•信息工程•

DOI:10.15961/j.jsuese.202100984

本刊网刊

基于特征选择和时间卷积网络的工业控制系统入侵检测

石乐义[1,2], 侯会文[1], 徐兴华[2], 许翰林[1], 陈鸿龙[3]

(1.中国石油大学(华东)计算机科学与技术学院, 山东 青岛 266580; 2.中国石油大学(华东)海洋与空间信息学院, 山东 青岛 266580; 3.中国石油大学(华东)控制科学与工程学院, 山东 青岛 266580))

摘　要: 针对工业控制系统流量数据存在特征冗余及深度学习模型对较小规模数据集检测能力较差的问题, 提出了一种基于特征选择和时间卷积网络的工业控制系统入侵检测模型。首先, 对源域数据集的异常特征和样本不平衡数据进行处理, 提高源域数据集质量。其次, 针对流量数据的特征冗余, 利用信息增益率和主成分分析法构建IGR–PCA特征选择算法, 筛选出最优特征子集实现数据降维。然后, 根据工业控制系统流量数据的时间序列特性, 在较大规模的源域数据集上, 利用时间卷积网络(temporal convolution network, TCN)对时间序列数据优异的处理能力, 构建源域时间卷积网络预训练模型。最后, 在较小规模的目标域数据集上, 结合迁移学习(transfer learning, TL)微调策略, 获取源域样本数据的流量特征, 构建目标域TCN–TL模型。利用公开的工业控制系统数据集进行实验测试, 实验结果表明: 流量数据经本文特征算法处理后, 相较于其他方法, 在降低数据维度减少计算量的同时仍具有良好的检测效果; 在较大规模的源域数据集和较小规模的目标域数据集上, 本文模型均取得了良好的检测效果; 在目标域中利用迁移学习微调策略能够学习到源域中的知识, 模型检测准确率为99.06%; 在训练时间对比中, 本文模型训练时间消耗更少, 具有更好的泛化能力, 能够更好地保护工业控制系统安全。

关键词: 工业控制系统; 入侵检测; 特征选择; 时间卷积网络; 迁移学习

中图分类号: TP391　　**文献标志码**: A　　**文章编号**: 2096-3246(2022)06-0238-10

Industrial Control System Intrusion Detection Based on Feature Selection and Temporal Convolutional Network

SHI Leyi[1,2], *HOU Huiwen*[1], *XU Xinghua*[2], *XU Hanlin*[1], *CHEN Honglong*[3]

(1.School of Computer Sci. and Technol., China Univ. of Petroleum (East China), Qingdao 266580, China; 2.School of Oceanography and Space Info., China Univ. of Petroleum (East China), Qingdao 266580, China; 3.School of Control Sci. and Eng., China Univ. of Petroleum (East China), Qingdao 266580, China)

Abstract: Aiming at the problem of feature redundancy in industrial control system traffic data and the poor detection ability of deep learning models for small-scale data sets, an industrial control system intrusion detection model based on feature selection and temporal convolutional networks was proposed. First, the abnormal features and sample imbalance data of the source domain dataset were processed to improve the quality of the source domain dataset. Secondly, in view of the feature redundancy of traffic data, a IGR–PCA feature selection algorithm was constructed by using the information gain rate and principal component analysis method, and the optimal feature subset was selected to achieve data dimensionality reduction. Then, according to the time series characteristics of industrial control system traffic data, the excellent processing ability of temporal convolution network (TCN) for time series data was used to construct a source domain temporal convolution network pretrained model on a large-scale source domain data set. Finally, combined with the transfer learning (TL) fine-tuning strategy, the traffic characteristics of the source domain sample data were obtained on a small-scale target domain dataset, and the target domain TCN–TL model was constructed. The experimental test was carried out using the public industrial control system data set. The experimental results showed that compared with other

收稿日期:2021 – 09 – 28

基金项目:国家自然科学基金项目(61772551); 山东省自然科学基金项目(ZR2019MF034)

作者简介:石乐义(1975—), 男, 教授, 博士, 博士生导师. 研究方向: 网络安全; 博弈论和移动计算. E-mail: shileyi@upc.edu.cn

网络出版时间:2022 – 07 – 26 10 : 39 : 59　　**网络出版地址**:https://kns.cnki.net/kcms/detail/51.1773.TB.20220725.0947.004.html

methods, the proposed method can reduce the data dimension and reduce the calculation amount while still having a superior detection effect. The model proposed in this paper has achieved good detection results on both large-scale source domain data sets and small-scale target domain data sets. In the target domain, the transfer learning fine-tuning strategy can be used to learn the knowledge in the source domain, and the detection accuracy rate is 99.06%. In the training time comparison, the proposed model consumes less training time. Meanwhile, it also has better generalization ability and can better protect the security of industrial control systems.

Key words: industrial control system; intrusion detection; feature selection; temporal convolutional network; transfer learning

工业控制系统（industrial control system, ICS）是用于工业生产的控制系统的统称[1]，由各种工业设备组件构成，负责工业生产流程中的监测控制和资源调度，实现设备的自动化运行，为现代化工业提供支撑，是国家基础设施建设的重要组成部分。近年来，针对工业控制系统的安全事件频发，如伊朗震网病毒、火焰病毒、WannaCry勒索病毒事件等，对国家经济和社会都带来了严重的危害。

工业控制系统安全事件揭示了所面临的安全问题，究其原因主要有以下几个方面：首先，工业控制系统的软硬件设备固有的安全漏洞和在更新迭代过程中产生的安全漏洞使黑客有可乘之机；其次，现代化的工业控制系统网络环境日趋开放，更具有开放性和共享性，使入侵途径增多，攻击风险大大增加；最后，工业控制系统的重要性日益显现，也使其逐步成为网络攻击的首要目标。

