计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2020,29(9):47−56 [doi: 10.15888/j.cnki.csa.007620]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于深度学习的工业物联网智能入侵检测①

胡向东, 周　巧

(重庆邮电大学 自动化学院, 重庆 400065)
通讯作者: 周　巧, E-mail: 1397195833@qq.com

摘　要: 如何有效识别工业物联网入侵攻击行为是一个新挑战. 针对工业物联网中入侵检测特征提取不高、检测效率低、适应能力差等问题, 提出一种基于深度学习的工业物联网智能入侵检测方法. 首先, 在数据处理上改进采样算法用于调节少数类别样本数量, 提高检测精度; 其次, 构建堆叠降噪卷积自编码网络提取关键特征, 结合卷积神经网络和降噪自编码器, 加强特征识别能力; 为了避免信息丢失和信息模糊, 改进池化操作以增加其自适应处理能力, 并在模型训练过程中采用 Adam 算法获取最优参数; 最后, 采用 NSL-KDD 数据集测试提出方法的性能. 实验结果表明, 该方法相比现有的 RNN、DBN 和 IDABCNN 的准确率分别提高了 3.66%、4.93% 和 4.6%; 与未经采样算法的 SDCAENN 试验对比, U2R 和 R2L 的检测精度分别提高 17.57% 和 3.28%.

关键词: 工业物联网; 入侵检测; 自适应采样算法; 堆叠卷积自编码; Adam 算法

引用格式: 胡向东,周巧.基于深度学习的工业物联网智能入侵检测.计算机系统应用,2020,29(9):47–56. http://www.c-s-a.org.cn/1003-3254/7620.html

IIoT Intelligent Intrusion Detection Based on Deep Learning

HU Xiang-Dong, ZHOU Qiao

(College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: How to effectively identify the intrusion attack behavior of the Industrial Internet of Things (IIOT) is a new challenge. Aiming at the problems of low intrusion detection feature extraction, low detection efficiency, and poor adaptability in IIOT, an intelligent intrusion detection method based on deep learning is proposed. First, improve the sampling algorithm in data processing for adjusting the number of samples in a few categories to improve the detection accuracy. Second, build a stacked denoising convolutional self-encoding network to extract key features. Combine the convolutional neural network and the denoising self-encoder to enhance feature recognition ability. In order to avoid information loss and information ambiguity, improve the pooling operation to increase its adaptive processing ability, and use Adam algorithm to obtain the optimal parameters during model training. Finally, use the NSL-KDD dataset to test the performance of the proposed method. Experimental results show that the accuracy of the method is 3.66%, 4.93%, and 0.04% higher than the existing RNN, DBN, and IDMBCNN, respectively. Compared with the SDCAENN test without sampling algorithm, the detection accuracy of U2R and R2L is improved by 17.57 % and 3.28%.

Key words: Industrial Internet of Things (IIoT); intrusion detection; adaptive sampling algorithm; stacked convolution self-coding; Adam algorithm

① 基金项目: 重庆市教委科学研究项目 (KJ1602201); 教育部-中国移动联合基金 (MCM20150202)
Foundation item: Scientific Research Program of Chongqing Municipal Education Commission (KJ1602201); Joint Fund of MOE and China Mobile (MCM20150202)
收稿时间: 2020-02-18; 修改时间: 2020-03-17; 采用时间: 2020-04-07; csa 在线出版时间: 2020-09-04

随着工业物联网应用程度的加深, 复杂的网络环境和层出不穷的攻击手段使得其面临许多挑战, 诸如黑客入侵、安全漏洞攻击、蠕虫[1] 等. 工业物联网可分为感知层、网络层、应用层, 其感知层安全需求主要致力于保障数据安全, 表现为防止恶意节点攻击、采集样本与节点数据伪造破坏等. 网络层安全表现为阻止 Dos 攻击, 保证路由安全. 应用层安全则满足用户隐私和访问控制等. 目前针对工业物联网的安全机制大多偏向被动防御, 而入侵检测 (Intrusion Detection, ID) 是可以在不影响网络内部的情况下, 对网络的传输数据进行实时监控并采取措施对入侵行为进行监测、分析、预警等处理, 从而提高网络应对外部威胁的能力. 传统的基于深度学习的入侵检测研究不完善, 仍然存在以下问题.

(1) 工业物联网环境复杂, 采集的网络流量数据是高维度的, 目前许多的入侵检测模型手动选取特征, 不够有效且依据较少, 可能会丢失重要特征而保留冗余特征.

(2) 自适应能力差. 随着工业物联网运行环境和结构变化, 要想检测出现的新型未知攻击, 需要不断更新模型.

(3) 低频攻击检测困难. 在实际的网络环境中, 不同类型的流量数据是不平衡的, 这使得分类器偏向于数量大的类, 少数类的攻击检测难度大且检测率不高.

(4) 模型拟合能力差. 传统机器学习模型结构简单, 特征提取及学习能力有限, 当面临大规模数据集时无法对数据分布形成有效的非线性映射.

考虑到上述因素, 本文提出堆叠降噪卷积自编码神经网络 (SDCAENN) 入侵检测模型, 将降噪自编码与卷积神经网络结合, 利用卷积神经网络 (Convolutional Neural Network, CNN) 的卷积特性, 充分学习网络特征, 通过重构误差来求解模型, 同时采用在随机梯度下降法 (Stochastic Gradient Descent, SGD) 改进下的自适应算法 (Adam) 优化网络; 此外增加卷积自编码网络的池化自适应能力, 使其尽可能地学习入侵特征; 并针对低频攻击检测困难问题, 对数据进行区域自适应过采样操作, 从算法角度平衡数据, 再转化为二维灰度图像进入深度卷积自编码网络进行训练.

