得对于异常和恶意攻击的检测变得至关重要.随着攻击数量的增加,迫切需要能够快速准确检测异常的工具.本文提出了一种称为 TDF-IDS 的新方法,它融合了 DRSN 模型和 TCN 模型的优势,并解决了入侵数据集样本不平衡的问题.TDF-IDS 能够从输入数据中提取时空特征,并通过改进的焦点损失函数增强了对小样本类别和难分类样本的检测能力.实验结果表明,TDF-IDS 在训练过程中收敛迅速,在测试集上表现出色,尤其是对于极少类别 mitm,精确率、召回率和 F1 值都达到了 100%,AUC 则达到了 99.95%.与其他先进模型相比,本文提出的方法在准确率、精确率、召回率、F1 值和 AUC 方面表现出色,所有指标均优于其他模型.本文的方法为物联网入侵检测提供了有益的探索和参考.在未来的工作中,我们计划将 TDF-IDS 应用于更多物联网数据集,以验证其可靠性和泛化性.此外,我们也将关注如何在提高模型对小样本的检测率的同时不影响对大样本的识别能力,这将是接下来的重点研究方向.

References:

[1] Madakam S, Lake V, Lake V, et al. Internet of Things(IoT): a literature review[J]. Journal of Computer and Communications, 2015, 3(5): 164-173.

[2] Vailshery L S. IoT and non-IoT connections worldwide 2010-2025[EB/OL]. https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/(Accessed on 3 January 2023), 2021.

[3] Sicari S, Rizzardi A, Grieco L A, et al. Security, privacy and trust in internet of things: the road ahead[J]. Computer Networks, 2015, 76: 146-164, doi: 10.1016/j.comnet.2014.11.008.

[4] Pajouh H H, Javidan R, Khayami R, et al. A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks[J]. IEEE Transactions on Emerging Topics in Computing, 2016, 7(2): 314-323.

[5] Yang L, Manias D M, Shami A. Pwpae: an ensemble framework for concept drift adaptation in iot data streams[C]//IEEE Global Communications Conference(GLOBECOM), 2021: 1-6.

[6] Abu Al-Haija Q, Al-Dala'ien M. ELBA-IoT: an ensemble learning model for botnet attack detection in IoT networks[J]. Journal of Sensor and Actuator Networks, 2022, 11(1): 18, doi: 10.3390/jsan11010018.

[7] Zhao M, Zhong S, Fu X, et al. Deep residual shrinkage networks for fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2019, 16(7): 4681-4690.

[8] Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv preprint arXiv: 1803.01271, 2018.

[9] Halbouni A, Gunawan T S, Habaebi M H, et al. CNN-LSTM: hybrid deep neural network for network intrusion detection system[J]. IEEE Access, 2022, 10: 99837-99849, doi: 10.1109/ACCESS.2022.3206425.

[10] Yin C, Zhu Y, Fei J, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. Ieee Access, 2017, 5: 21954-21961, doi: 10.1109/ACCESS.2017.2762418.

[11] Hu Z, Wang L, Qi L, et al. A novel wireless network intrusion detection method based on adaptive synthetic sampling and an improved convolutional neural network[J]. IEEE Access, 2020, 8: 195741-195751, doi: 10.1109/ACCESS.2020.3034015.

[12] Li Y, Xu Y, Liu Z, et al. Robust detection for network intrusion of industrial IoT based on multi-CNN fusion[J]. Measurement, 2020, 154: 107450, doi: 10.1016/j.measurement.2019.107450.

[13] Li Y, Yao R. Botnet detection method based on parallel CBAM-ResNet and self-attention[C]//IEEE International Conference on Signal Processing, Communications and Computing(ICSPCC), 2022: 1-6.

[14] Bhavsar M, Roy K, Kelly J, et al. Anomaly-based intrusion detection system for IoT application[J]. Discover Internet of Things, 2023, 3(1): 5, doi: 10.1007/s43926-023-00034-5.

[15] Yan B H, Han G D. Combinatorial intrusion detection model based on deep recurrent neural network and improved SMOTE algorithm[J]. Chinese Journal of Network and Information Security, 2018, 4(7): 48-59.

[16] Zhou X, Hu Y, Wu J, et al. Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT[J]. IEEE Transactions on Industrial Informatics, 2022, 19(1): 570-580.

[17] Khan S H, Hayat M, Bennamoun M, et al. Cost-sensitive learning of deep feature representations from imbalanced data[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 29(8): 3573-3587.

[18] Alsaedi A, Moustafa N, Tari Z, et al. TON-IoT telemetry dataset: a new generation dataset of IoT and IIoT for data-driven intrusion detection systems[J]. IEEE Access, 2020, 8: 165130-165150, doi: 10.1109/ACCESS.2020.3022862.

[19] Li S, Chai G, Wang Y, et al. CRSF: an intrusion detection framework for industrial internet of things based on pretrained CNN2D-RNN and SVM[J]. IEEE Access, 2023, doi: 10.1109/ACCESS.2023.3307429.

[20] Elsayed R A, Hamada R A, Abdalla M I, et al. Securing IoT and SDN systems using deep-learning based automatic intrusion detection[J]. Ain Shams Engineering Journal, 2023, 14(10): 102211, doi: 10.1016/j.asej.2023.102211.

[21] Ahmad J, Shah S A, Latif S, et al. DRaNN_PSO: a deep random neural network with particle swarm optimization for intrusion detection in the industrial internet of things[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(10): 8112-8121.

[22] Gad A R, Nashat A A, Barkat T M. Intrusion detection system using machine learning for vehicular ad hoc networks based on TON-IoT dataset[J]. IEEE Access, 2021, 9: 142206-142217, doi: 10.1109/ACCESS.2021.3120626.