为应对日益严峻的工业控制系统安全形势，如何保护工业控制系统的安全已成为亟待解决的问题。安全领域研究人员已经提出了防火墙、访问控制、信息加密等防御策略，但以上安全方案因自身安全防御策略的不足，虽能抵御攻击入侵，但不能快速、高效地实现安全防御。入侵检测是一种对流量数据进行检测并异常报警的安全防护技术，能够发现潜在的恶意活动或入侵，是对其他安全技术方案的升级，能够有效地实现对工业控制系统的实时监测，保证工业控制系统的安全运行。

深度学习技术作为机器学习领域的前沿技术，模型的学习能力会随着模型深度的增加而呈指数增长，在计算机视觉[2]、自然语言处理[3]、语音识别[4]等领域得到广泛的应用。在入侵检测领域，Liu等[5]提出一种新的两级检测框架，一级检测采用CNN进行特征提取，二级检测提出状态转移算法，将CNN模型提取的特征作为算法的输入，建立ICS的正常状态的转换方程，有效地实现了入侵检测。Mirza等[6]使用LSTM处理计算机网络数据，能够同时处理固定长度和可变长度的数据序列。石乐义等[7]利用相关信息熵进行特征选择，运用CNN–BiLSTM并融合多头注意力机制进行入侵检测，取得了较低的漏报率。Chawla等[8]提出一种基于递归神经网络的高效计算的入侵检测系统，将堆叠CNN和GRU相结合以检测异常。Yan等[9]构建了一种基于自动编码器和LSTM的入侵检测模型，将深度学习方法用于特征提取。以上方法通过对大量流量数据进行训练，能够及时地发现入侵攻击行为，大幅提升了模型的检测效果，但这些方法也有一定的局限性，带来了更大的系统开销和更长的模型训练时间。同时，面对复杂多变的安全威胁，在收集到的攻击流量数据不足、数据集规模小时，深度学习模型无法进行有效的训练，使入侵检测性能降低。

针对流量数据不足，数据集规模小等问题，Mathew等[10]结合迁移学习方法，使用Inception模型作为初始模型进行微调，在专有的ATM监控训练集上训练模型，获得了更好的准确率。迁移学习（transfer learning, TL）被设计成利用数据、模型和任务之间的相似性，将从源域中学习到的知识迁移到目标域，来帮助目标域训练，能够有效地解决目标域训练数据不足的问题，具有更广泛的应用前景[11]。然而，在工业控制系统入侵检测领域，系统通过访问时间戳顺序收集流量数据信息，致使采集到的流量数据具有强烈的时间特性，仅将迁移学习引入到工业控制系统入侵检测中不能充分利用工业控制系统流量数据的时间特性为入侵检测带来更好的效果。

针对时间序列检测，循环神经网络获得了广泛的应用，但大部分存在梯度消失、并行性差、模型难以收敛等问题，导致入侵检测的精度较低。相比于传统的循环神经网络，时间卷积网络（temporal convolution network, TCN）被用于时间序列分析时，具有更好的性能表现[12]，可以为工业控制系统入侵检测带来更好的检测效果。

由于工业控制系统需要处理大量的数据导致采集的流量数据具有特征冗余和特征不相关特性，不仅增加了模型的复杂度，降低了检测速度，而且需要消耗大量的计算资源[13]。针对该问题，许多研究者采用特征选择的方法实现数据降维并达到去除数据冗余的目的。利用特征选择算法寻找最优的特征子集以实现数据降维，不仅可以提高入侵检测效率还可以降低系统的开销，节约成本，这对工业控制系统入侵检测至关重要。Chen等[14]提出一种基于主成分分

析(PCA)、决策树和朴素贝叶斯的自适应网络入侵检测模型,使用PCA去除不重要信息,实现数据降维。唐成华等[15]通过FCM算法全局搜索,利用信息增益算法进行特征排序,结合约登指数删减冗余特征。Jadhav等[16]提出一种信息增益定向特征选择算法(IGDF),利用信息增益执行特征的排序筛选。上述方法由于仅采用单一的特征选择方法,不能同时兼顾特征本身及特征之间的联系导致特征选择效果并没有达到最优。

针对上述工业控制系统流量数据不足、特征冗余等问题,为更好地处理工业控制系统数据流量,提高模型的检测能力,本文结合迁移学习提出一种工业控制系统入侵检测方法。首先,结合IGR–PCA特征选择算法对数据进行预处理,在降维的同时提高源域数据集质量。其次,在较大规模数据集上利用时间卷积网络搭建源域TCN预训练模型学习源域知识。最后,在源域TCN预训练模型的基础上,引入迁移学习的思想,在较小规模数据集上对模型进行微调,搭建目标域TCN–TL入侵检测模型,减少训练过程的时间消耗。

1 工业控制系统入侵检测模型

为降低工业控制系统数据集的特征冗余,实现对流量数据的有效检测,本文建立了工业控制系统入侵检测模型,该模型整体流程如图1所示。本文模型整体流程主要包含5个部分:数据预处理模块、特征选择模块、源域TCN预训练模型、目标域TCN–TL入侵检测模型、入侵检测结果输出模块。

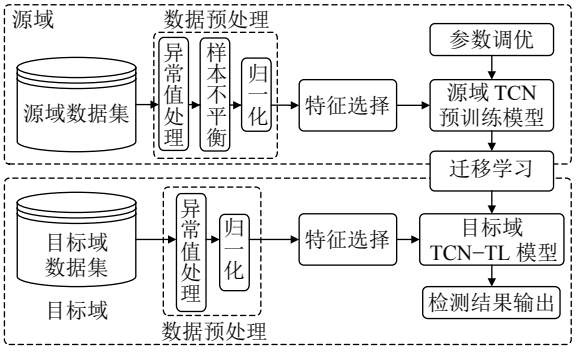

图 1 本文模型整体流程

Fig. 1 Overall process of the proposed model

1.1 数据预处理

在工业控制系统中,数据采集与监控系统(SCADA)负责对现场设备节点进行集中监测和远程控制,利用传感器收集流量数据信息,其中,包括报警信息、过程状态数据等,对数据的采集发挥着重要的作用。在工控系统中,网络攻击会使传感器测量值严重越界,明显超出警报点设置范围的过程测量值,从而产生异常值;同时,SCADA通常用于监测流量必经的链路并进行数据采集,因此获得的数据流量中正常的样本总是占多数,某些攻击类型的样本数量偏少,存在样本不平衡问题;另外,采集到的数据参数数值范围不同,存在数值量差异大的问题。因此,需要先对数据集进行数据预处理。