1 相关工作

随着互联网和工业物联网等的飞速发展, 网络安全成为信息安全国防系统基础设施中的重要部分. 为了检测恶意入侵, 将深度学习应用在网络安全领域. 文献 [2] 采用贪婪多层深度置信网络 (DBN) 模型, 首先利用受限玻尔兹曼机 (RBM) 消除噪声和异常数据对网络的负面影响, 然后采用反向传播 (BP) 算法来微调 DBN 实现分类任务. 文献 [3] 使用深度自动编码 (DAE) 模型, 其中前一层每个自动编码器的输出用作下一层的输入, 以逐层贪婪的分层方式进行训练, 以避免过度拟合和局部最优. 文献 [4] 中, 张宝安提出基于栈式稀疏自编码网络并结合分阶段抽样算法的集成学习, 将多分类集成学习加权融合, 在入侵病毒初期就能有很好的检测. 文献 [5] 改进卷积神经网络, 与传统的“卷积-池化-全连接”结构不同, 采用跨层聚合网络的设计, 将两层卷积-池化-全连接聚合输出到分类决策, 具有较高的准确率. 文献 [6] 提出了基于层次化时空特征学习的网络流量异常检测方法应用在工业物联网, 结合 CNN 和 LSTM, 取得高检测率和低误警率. 文献 [7] 提出了一种基于深度学习模型的 IICS 异常检测技术, 该模型可以使用从 TCP/IP 数据包收集的信息进行学习和验证, 它包括自动编码器和深度前馈神经网络训练过程和实验. 文献 [8] 使用 BiLSTM-RNN 检测工业物联网攻击. 使用新型 UNSWNB15 数据集对多层深度神经网络进行了训练, BiLSTM 模型在攻击检测中达到了 95% 以上的准确率.

分析这些研究发现, 由于存在大量冗余和噪声等干扰, 占用系统资源, 现有工业物联网入侵检测仍然存在检测时间长、准确率低、时效性差等问题. 因此, 提出一种堆叠降噪卷积自编码入侵检测模型, 以改善这些问题.

2 基于深度学习的工业物联网入侵检测模型

本文的模型框架是一种基于堆叠降噪卷积自编码神经网络的入侵检测模型, 该模型的总体框架如图 1 所示.

2.1 入侵检测模型总体架构

由图 1 可知, 该模型对工业物联网入侵检测的识别主要有以下 3 个步骤:

(1) 数据预处理. 搭建工业物联网环境, 利用数据包捕获实时网络数据, 包含源地址、目标地址、连接属性等相关信息. 对其进行预处理转化为构建的堆叠

降噪卷积自编码器可以处理的格式. 本文中数据预处理分为 3 个部分.

① 属性映射, 将 protocol-type、service 和 flag 等字符型数据转换为数值型数据.

② 数据归一化, 将数据归一化到 [0,1] 区间内, 以消除网络连接中不同特征取值量纲过大对入侵检测模型的训练造成影响.

③ 区域自适应过采样算法 (RASmote), 对于少数类样本, 在算法层面上生成新样本, 适当处理数据分布不平衡问题, 再进行下一步操作, 可以优化少数类数据.

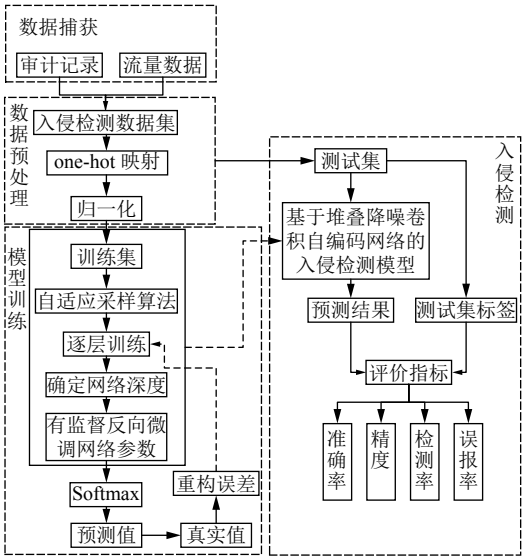

图 1 工业物联网入侵检测模型

(2) 通过建立 SDCAENN 模型, 训练网络模型, 提取分析特征, 结合卷积神经网络和自编码器特性, 对标准数据集进行预训练和调参处理, 实现标准数据特征的最优提取.

① 输入层: 为后续神经网络做准备, 将数据集映射为二维灰度图像处理格式.

② 隐藏层: 由卷积层、池化层和全连接层的编码和解码构成. 其中卷积层激活函数采用 ReLU, 自主学习特征信息, 对池化层进行改进, 全连接层引入 Dropout 方法, 防止由于训练集不足或过度训练造成过拟合.

(3) 决策输出. 通过 Softmax 分类器输出分类决策, 其中 Softmax 权重参数可以与神经网络一起反向传播微调得出.

2.2 堆叠降噪卷积自编码网络

结合降噪自编码和 CNN 提出堆叠降噪卷积自编码神经网络 (Stacking Denoising Convolutional Auto-Encoder Neural Network, SDCAENN), 降噪自编码通过加入卷积系列操作, 实现局部感受野和权值共享, 能更好的解决工业物联网中各类数据信息冗余失真等问题, 有效提高检测率.

由于卷积层与池化层交替设置的网络结构, 池化操作频繁, 会使特征信息模糊, 可能造成不能正确描述入侵, 因此本文改进卷积层与池化层的结构, 使得每经过两个卷积操作进行一次池化, 加强网络的学习能力. 整体网络结构如图 2 所示.