1.1.1 异常值处理

在工控数据集中,“measurement”特征表示天然气管道压力,响应注入攻击(MRI)和侦察攻击(RECO)会造成该特征异常,如该特征数值为8.66E+26,数值大小明显异常。从工业控制系统实际情况出发,为特征“measurement”设定阈值,超出正常压力范围则重新赋值。数据流量二分类时,可考虑将“measurement”数据异常直接判定为异常流量;数据流量多分类时,为该特征大小设定阈值,结合其他数据流量特征继续细分攻击类别,以便更好地进行安全防护。

1.1.2 不平衡数据集处理

在工控数据集中存在着样本不平衡问题,为解决数据样本不平衡,本文使用SMOTE–Tomek Links方法[17],先使用SMOTE方法对少数类样本进行过采样处理,再结合数据清洗技术Tomek Links解决SMOTE方法在生成少数类样本时容易产生样本重叠的问题,使合成后的样本数据更加合理。

SMOTE–Tomek Links的过程为:

1)首先,随机选择少数类样本V_i;然后,从数据中设置K近邻,在样本V_i和随机选择的K近邻之间生成合成数据,将合成样本添加到少数类中,重复此步骤,直到达到所需的少数类样本比例,新样本合成如式(1)所示:

$$V_{\text{new}} = V_i + (V_j - V_i)\delta \tag{1}$$

式中,V_j为选出的邻近点,δ表示范围为[0,1]的随机数。

2)寻找并删除Tomek Links对。

设$d(V_m,V_n)$表示V_m与V_n之间的欧几里得距离,其中,V_m为少数类的样本,V_n为多数类的样本;如果没有样本V_k满足$d(V_m,V_k)<d(V_m,V_n)$或$d(V_n,V_k)<d(V_m,V_n)$,则(V_m,V_n)对是Tomek Links对,应删除Tomek Links对。

1.1.3 归一化处理

工控数据集在异常值处理后,还存在着数据差异大的问题,采用归一化方法可以消除不同特征之间的差异,本文采用MinMax归一化方法将特征值映射到区间[0,1]之间:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

式中,x'为经过归一化之后的数据,$x_{\min}$为该特征数据的最小值,$x_{\max}$为该特征数据的最大值。

1.2　基于IGR–PCA的特征选择算法

工业控制系统数据维度更高，特征之间存在相关性且有冗余特征，会增加模型复杂度，降低分类精度。基于此背景，为筛选出最优特征子集，实现数据的降维，本文提出一种基于混合信息增益率（information gain ratio，IGR）和主成分分析（principal component analysis，PCA）的IGR–PCA特征选择算法。

信息增益被广泛用于特征选择，信息增益率相比于信息增益，引入分裂信息能够对维度较高的特征进行惩罚，减少对维度较高特征的倾向[18]。假设工控数据集$X=\{x_1,x_2,x_3,\cdots,x_n\}$，其中$n$为样本总数；数据集特征集合$F=\{f_1,f_2,f_3,\cdots,f_m\}$，其中$m$为特征种类的数量，则信息熵的计算可表示为：

$$Entropy(X)=-\sum_{i=1}^{n}P(x_i)\mathrm{lb}\ P(x_i) \tag{3}$$

进而，得到特征f的信息增益率：

$$GainRatio(X,f)=\frac{Entropy(X)-\sum_{i=1}^{k}\frac{f_i}{X}Entropy(f_i)}{SplitInfo(X,f)} \tag{4}$$

式中：k为根据特征f划分子集数量的种类；f_i为第i类子集的数量；$SplitInfo(X,f)$表示将数据按照特征f划分为k个子集的分裂信息，其公式如下：

$$SplitInfo(X,f)=-\sum_{i=1}^{k}\frac{|f_i|}{|X|}\mathrm{lb}\frac{|f_i|}{|X|} \tag{5}$$

通过信息增益率去除冗余特征，筛选出对分类影响较大的特征，初步实现数据的降维，但在降维过程中忽略了特征之间的相关性。为保留数据的原始信息，减少特征相关性对流量数据分类的影响，接下来使用主成分分析进行降维。主成分分析是有效的数据分析方法，能够最大限度地在保留原始数据信息的基础上，实现高维数据向低维数据的转换[19]。主成分分析的主要步骤如下：

1）标准化样本数据：

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

式中，x_s为标准化后的样本数据，μ为特征数据的均值，σ为特征数据的标准差。

2）计算协方差矩阵$\boldsymbol{C}$：

$$\boldsymbol{C}=\frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{X}_s\cdot\boldsymbol{X}_s^{\mathrm{T}}) \tag{7}$$

式中，$\boldsymbol{X}_s$为标准化后的样本矩阵，N为变量的个数。

3）计算协方差矩阵$\boldsymbol{C}$的特征值λ_j与特征向量$\boldsymbol{u}_i$，根据特征值计算主成分的贡献率和累计贡献率，选取前k个累计贡献率达85%的特征值对应的特征向量构成特征变换矩阵$\boldsymbol{T}$。

4）根据标准化后的样本数据$\boldsymbol{X}_s$和特征变换矩阵$\boldsymbol{T}$计算获得新的样本矩阵$\boldsymbol{Y}$。

IGR–PCA算法的基本思想：使用信息增益率从数据集特征中筛选出重要的特征，引入主成分分析法，去除重要特征之间的相关性，再次降低维度，实现特征选择，从而提高模型的训练速度。IGR–PCA特征选择算法的详细内容如下：