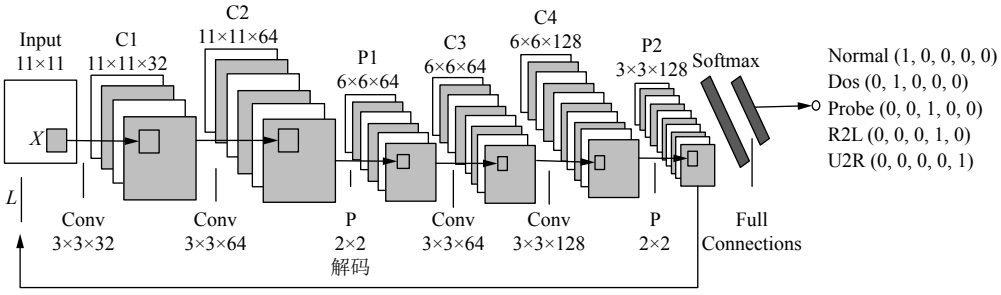

图 2 卷积自编码模型

该神经网络由二个卷积自编码构成. 输入经过两次卷积操作之后得到特征 C2, C2 池化后得到 P1 作为第二个卷积自编码的输入, 第二个卷积自编码通过卷积得到特征 C4 并池化输出 P2. P2 输入两个全连接层 FC1, FC2, 结果作为输出层 Softmax 的输入, 训练得到五分类.

2.2.1 卷积自编码网络

卷积自编码网络结构如图 3 所示, 详细的编码解码过程推导如下:

(1) 编码过程

卷积层输出可以表示为:

$$h_1 = S_f\left(x \otimes W'_{11} + b'_{11}\right) \tag{1}$$

式中, x 表示输入特征向量, $\otimes$为卷积操作, W'_{11}表示第 1 层权重, b'_{11}表示第 1 层偏置, 而S_f为非线性激活函数, 常见的有 Sigmoid, Tanh, ReLU, 由于 ReLU 相较于其他激活函数可以使网络更快收敛, 减小训练时间, 因此本文采取 ReLU 激活函数, 即:

$$f(x)_{\text{relu}} = \begin{cases} 0\ (x \le 0) \\ x\ (x > 0) \end{cases} \tag{2}$$

池化层层输出可以表示为:

$$h_2 = pool(h_1) = S_f\left(down(x) + b'_{11}\right) \tag{3}$$

其中, $pool$表示池化操作, $down(\cdot)$表示下采样.

(1) 解码过程

$$h'_2 = S_g\left(h_2 \otimes W'_{22} + b'_{22}\right) \tag{4}$$

$$h'_1 = upsample\left(h'_2\right) = S_g\left(upsample(x) + b'_{22}\right) \tag{5}$$

$$x' = S_g\left(h'_1 \otimes W'_{21} + b'_{21}\right) \tag{6}$$

式中, x'为重构后的x, W'_{22}和b'_{22}为解码时第 1 层卷积的权重和偏置. W'_{21}和b'_{21}为解码时第 2 层卷积的权重和偏置, h'_2为解码卷积输出, $upsample$为上采样, h'_1为解码池化输出, S_g为解码激活函数, 同编码器中一样.

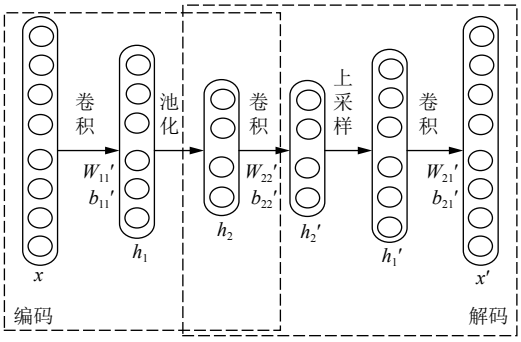

图 3 卷积自编码网络结构

2.2.2 卷积自编码网络的训练

SDCAENN 的训练流程如图 4 所示.

(1) 前向传播

① 从标准数据集随机选取 batch 输入, 输入数据参数维度 ($batch\_size$,h,w,c).

② 输入图 2 所示卷积降噪自编码神经网络, 卷积运算和池化运算分别如式 (7) 和式 (8).

$$h_j^l = S_f\left(\sum_{i \in M_J} h_j^{l-1} \otimes W'_{ij} + b'_j\right) \tag{7}$$

$$Z_j^l = \beta\left(W_j^l down\left(Z_j^{l-1}\right) + b_j^l\right) \tag{8}$$

③ 利用 Tensorflow 中 conv2d_transpose 函数和 $upsample$ 函数进行反卷积池化解码. 并输入到全连接层输出结果. 全连接层 (FC) 计算如式 (9):

$$y_j^l = S_f\left(\sum_{i \in M_J} y_j^{l-1} \otimes W_{ij}^l + b_j^l\right) \tag{9}$$

④ 求解重构误差, Softmax 决策输出进行数据分类.

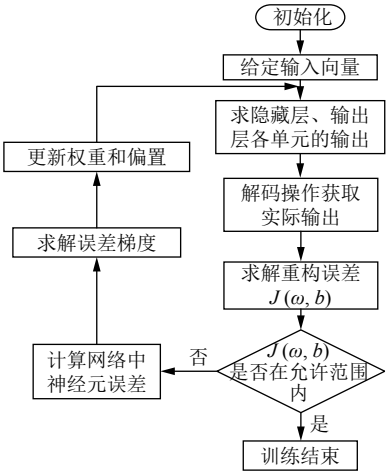

图 4 SDCAENN 模型训练流程图

(2) 反向传播

① 根据训练集样本分类结果, 计算整体损失函数 $J(\omega,b)$.