算　法　基于IGR–PCA的特征选择算法

输入：工控数据集X，流量特征集合F

输出：特征选择后的特征矩阵$\boldsymbol{Y}$

1. For each $f_i\in F$
2. 　　根据式（3）计算f_i的信息熵；
3. 　　根据式（4）计算f_i的信息增益率；
4. 　　数据按照IGR降序排列，筛选重要特征；
5. 　　得到新的样本矩阵$\boldsymbol{F}_{\mathrm{IGR}}$；
6. End For
7. 根据式（6）对$\boldsymbol{F}_{\mathrm{IGR}}$进行标准化处理得到矩阵$\boldsymbol{X}_s$；
8. 根据式（7）计算协方差矩阵$\boldsymbol{C}$；
9. 计算$\boldsymbol{C}$的特征值λ_j与特征向量$\boldsymbol{u}_i$，按照累计贡献率选取前k个特征向量构成特征变换矩阵$\boldsymbol{T}$；
10. 获得新的样本矩阵$\boldsymbol{Y}$；
11. 返回$\boldsymbol{Y}$。

1.3　源域TCN预训练模型

为解决时间序列预测问题，TCN引入1维卷积、因果卷积、扩张卷积和残差块等，在时间序列中表现优异[20]。本文将依据工业控制系统数据流量的时间序列特性，构建TCN入侵检测模型，利用TCN对时间序列的优异表现，将TCN模型作为源域预训练模型，TCN网络模型图如图2所示。

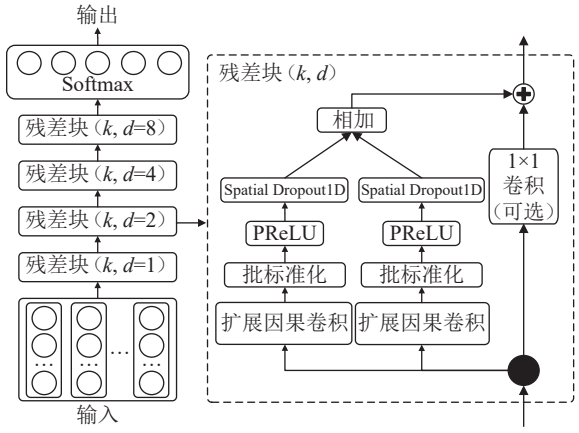

图 2　TCN模型图

Fig. 2　TCN model diagram

时间卷积网络使用1维全卷积网络（FCN）架构，网络产生与输入相同长度的输出；使用因果卷积，在t时刻的入侵检测报警输出只与时间t或者更早的时

间序列进行卷积，不会丢失历史数据，能够保存较长时间的数据流量信息，工业控制系统的数据流量相对固定，当出现相同类型的攻击流量时，能够更好地进行检测。

TCN的体系结构可以简单描述为1维全卷积网络和因果卷积网络的组合，在处理时间序列时，期望获得较长时间的数据流量信息，为使模型获得更大范围的感受野，引入扩张卷积。对于1维序列$x \in R^n$，滤波器$g:\{0,1,\cdots,w-1\} \to R$，扩张卷积运算$G$定义为：

$$G(s) = (x \cdot g_d)(s) = \sum_{i=0}^{w-1} g(i) \cdot x_{s-d \cdot i} \tag{8}$$

式中，g_d为带有扩张因子d的卷积操作，w为卷积核大小，$s-d \cdot i$为过去的方向。

扩张卷积通过扩张因子大小来控制感受野的大小，当扩张因子d=1时，变成普通卷积。图3描述了当扩张因子d=1,2,4和卷积核大小为3时的扩张因果卷积。使用更大的扩张因子可以增加感受野，也可以在参数量不变的情况下减少网络深度。

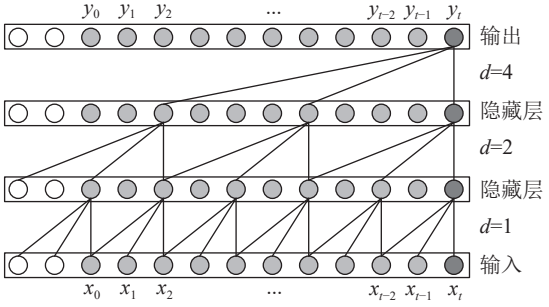

图 3　扩张因果卷积图

Fig. 3　Diagram of dilated causal convolution

残差块使用跳跃连接，通过一系列变换φ，将输出添加到输入x中获得最终的输出结果o：

$$o = Activation(x + \varphi(x)) \tag{9}$$

式中，$Activation(\ )$为激活函数。

使用残差块代替卷积层，以较少的层获得较长的依赖关系，使网络更加易于训练和收敛，避免了深度学习模型中梯度消失的问题。如图2所示，本文提出的并行结构残差块，残差块包含两层扩张的因果卷积和参数化线性修正单元（PReLU）。在扩张卷积层后，使用正则化减少过拟合；同时为确保输入和输出具有相同的宽度，使用额外的1维卷积来确保输出具有相同形状的张量，增加系统的稳定性。

1.4　目标域TCN–TL入侵检测模型

迁移学习通过迁移源域中的知识来提高目标域的表现[21]，可以将训练好的模型参数迁移到新的模型来帮助新的模型训练，将迁移学习应用于入侵检测领域，可以减轻深度学习模型对大量数据的依赖，更好地实现检测。

将源域（source domain）表示为D_S，源域数据可表示为$D_S:(X_S,Y_S)=\{(X_{S1},Y_{S1}),(X_{S2},Y_{S2}),\cdots,(X_{Sn},Y_{Sn})\}$；将目标域（target domain）表示为$D_T$，目标域数据可表示为$D_T:(X_T,Y_T)=\{(X_{T1},Y_{T1}),(X_{T2},Y_{T2}),\cdots,(X_{Tn},Y_{Tn})\}$。将TCN作为源域预训练模型，利用迁移学习微调策略，对源域预训练模型进行微调，构建目标域TCN–TL模型，构建过程如图4所示。具体步骤为：1）以TCN模型当作源域特征提取层，获取TCN预训练模型的权值。2）对时间卷积网络较浅的层数进行冻结，不再参与后续模型的训练；在冻结部分网络的基础上微调网络模型。3）使用目标领域的数据集进行训练，使用Adam算法进行参数更新；当模型精度不再变化，保留模型新的参数和结构，模型训练完毕。

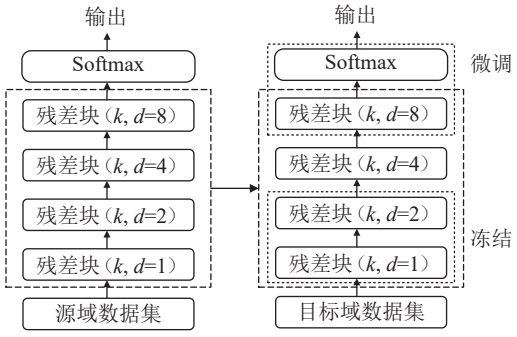

图 4　TCN–TL模型构建过程

Fig. 4　TCN–TL model building process

2　实验与结果

2.1　数据集与评价标准

实验环境：Windows10操作系统、Intel(R) Core(TM) i7–7700 CPU、24 GB内存，基于Tensorflow2.0和Python3.6实现。使用密西西比州立大学于2014年公开的工控入侵检测数据集[22]，数据集中包含正常网络流量和7种攻击网络流量，详细类别描述如表1所示。

表 1　数据集类别描述

Tab. 1　Dataset category description

样本类别	标签编码	标签描述
Normal	0	正常样本数据
NMRI	1	普通的恶意响应注入攻击
CMRI	2	复杂的恶意响应注入攻击
MSCI	3	恶意状态命令注入攻击
MPCI	4	恶意参数命令注入攻击
MFCI	5	恶意功能命令注入攻击
DoS	6	拒绝服务攻击
RECO	7	侦察攻击

该数据集总共包含4个数据集，为最大程度地减

少内存需求和处理时间，使用精简的10%天然气数据集。该天然气数据集中包含10 619条样本数据，每条样本数据中包含26个流量特征和1个类别标签。为获得较大规模的源域数据集和较小规模的目标域数据集，以满足实验要求，对天然气数据集进行拆分，天然气数据集划分如表2所示。源域的数据集D_S进行了样本不平衡处理，最终获取9 320条数据；目标域数据集D_T未做样本不平衡处理，最终获取1 598条数据。

表 2　天然气数据集和划分

Tab. 2　Gas Pipeline dataset and partition

样本类型(标签编码)	天然气数据集	D_S数据集	D_T数据集
Normal(0)	6 672	5 610	1 029
NMRI(1)	355	275	49
CMRI(2)	1 664	1 420	242
MSCI(3)	93	190	14
MPCI(4)	842	711	124
MFCI(5)	41	182	5
DoS(6)	189	257	22
RECO(7)	783	675	108
合计	10 619	9 320	1 598

为评估本文模型的检测性能，使用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1值(F1-measure)作为评价指标，具体公式如下：

$$A_c = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{10}$$

$$P_r = \frac{T_p}{T_p + F_p} \tag{11}$$

$$R_e = \frac{T_p}{T_p + F_n} \tag{12}$$

$$F_1 = \frac{2T_p}{2T_p + F_p + F_n} \tag{13}$$

式中，T_p为正确分类正常行为，T_n为正确分类攻击行为，F_p为错误分类攻击行为，F_n为错误分类正常行为。

2.2　数据预处理分析

为了验证数据处理的有效性，本文进行多组对比实验，包括异常值处理前后效果对比、使用SMOTE–Tomek Links方法处理样本不平衡前后效果对比。

2.2.1　异常值处理前后效果

在工业控制系统中，恶意响应注入攻击会使传感器测量值严重越界，明显超出警报设置点范围的过程测量值，因此产生了一些特征异常值。为说明数据异常值处理的必要性进行了对比实验，异常值处理前后的实验结果如表3所示。

表 3　异常值处理前后结果对比

Tab. 3　Comparison of results before and after outlier processing

实验方法	准确率/%	精确率/%	召回率/%	F1值/%
异常值未处理	94.68	85.60	78.61	81.05
异常值处理后	97.36	98.59	90.64	93.79

由表3可知：对特征“measurement”进行异常值处理后，在准确率、精确率、召回率、F1值方面均获得了明显提升，大幅提高了检测效果。这是因为：特征异常值未进行处理时，特征数据之间差异极大，在对数据进行归一化，会造成正常流量样本和攻击流量样本差异过小，使攻击流量样本被误分类为正常流量样本。

2.2.2　样本不平衡处理效果对比

仅对源域数据集进行样本不平衡处理，采用SMOTE–Tomek Links方法对少数类攻击样本进行过采样并进行数据清洗。仅对MSCI、MFCI、DoS这3类样本进行样本不平衡处理。表4是利用初始配置的TCN模型对样本不平衡处理前后的源域数据集进行对比实验的结果。

表 4　样本不平衡处理前后结果对比

Tab. 4　Comparison of results before and after sample imbalance handling

实验方法	准确率/%	精确率/%	召回率/%	F1值/%
样本不平衡处理前	97.36	98.59	90.64	93.79
样本不平衡处理后	98.67	99.18	95.06	96.91

由表4可知：通过SMOTE–Tomek Links算法对数据集进行处理后，提高了数据集的质量；SMOTE–Tomek Links算法可以有效地提升召回率和F1值，准确率和精确率也得到相应提升。

2.3　IGR–PCA特征选择算法实验

利用信息增益率进行初步的特征筛选，信息增益率越大表示该特征对分类的贡献度越高，通过排序选取了18个特征构成新的数据集X。在PCA降维阶段，对数据集X的18维特征进行分析，最终选取了贡献度最高的14个主成分，即k=14，构成最优特征子集。本文采用基于IGR–PCA的特征选择算法进行特征选择，并与其他特征选择算法IGR、PCA、KPCA算法进行对比实验，结果如图5所示。图5中，None表示未采用任何特征选择算法。