② 反向传播训练网络的权重和偏置, 直到收敛. 在模型的训练过程中, 为了加快收敛时间和提高收敛精度, 本文使用 Adam 对网络模型参数更新, 该方法解决了收敛速度慢和容易局部最优问题, 并节约了计算机的资源.

本模型的损失函数为:

$$J(\omega,b) = J\left(\omega,b;x^i,y^i\right) + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_i}\sum_{j=1}^{s_j+1}\left(\omega_{ji}^l\right)^2 \tag{10}$$

$$J\left(\omega,b;x^i,y^i\right) = \frac{1}{N}\sum_{i=1}^{N} y\ln a + (1-y)\ln(1-a) \tag{11}$$

其中, $J\left(\omega,b;x^i,y^i\right)$为交叉熵损失函数, 可以减少不平衡数据中不同分类难度差异问题. $\frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_i}\sum_{j=1}^{s_j+1}\left(\omega_{ji}^l\right)^2$为正则化项, 即权重衰减, 其目的是减小权重幅度, 防止训练过拟合. 其中$a=\sigma(h)$, $h=\omega*x+b$. 激活函数 Sigmoid

及其导数分别如式 (12) 和式 (13) 所示:

$$\sigma(h)=\frac{1}{1+e^{-h}} \tag{12}$$

$$\sigma'(h)=\frac{e^{-h}}{(1+e^{-h})^2}=\sigma(h)(1-\sigma'(h)) \tag{13}$$

交叉熵对权重和偏置的求导如下:

$$\frac{\partial J\left(\omega,b;x^i,y^i\right)}{\partial \omega_j}=\frac{1}{N}\sum_{i=1}^{N}\frac{\sigma'_{\omega_j}(h)x_j}{\sigma(h)(1-\sigma(h))}(\sigma(h)-y) \tag{14}$$

$$\frac{\partial J\left(\omega,b;x^i,y^i\right)}{\partial b_j}=\frac{1}{N}\sum_{i=1}^{N}\frac{\sigma'_{b_i}(h)}{\sigma(h)(1-\sigma(h))}(\sigma(h)-y) \tag{15}$$

代入式 (13) 可得:

$$\frac{\partial J\left(\omega,b;x^i,y^i\right)}{\partial \omega_j}=\frac{1}{N}\sum_{i=1}^{N}x_j(\sigma(h)-y) \tag{16}$$

$$\frac{\partial J\left(\omega,b;x^i,y^i\right)}{\partial b_j}=\frac{1}{N}\sum_{i=1}^{N}(\sigma(h)-y) \tag{17}$$

本文采用 Adam 优化算法对权重和偏置更新. 算法如算法 1 所示.

算法1. Adam算法

1. Require: 步长α;
2. Require: 矩估计的指数衰减速率$\beta_1,\beta_2\in(0,1)$;
3. Require: 有参数的损失函数$J(\theta)$;
4. Require: 初始参数θ;
5. 一阶和二阶矩变量进行初始化$s=0,\gamma=0$, 始化时间同步$t=0$;
6. While没有达到停止准则
7. 从训练集中采取N样本$\left\{x^1,x^2,\cdots,x^n\right\}$, 对应目标为$y^i$;
8. 计算梯度: $g\leftarrow\frac{1}{N}\nabla_\theta\sum_i L\left(f\left(x^i;\theta\right),y^i\right)$;
9. $t\leftarrow t-1$;
10. 更新偏一阶矩阵$S_t\leftarrow\beta_1 S_{t-1}+(1-\beta_1)g$;
11. 更新偏二阶矩阵: $\gamma_t\leftarrow\beta_2\gamma_{t-1}+(1-\beta_2)g^2$;
12. 修正偏一阶矩阵$\hat{s}_t\leftarrow\frac{s_t}{1-{\beta_1}^t}$;
13. 修正偏二阶矩阵$\hat{\gamma}_t\leftarrow\frac{\gamma_t}{1-{\beta_2}^t}$;
14. 更新参数$\theta_t\leftarrow\theta_{t-1}-\alpha\frac{\hat{s}_t}{\sqrt{\hat{\gamma}^t+\xi}}$;
15. 应用更新$\theta\leftarrow\theta+\Delta\theta$.

一般情况下, $\alpha=0.001$、$\beta_1=0.9$、$\beta_2=0.999$、$\xi=10^{-8}$.

2.2.3 改进的池化方法

池化技术可以对特征进行缩放、位移等, 且保持特征不变, 可以减少网络负担, 筛选冗余特征. 传统的池化方法有平均池化 (Average Pooling) 和最大池化 (Max Pooling). 平均池化选取池化核平均值, 可能会有较强信息被弱化, 同理最大池化也可能使得关键信息丢失, 因此, 本文提出式 (18) 的自适应池化算法. 可以对池化核的不同元素动态分配合适的池化权值, 能更好的表达特征信息.

自适应池化:

$$S_{ij}=\mu_{ij}\max_{i=1,j=1}^{c}(F_{ij})+b \tag{18}$$

其中, μ_{ij}为池化因子, $\mu_{ij}=\frac{2}{c^2\left(1+e^{\frac{-F_{ij}}{F_{\text{sum}}}\sigma^2}\right)}$.

F_{ij}表示卷积特征 F 中池化核$c*c$对应的元素, c表示当前池化核的大小, F_{sum}表示池化核所有元素和, σ是标准差. 自适应池化算法克服了最大池化、平均池化片面性问题, 可以获取更为准确的信息.

3 实验及结果分析

为了验证本文提出的入侵检测方法在工业物联网背景下的优势, 对本文入侵检测模型进行仿真, 设计评价指标来对性能进行测试.