由图5可知，数据经本文提出的特征选择算法处理后，其准确率、召回率和F1值得到了明显的提升，其中，对准确率的提升最为明显，其准确率为98.55%，相较于仅经过IGR处理的方法准确率提升了1.38%。实验结果表明，本文提出的基于IGR–PCA的特征选择算法在降低数据维度减少模型计算量的同时仍能

保持良好的检测效果。

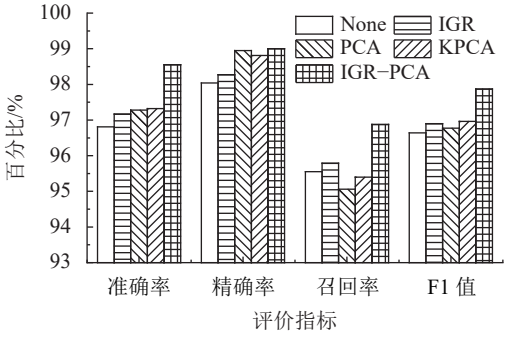

图 5 不同特征选择算法实验结果对比

Fig. 5 Comparison of experimental results of different feature selection algorithms

2.4 源域TCN预训练模型实验分析

为了取得较好的模型参数配置，首先，进行多组参数设置测试；其次，为了验证本文提出的源域TCN预训练模型的有效性，与其他入侵检测方法进行检测效果对比；最后，采用本文源域TCN模型进行源域数据多分类测试。

2.4.1 参数设置测试

为获得最佳的模型参数配置，对模型进行了多组参数设置测试，主要影响参数包括扩张因子d和卷积核大小w。实验评价指标包括准确率、精确率、召回率和F1值，同时，考虑模型训练时间t，将其作为新的评价指标。对参数d、w采用不同设置进行测试的实验结果如表5所示。

表 5 参数d、w的不同设置的测试结果

Tab. 5 Test results for different setting of parameters d and w

参数		准确率/%	精确率%	召回率/%	F1值/%	t/s
d	w					
2	1	98.35	99.07	95.32	96.96	9.92
2	3	98.54	99.27	93.81	96.22	12.72
2	5	98.68	99.57	93.86	96.38	14.50
4	1	98.87	99.36	96.60	97.89	16.42
4	3	99.01	99.71	95.97	97.65	22.04
4	5	98.92	99.69	95.29	97.21	25.15
8	1	99.19	99.40	97.75	98.55	30.94
8	3	99.10	99.75	96.28	97.86	35.67
8	5	99.05	99.74	95.98	97.67	38.48

由表5可知：当本文源域TCN预训练模型的扩张因子d取8，卷积核大小w取1时，准确率、召回率、F1值的效果最好。随着扩张因子d的增加，模型的层数和参数量在递增，造成训练时间变长。综合考虑，本文选取了扩张因子d=8，卷积核大小w=1的参数配置。

深度学习模型中良好的参数设置能够有效提升检测分类的性能，经多组参数设置实验，最终确立了源域TCN预训练模型的参数配置，模型参数设置如表6所示。

表 6 源域TCN模型参数设置

Tab. 6 Parameter settings of source domain TCN model

参数名称	参数值
扩张因子	8
卷积核大小	1
隐藏层节点	32
激活函数	PReLU
优化器	Adam
损失函数	多分类交叉熵

2.4.2 源域的不同模型对比实验分析

为了验证本文提出的源域TCN预训练模型的检测性能，将RNN、LSTM[23]、BiLSTM、OCC–eSNN[24]、CNN–LSTM[25]、PPO2[26]、HCIPSO–OCSVM[27]在入侵检测中的相关方法及本文的源域TCN模型进行检测效果对比实验，结果如表7所示。

表 7 源域的不同深度学习模型检测效果

Tab. 7 Detection effect of different deep learning models in source domain

方法	准确率/%	精确率/%	召回率/%	F1值/%	t/s
RNN	95.11	98.10	82.10	84.83	58.40
LSTM	95.81	97.98	85.05	88.87	64.41
BiLSTM	98.71	99.38	95.33	97.03	84.96
OCC–eSNN	98.82	98.80	96.80	97.40	76.24
CNN–LSTM	98.66	99.40	94.90	96.82	91.35
PPO2	99.10	97.17	98.56	97.85	—
HCIPSO–OCSVM	98.58	98.82	—	98.89	—
本文源域TCN	99.25	99.52	97.57	98.51	55.45

从表7可知：本文的源域TCN模型在拥有较大规模数据集的源域上分类准确度和精确度最高，分别达到99.25%和99.52%；对于时间序列数据，时间卷积网络保存较长时间的数据信息，更好地进行特征提取，能够满足工控系统的安全需求。从召回率和F1值看出，本文模型明显优于传统的循环神经网络RNN和LSTM，因为RNN和LSTM在训练过程中容易出现梯度爆炸，降低了检测性能。在训练时间消耗方面，本文模型具有更短的训练时间，能够满足工控系统对时间的需求。

2.4.3 源域数据多分类结果测试

为了说明本文源域TCN预训练模型对Normal（正常流量）及NMRI、CMRI、MSCI、MPCI、MFCI、DoS、RECO 7种攻击类型的检测效果，图6展示了本文的源域TCN模型在源域数据集各类样本上的检测效果。

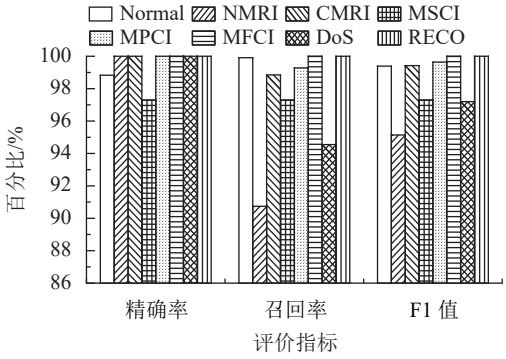

图 6　本文模型的源域多分类结果

Fig. 6　Source domain multiple classification results of the proposed model

从图6可知：本文模型对MFCI和RECO两种攻击类型具有最好的检测性能；对MSCI和DoS两种攻击类型检测效果欠佳；对NMRI攻击类型的召回率较低，召回率为90.74%。