3.1 实验环境及参数选择

本实验使用 Tensorflow 来进行实验模拟, 选择 Python 编程语言. 计算机硬件配置为 Inter(R) Core(TM) i7-6700CPU@3.40 GHz 处理器, 8 GB 内存, 操作系统为 64 位 Windows10 系统.

在模型中主要的参数变量包含卷积自编码网络结构参数、学习率、连接概率和训练次数等. 参数的具体数值如表 1 所示.

3.2 实验数据来源

目前, 工业物联网入侵的公共数据集主要有 KDDCup99, NSL-KDD[9], GasPipeline Datasets, Water Datasets, UNSW-NB15 等, 这些数据集存在数据和属性冗余重复等问题, 本文选用 NSL-KDD 数据集作为实验基准数据. 它解决了 KDDCup99 数据集冗余数据的问题, 其原始训练集 KDDTrain 包含 125 973 条数据, 原始测试集 KDDTest 包含 22 544 条数据, 本文选用 KDDTrain+20% 的 25 192 条数据作为实验数据. 数据集中的每一行数据都有 41 个特征属性和 1 个标签属性, 其中主要包括 4 种类型的攻击: Dos (拒绝服务攻

击)、Probe (端口漏洞扫描攻击)、R2L (远程非法访问攻击)、U2R (越权访问攻击). NSL-KDD 训练数据集包含 22 种攻击, 测试数据集包含 17 种攻击, 具体分布占比如表 2 所示.

3.3 数据预处理

NSL-KDD 数据集中包含 41 个特征属性, 其中包括符号型特征 (tcp,udp,icmp,…) 和数值型特征, 需要将数据进行标准化预处理才能应用到检测算法之中.

(1) 字符型映射数值型

“0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0,0,0,0,1,0,0,150,25,0.17,0.03,0.17,0,0,0,0.05,0,Normal”是该数据集中的一条数据, 分析可知, 数据的第 2,3,4 维数值是字符类型, 需要转化为数值类型, 例如第 2 维有 (tcp,udp,icmp)3 种类型, 第 3 维有 (‘auth’, ‘bgp’,‘courier’等)70 种类型, 第 4 维有 (‘OTH’,‘REJ’, ‘RSTO’等)11 种类型, 按照图 5 的 one-hot 编码来处理, 最终将 41 维转化为 122 维属性.

表 1 实验变量参数

变量名称	变量值
输入层	Input_data=(11,11,1)
卷积层1	Conv1=(3,3,32), ReLU
卷积层2	Conv2=(3,3,64), ReLU
池化层1	Pool1=Adaptive pooling
卷积层3	Conv3=(3,3,64), ReLU
卷积层4	Conv4=(3,3,128), ReLU
池化层2	Pool2=Adaptive pooling
上采样层U1	Upsample = (6,6)
卷积层5	Conv5=(3,3,32), ReLU
卷积层6	Conv6=(3,3,64), ReLU
池化层3	Pool3=Adaptive pooling
卷积层7	Conv7=(3,3,64), ReLU
卷积层8	Conv8=(3,3,128), ReLU
池化层4	Pool4=Adaptive pooling
全连接层	节点数512, ReLU
输出层	节点数5, Softmax
Dropout	$p = 0.6$
学习率	$\varepsilon = 0.001$
训练轮数	Epochs=130

表 2 NSL-KDD 数据集占比情况

数据集	数量	Normal		Dos		Probel		R2L		U2R	
		数据量	占比(%)	数据量	占比(%)	数据量	占比(%)	数据量	占比(%)	数据量	占比(%)
KDDTrain+	125 973	67 343	53.46	45 927	36.46	11 656	9.25	995	0.79	52	0.04
KDDTest+	22 544	9711	43.08	7460	31.79	2421	12.04	2885	12.80	67	0.30
KDDTrain20%	25 192	13 449	53.39	9234	36.65	2289	9.09	209	0.83	11	0.04
KDDTest50%	11 850	2152	18.15	4244	36.66	2402	20.27	2885	24.35	67	0.57

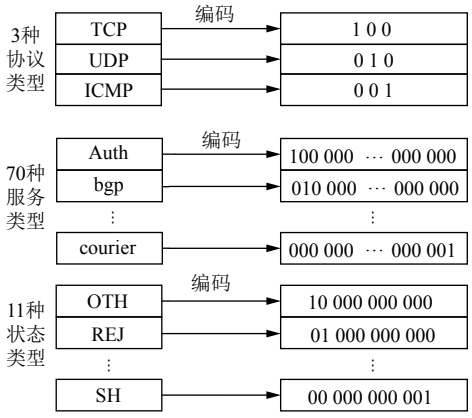

图 5 One-hot 编码数值化

(2) 数值归一化

不同的特征属性其数据量纲和对应取值范围都有明显的差异, 为了方便实验结果分析, 采用 Min-Max 标准化方法将数值型数据统一映射到 [0,1] 区间, 使得数据处于同一量级.

$$X_{\text{normal}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{19}$$

其中, x表示样本特征原始值, $x_{\min}$, $x_{\max}$分别表示该条数据中的最小值和最大值, X_{normal} 表示每条数据归一化后新特征值.