2.5　目标域TCN–TL实验分析

本文使用源域的TCN预训练模型，结合迁移学习的微调策略，学习源域模型的权重，将源域学习到的知识迁移到目标域。使用1 598条样本数据进行实验，先对目标域的不同模型的检测效果进行对比，再进行目标域数据多分类测试。

2.5.1　目标域的不同模型对比实验分析

参照源域实验设置，在目标域数据集上进行检测效果实验，利用RNN、LSTM[23]、BiLSTM、CNN–LSTM[25]、未采用迁移学习的TCN模型与本文采用迁移学习的目标域TCN–TL模型进行检测效果测试对比实验，结果如表8所示。

表 8　目标域的不同深度学习模型检测效果

Tab. 8　Detection effect of different deep learning models in target domain

方法	准确率/%	精确率/%	召回率/%	F1值/%	t/s
RNN	93.41	54.80	56.88	55.80	11.00
LSTM	93.10	53.07	56.81	54.84	13.31
BiLSTM	84.64	75.95	77.55	74.52	16.85
CNN–LSTM	93.10	98.10	91.27	93.74	18.45
未采用迁移学习的TCN	98.75	99.46	90.95	93.83	14.00
本文采用迁移学习的目标域TCN–TL	99.06	99.53	97.89	98.63	6.90

由表8可知：本文的TCN–TL模型在较小规模的目标域数据集上的检测效果最好；RNN、LSTM等深度学习模型需要干净的大规模数据集，在较小规模的目标域数据集上表现差。其原因是：本文的TCN–TL模型使用迁移学习的方法，降低了深度学习模型对大规模数据集的要求，提升了模型的适应性；同时，本文模型使用迁移学习微调策略，继承了源域TCN模型参数，相比于原始的TCN模型，其训练时间节省7.1 s，在数据集训练阶段的时间消耗上耗时最少，更加符合工业控制系统入侵检测的要求。

2.5.2　目标域数据多分类结果测试

为了验证本文目标域TCN–TL模型对Normal（正常流量）及NMRI、CMRI、MSCI、MPCI、MFCI、DoS、RECO 7种攻击类型的检测效果，对目标域数据集各类样本的精确率、召回率和F1值进行了统计，目标域模型的多分类结果如图7所示。

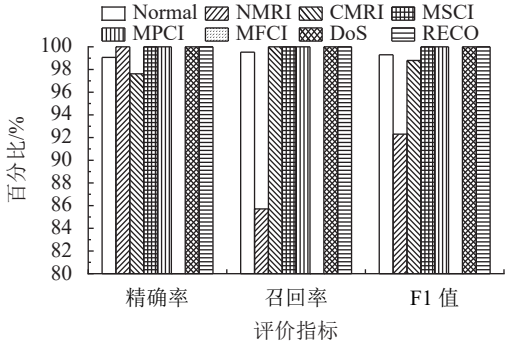

图 7　本文模型的目标域多分类结果

Fig. 7　Target domain multiple classification results of the proposed model

从图7可以看出，本文的目标域模型对MSCI、MPCI、MFCI和RECO 4种攻击类型具有最好的检测性能，对NMRI攻击类型的检测效果欠佳，对MFCI攻击样本的检测效果最差。其原因是在目标域测试阶段，MFCI只有一个样本用于测试集测试，本文目标域TCN–TL模型对该样本分类错误，导致精确率、召回率和F1值均为0。对多分类的实验结果综合考虑，本文模型取得了较好的分类效果，但对小样本攻击数据的检测有待进一步提升。

3　结　论

本文提出一种基于特征选择和时间卷积网络的工业控制系统入侵检测方法。针对流量数据特征冗余，将IGR–PCA特征算法应用于工控系统数据降维，筛选出最优特征子集。针对工业控制系统的时间特性，为更好地实现对小规模数据集的检测，结合迁移学习微调策略，依次构建了源域TCN预训练模型和目标域TCN–TL模型。实验结果表明，本文模型在源域和目标域上都具有较好的检测效果，针对目标域的小规模数据集，采用迁移学习方法，利用源域数据集进行预训练，能够有效地降低对训练样本数量的依赖，在提升模型检测性能的同时减少训练时间的消耗。在未来的研究工作中，将考虑源域数据与目标域数据的分布特征，构建合适的迁移算法，提升模型

的适应性和泛化能力，更好地提高检测性能；同时，对小样本攻击数据的检测也需要进一步研究。

参考文献：

[1] Yang An,Sun Limin,Wang Xiaoshan,et al.Intrusion detection techniques for industrial control systems[J].Journal of Computer Research and Development,2016,53(9):2039–2054.[杨安,孙利民,王小山,等.工业控制系统入侵检测技术综述[J].计算机研究与发展,2016,53(9):2039–2054.]

[2] O'Mahony N,Campbell S,Carvalho A,et al.Deep learning vs.traditional computer vision[M]//CVC 2019:Advances in Computer Vision.Cham:Springer,2020:128–144.

[3] Guo Jian,He He,He Tong,et al.GluonCV and GluonNLP:Deep learning in computer vision and natural language processing[J].Journal of Machine Learning Research,2020,21(23):1–7.

[4] Zhang Zixing,Geiger J,Pohjalainen J,et al.Deep learning for environmentally robust speech recognition[J].ACM Transactions on Intelligent Systems and Technology,2018,9(5):1–28.

[5] Liu Junjiao,Yin Libo,Hu Yan,et al.A novel intrusion detection algorithm for industrial control systems based on CNN and process state transition[C]//Proceedings of the 2018 IEEE 37th International Performance Computing and Communications Conference.Orlando:IEEE,2018:1–8.