(3) 低频样本处理

尽管当前工业物联网攻击呈快速增长的趋势, 但攻击类别以及个别攻击类别相较于正常数据流量仍然属于低频范畴, 导致难以捕捉其特征记录, 还由于大多数人工智能模型以最大样本整体分类准确率为目的, 因此具有明显的分类偏向性. 因此本文改进采样算法, 引入区域自适应合成过采样算法 (RASmote) 增量处理低频样本, 算法公式如下:

$$K = |X_r - X_{mi}| = \sqrt{\sum_{i=1}^{n} (X_{ri} - X_{mii})^2} \tag{20}$$

利用欧氏距离计算最近邻半径内低频样本距离. r 为最近邻半径, X_r为最近邻样本集合, X_{mi}为低频样本,

X_{new}为新样本集合. $X \in X_{mi}$

$$K = r, X \in IPR, X_{\text{new}} = X \tag{21}$$

$$0 \le K \le \frac{r}{2},\ X \in SPR,\ X_{\text{new}} = 0 \tag{22}$$

$$\frac{r}{2} < K < r, X \in DPR, X_{\text{new}} = X + \mu(0,1)\left(\left(\frac{1}{r-k}\sum_{i=1}^{r-k} X_i\right) - X\right) \tag{23}$$

其中, $\left(\frac{1}{r-k}\sum_{i=1}^{r-k} X_i\right)$为低频样本.

3.4 评价指标

判断入侵检测模型的性能可以从模型对比和分类检测两方面考虑, 模型的对比主要与传统的入侵检测技术进行对比, 具体见 3.5.3 节. 系统检测准确性主要指标有准确率、精确率、检测率等, 本文采取混淆矩阵度量实验结果, 如表 3 所示.

表 3 入侵检测混淆矩阵

实际	预测	
	Normal	Attack
Normal	*TN*	*FP*
Attack	*FN*	*TP*

则评价指标分别如下:

准确率 (Accuracy, *ACC*): 分类正确的样本数与样本总数之比.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{24}$$

精确率 (Precision, *P*): 正确判定为入侵/正常的数据占预测为入侵/正常数据的总数.

$$P = \frac{TP}{TP + FP} \tag{25}$$

检测率/召回率 (Recall, *R*): 正确判定为入侵/正常的数据占实际为入侵/正常数据的总数.

$$R = \frac{TP}{TP + FN} \tag{26}$$

误报率 (False Alarm Rate, *FAR*): 正常数据被预测错误的个数占正常数据总数.

$$FAR = \frac{FP}{FP + TN} \tag{27}$$

*F*1-*Score*: 该指标是 Precision 与 Recall 的调和平均.

$$F1 - Score = \frac{2 \times P \times R}{P + R} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{28}$$

3.5 结果分析

3.5.1 RASmote 算法对检测率影响

为了验证 RASmote 算法的有效性, 实验将 RASmote 处理前后的数据集在本文模型上进行验证. 实验表明采样率为 60%, 最近邻半径 *r*=55 时, 获取的新数据集分布最均衡, 效果最好. 实验结果如表 4 所示.

表 4 采样算法比较试验结果 (%)

样本类别	评价指标	SCAE	RASmote-SDCAENN
Normal	精确率	97.14	96.91
	检测率	98.37	98.22
Dos	精确率	98.05	97.92
	检测率	97.23	95.27
Probel	精确率	84.74	92.43
	检测率	78.92	93.27
U2R	精确率	12.87	30.44
	检测率	37.25	63.63
R2L	精确率	64.34	67.62
	检测率	47.56	78.95

由表可知, 经过 RASmote 处理的数据比未经其处理的数据在检测率和精确率上明显提高, 其中 Normal 因为数量多检测率已经较高, 因此变化微小, 还可以得出, 其中 Dos 的精确率下降 0.13%, 但是稀少类样本 U2R 的精确率提高 17.57%, 检测率提高 26.38%. R2L 的精确率提高 3.28%, 检测率提高 31.39%. 结果证明, 采用自适应采样算法可以适当平衡数据, 提高稀少类的检测性能.

3.5.2 网络结构性能分析

本节针对提出的神经网络结构进行实验和测试, 采用 KDDTrain+20% 数据作为本次实验数据, 取 70% 作为训练集, 30% 作为测试集, 数据分配如表 5 所示.

表 5 测试集和训练集分布情况

类型	训练集	测试集
Normal	9396	4053
Dos	6500	2734
Probe	1588	701
U2R	5	6
R2L	145	64
Total	17634	7558

由于卷积核过大会导致模型训练复杂, 降低检测性能, 卷积核过小容易导致特征学习不完整, 因此本文选取卷积核大小 3×3. 为了检测提出的堆叠卷积自编码入侵检测模型的性能, 本文设置 3 组卷积自编码网络结构来做实验对比, 分别为 1 组卷积+池化, 2 组卷

积+卷积+池化, 2 组卷积+池化. 实验准确率和精确率结果如图 6 所示. 并在这 3 种网络结构下每类攻击准确率实验结果如图 7 所示.

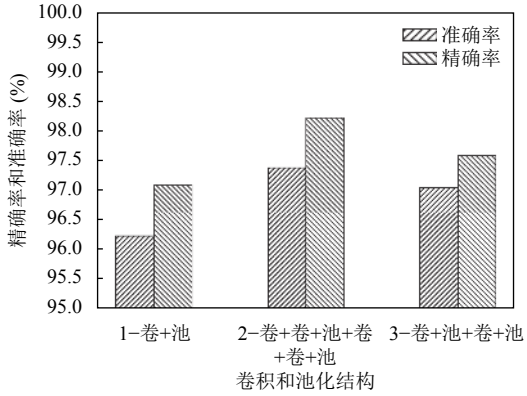

图 6 准确率和精确率对比

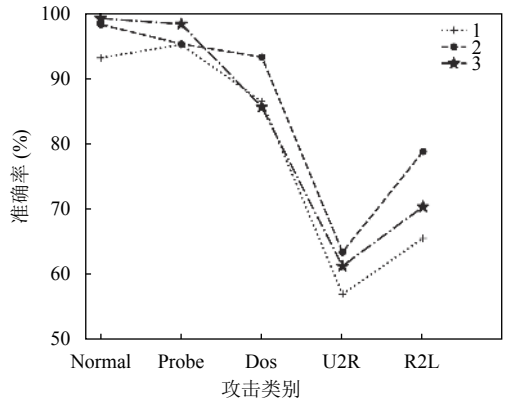

图 7 各类攻击精确率对比

由图 6 可以看出, 在两组卷积卷积池化结构下, 整体准确率和精确率占优势, 且每一类准确率均高于一组卷积池化结构, Normal 和 Probel 的准确率分别低于两组卷积池化结构 0.91% 和 3.01%. 因此两组卷积卷积池化结构比其他两组更有优势, 不会造成特征提取不充分, 也不会引起特征模糊稀疏.

本文选取两组卷积卷积池化的结构在进行 130 次迭代后, 其准确率及损失函数变化结果如图 8, 呈稳定趋势且收敛, 因此该结构效果良好, 且可以避免过拟合.

3.5.3 本文模型与其他模型性能比较

为了评估本文模型的性能, 设计实验与常规模型和局部算法进行比较. 将本文模型与传统检测模型 DBN、RNN、DCNN 和 LeNet-5 进行实验对比, 分别得到如表 6 所示结果.

由表 6 可知, 本文提出模型在准确率和检测率上均高于其他 4 种模型, 而误报率均比 3 种模型略高, 较 LeNet-5 低 0.25%, 这说明本文模型对入侵检测数据的识别能力较强.

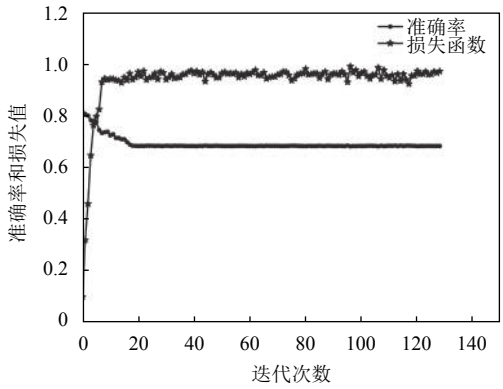

图 8 准确率和损失函数曲线

表 6 与经典网络模型比较 (%)

入侵检测模型	准确率	检测率	误报率
本文	97.38	96.42	1.78
DBN[10]	92.45	95.68	0.85
RNN[11]	93.72	92.70	0.91
DCNN[12]	96.50	92.60	0.97
LeNet-5[13]	86.54	91.31	2.03

此外, 为了更好地验证模型的有效性, 与改进的卷积神经网络入侵检测模型 IDABCNN[14]、NIDMBCNN[15] 和 CNN-Bi- LSTM[16] 进行对比, 均采用 NSL-KDD 数据集训练和测试, 实验结果如表 7 所示.

表 7 与改进的 CNN 模型对比 (%)

入侵检测模型	准确率	检测率	误报率	*F1-Score*
本文	97.38	96.42	1.78	96.90
文献[14]	92.78	90.61	0.95	91.56
文献[15]	97.34	91.33	0.82	94.24
文献[16]	94.33	73.24	6.49	82.45

由表 8 可知, 本文提出的模型在准确率上比文献 [14] 提高了 4.6%, 比文献 [15] 提高了 0.04%, 比文献 [16] 提高了 3.05%, 检测率分别提高 5.81%、5.09% 和 23.18%. 但是其误报率分别高于文献 [14]、文献 [15] 0.83% 和 0.96%, 低于文献 [16] 4.71%. 综上可以得出, 本文结合卷积神经网络和自编码网络特性, 提出的入侵检测方法能够保证应用到工业物联网入侵检测系统中, 并且训练出的网络模型具有较好的分类检测性能.

为了验证 R2L 和 U2R 分类检测结果, 本文与 CNN 和 SSAE-XGB 作比较, 结果如表 8. 结果表明本文检测

算法在 R2L 的 *F*1-*Score* 相对 CNN 提高了 48.06%, 而比 SSAE-XGB 降低了 0.03%, U2R 的 *F*1-*Score* 相对 CNN 和 SSAE-XGB 分别提高了 19.51% 和 28.78%. 因此可以得出结论, 引入的区域自适应过采样算法使得少数类攻击检测性能提高, 有效改善了工业物联网个别攻击问题.

表 8 少数类检测模型对比 (%)

模型	Class	*P*	*R*	*F*1-*Score*
本文	R2L	67.62	78.95	82.30
	U2R	30.44	63.63	41.18
CNN[17]	R2L	81.34	21.68	34.24
	U2R	65.00	13.00	21.67
SSAE-XGB[18]	R2L	97.18	71.42	82.33
	U2R	12.83	12.00	12.40

3.5.4 模型复杂度分析

时间复杂度即模型的运算次数, 可用浮点运算次数 (FLoating-point OPerations, FLOPs) 衡量, 其决定了模型前项传播训练时间, 如果时间复杂度较大, 会导致消耗大量时间, 无法快速验证预测. 空间复杂度决定了模型的参数数量, 包括总参数量加各层输出特征图. 计算公式分别如下:

时间复杂度:

$$Time \sim \mathrm{O}\left(\sum_{l=1}^{L} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l\right) \tag{29}$$

可见, 时间复杂度由输出特征图面积 M^2、卷积核面积 K^2、输入输出通道数决定. 其中, L 为网络深度, l 为第 l 个卷积层, C_l 第 l 个卷积层的卷积输出通道数.

输出特征图尺寸:

$$M = \frac{(X - K + 2 \times Padding)}{Stride} + 1 \tag{30}$$

空间复杂度:

$$Space \sim \mathrm{O}\left(\sum_{l=1}^{L} K_l^2 \cdot C_{l-1} \cdot C_l\right) \tag{31}$$

空间复杂度只与卷积核尺寸 K、通道数 C、层数 L 相关. 而与输入图片尺寸 M 无关.