[6] Mirza A H,Cosan S.Computer network intrusion detection using sequential LSTM Neural Networks autoencoders[C]//Proceedings of the 2018 26th Signal Processing and Communications Applications Conference(SIU).Izmir:IEEE,2018:1–4.

[7] Shi Leyi,Zhu Hongqiang,Liu Yihao,et al.Intrusion detection of industrial control system based on correlation information entropy and CNN–BiLSTM[J].Journal of Computer Research and Development,2019,56(11):2330–2338.[石乐义,朱红强,刘祎豪,等.基于相关信息熵和CNN–BiLSTM的工业控制系统入侵检测[J].计算机研究与发展,2019,56(11):2330–2338.]

[8] Chawla A,Lee B,Fallon S,et al.Host based intrusion detection system with combined CNN/RNN model[M]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases.Cham:Springer,2018:149–158.

[9] Yan Yu,Qi Lin,Wang Jie,et al.A network intrusion detection method based on stacked autoencoder and LSTM[C]//Proceedings of the 2020 IEEE International Conference on Communications(ICC 2020).Dublin:IEEE,2020:1–6.

[10] Mathew A,Mathew J,Govind M,et al.An improved transfer learning approach for intrusion detection[J].Procedia Computer Science,2017,115:251–257.

[11] Niu Shuteng,Liu Yongxin,Wang Jian,et al.A decade survey of transfer learning(2010—2020)[J].IEEE Transactions on Artificial Intelligence,2020,1(2):151–166.

[12] Bai Shaojie,Kolter J Z,Koltun V.Trellis networks for sequence modeling[EB/OL].[2021–09–01].https://doi.org/10.48550/arXiv.1810.06682.

[13] Zaman S,Karray F.Features selection for intrusion detection systems based on support vector machines[C]//Proceedings of the 2009 6th IEEE Consumer Communications and Networking Conference.Las Vegas:IEEE,2009:1–8.

[14] Chen Zhiguo,Kim S R.Combining principal component analysis,decision tree and naïve Bayesian algorithm for adaptive intrusion detection[C]//RACS'13:Proceedings of the 2013 Research in Adaptive and Convergent Systems.New York:ACM,2013:312–316.

[15] Tang Chenghua,Liu Pengcheng,Tang Shensheng,et al.Anomaly intrusion behavior detection based on fuzzy clustering and features selection[J].Journal of Computer Research and Development,2015,52(3):718–728.[唐成华,刘鹏程,汤申生,等.基于特征选择的模糊聚类异常入侵行为检测[J].计算机研究与发展,2015,52(3):718–728.]

[16] Jadhav S,He Hongmei,Jenkins K.Information gain directed genetic algorithm wrapper feature selection for credit rating[J].Applied Soft Computing,2018,69:541–553.

[17] 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.

[18] Lee C,Lee G G.Information gain and divergence-based feature selection for machine learning-based text categorization[J].Information Processing & Management,2006,42(1):155–165.

[19] Wold S,Esbensen K,Geladi P.Principal component analysis[J].Chemometrics and Intelligent Laboratory Systems,1987,2(1/2/3):37–52.

[20] Bai Shaojie,Kolter J Z,Koltun V.An empirical evaluation of generic convolutional and recurrent networks[EB/OL].[2021–09–01].https://arxiv.org/pdf/1803.01271.pdf.

[21] Zhuang Fuzhen,Qi Zhiyuan,Duan Keyu,et al.A comprehensive survey on transfer learning[J].Proceedings of the IEEE,2021,109(1):43–76.

[22] Morris T,Gao Wei.Industrial control system traffic data sets for intrusion detection research[M]//Critical Infrastructure Protection Ⅷ.Berlin:Springer,2014:65–78.

[23] Yu Bangbing,Wang Huazhong,Yan Bingyong.Intrusion de-

tection of industrial control system based on long short term memory[J].Information and Control,2018,47(1):54–59.[於帮兵,王华忠,颜秉勇.基于长短时记忆网络的工业控制系统入侵检测[J].信息与控制,2018,47(1):54–59.]

[24] Demertzis K,Iliadis L,Spartalis S.A spiking one-class anomaly detection framework for cyber-security on industrial control systems[M]//Engineering Applications of Neural Networks.Cham:Springer,2017:122–134.

[25] Vinayakumar R,Soman K P,Poornachandran P.Applying convolutional neural network for network intrusion detection[C]//Proceedings of the 2017 International Conference on Advances in Computing,Communications and Informatics(ICACCI).Udupi:IEEE,2017:1222–1228.

[26] Li Beibei,Song Jiarui,Du Qingyun,et al.DRL–IDS:Deep reinforcement learning based intrusion detection system for industrial Internet of Things[J].Computer Science,2021,48(7):47–54.[李贝贝,宋佳芮,杜卿芸,等.DRL–IDS:基于深度强化学习的工业物联网入侵检测系统[J].计算机科学,2021,48(7):47–54.]

[27] Zhang Ziying,Pan Sichen,Wang Yuhua.Intrusion detection of industrial control system based on FKPCA–HCIPSO–OCSVM[J/OL].Journal of Harbin Engineering University [2022–07–01].http://kns.cnki.net/kcms/detail/23.1390.U.20220401.1711.012.html.[张子迎,潘思辰,王宇华.基于FKPCA–HCIPSO–OCSVM的工业控制系统入侵检测[J/OL].哈尔滨工程大学学报[2022–07–01].http://kns.cnki.net/kcms/detail/23.1390.U.20220401.1711.012.html.]

（编辑　赵　婧）

引用格式：Shi Leyi,Hou Huiwen,Xu Xinghua,et al.Industrial control system intrusion detection based on feature selection and temporal convolutional network[J].Advanced Engineering Sciences,2022,54(6):238–247.[石乐义,侯会文,徐兴华,等.基于特征选择和时间卷积网络的工业控制系统入侵检测[J].工程科学与技术,2022,54(6):238–247.]