文献 [14] 为 3 个卷积 3 个池化, 卷积核大小为 2×2, 文献 [15] 为 3 个卷积 3 个池化, 卷积核大小为 3×3. 由于本文模型使用 Dropout 稀疏网络, 使得网络参数大幅减少, 但其涉及到的卷积操作较多, 根据公式可知, 其空间复杂度低于文献 [14] 和文献 [15], 时间复杂度高于文献 [14] 和文献 [15].

$$\begin{aligned} &Time \sim \mathrm{O}(\text{本文}) > Time \sim \mathrm{O}(15) > Time \sim O(14) \\ &Space \sim \mathrm{O}(15) > Space \sim \mathrm{O}(14) > Space \sim \mathrm{O}(\text{本文}) \end{aligned} \tag{32}$$

为了改进时间复杂度, 可以采用小卷积操作代替大卷积优化, 不仅可以减小网络训练参数, 而且可以在保证有效提取特征的情况下减少训练时间.

4 结果与展望

针对目前工业物联网深度学习入侵检测技术的检测效率低、特征丢失、低频检测率困难、自适应能力差等问题, 本文提出了一种基于区域自适应过采样算法与堆叠降噪卷积自编码结合的入侵检测模型, 与传统的多层自编码网络与卷积神经网络相比, 可以更充分地学习网络特征. 其中 Dropout 正则化避免了过拟合; Adam 优化重构误差, 加快收敛速度, 并避免局部最优; 自适应池化算法, 减少特征丢失和不平衡. 此外, 针对 NSL-KDD 数据集存在比例差别大的问题, 对少数类进行改进的采样算法, 有效提高检测精度. 对比实验结果表明, 本文提出的 RASmote-SDCAENN 模型准确率和检测率明显上升, 分别达到 97.38% 和 96.42%, 误报率稍微降低.

虽然本文方法对解决工业物联网入侵检测有改善, 但仍然有问题尚未解决, 下一步集中关注的问题主要有 3 方面: (1) 如何节省工业物联网节点存储空间, 并保证入侵检测效率; (2) 针对模型训练过程中可能出现的梯度弥散, 局部最优等问题, 考虑用改进的遗传算法等启发式算法来进行参数调优; (3) 目前背景下, 入侵检测数据集众多, 尝试采其他代表性的数据集进行验证试验, 进一步提高泛化能力.

参考文献

1 韩丽, 李孟良, 卓兰, 等. 《工业物联网白皮书 (2017 版)》解读. 信息技术与标准化, 2017, (12): 30–34.

2 Yan BH, Han GD. Effective feature extraction via stacked sparse autoencoder to improve intrusion detection system. IEEE Access, 2018, 6: 41238–41248. [doi: 10.1109/ACCESS.2018.2858277]

3 Hsu CM, Hsieh HY, Prakosa SW, *et al*. Using long-short-term memory based convolutional neural networks for network intrusion detection. Proceedings of the 11th EAI

International Conference on Wireless Internet. Taipei, China. 2018. 15–16.

4 Al-Qatf M, Lasheng Y, Al-Habib M, *et al*. Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access, 2018, 6: 52843–52856. [doi: 10.1109/ACCESS.2018.2869577]

5 张宝安. 基于深度学习的入侵检测研究与实现 [硕士学位论文]. 北京: 北京邮电大学, 2019.

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

7 Al-Hawawreh M, Moustafa N, Sitnikova E. Identification of malicious activities in industrial internet of things based on deep learning models. Journal of Information Security and Applications, 2018, 41: 1–11. [doi: 10.1016/j.jisa.2018.05.002]

8 Roy B, Cheung H. A deep learning approach for intrusion detection in Internet of Things using bi-directional long short-term memory recurrent neural network. Proceedings of 2018 28th International Telecommunication Networks and Applications Conference. Sydney, Australia. 2018. 1–6.

9 Tavallaee M, Bagheri E, Lu W, *et al*. A detailed analysis of the KDD CUP 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. Ottawa, ON, Canada. 2009. 53–58.

10 Chung J, Gülcehre Ç, Cho K, *et al*. Gated feedback recurrent neural networks. Proceedings of the 32nd International Conference on Machine Learning. Lille, France. 2015. 2067–2075.

11 Cui JJ, Long J, Min EX, *et al*. Comparative study of CNN and RNN for deep learning based intrusion detection system. Proceedings of the 4th International Conference on Cloud Computing and Security. Haikou, China. 2018. 159–170.

12 Naseer S, Saleem Y. Enhanced network intrusion detection using deep convolutional neural networks. KSII Transactions on Internet and Information Systems, 2018, 12(10): 5159–5178.

13 LeCun Y, Bottou L, Bengio Y, *et al*. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278–2324. [doi: 10.1109/5.726791]

14 贾凡, 孔令智. 基于卷积神经网络的入侵检测算法. 北京理工大学报, 2017, 37(12): 1271–1275.

15 Wang M, Li J. Network intrusion detection model based on convolutional neural network. Journal of Information Securyity Research, 2017, 3(11): 990–994.

16 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. 计算机研究与发展, 2019, 56(11): 2330–2338. [doi: 10.7544/issn1000-1239.2019.20190376]

17 Qian TY, Wang Y, Zhang MM, *et al*. Intrusion detection algorithm based on convolutional neural network. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2018, 46(1): 6–10.

18 Zhang BA, Yu YH, Li J. Network Intrusion detection based on stacked sparse autoencoder and binary tree ensemble method. Proceedings of 2018 IEEE International Conference on Communications Workshops. Kansas City, MO, USA. 2018. 1–6.