计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2020,29(8):179−184 [doi: 10.15888/j.cnki.csa.007539]

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

基于卷积神经网络的电网工控系统入侵检测算法①

赵智阳[1,2], 夏筱筠[2]

[1](中国科学院大学, 北京 100049)
[2](中国科学院 沈阳计算技术研究所, 沈阳 110168)
通讯作者: 赵智阳, E-mail: zzybry@163.com

摘　要: 传统的电网工控系统主要通过防火墙等工具, 与外部网络进行隔离, 但是随着云计算、物联网等新技术的应用, 网络之间互联程度不断深入, 安全防护难度大大提高, 如何有效检测出网络入侵行为变得至关重要. 与传统入侵检测技术相比, 卷积神经网络具有更好的提取入侵特征的能力. 本文提出一种基于卷积神经网络的电网工控系统入侵检测算法, 使用经过处理的 KDD99 数据集进行模型训练, 并添加级联卷积层优化网络结构. 在参数规模不大的前提下, 保证了模型运行的实时性要求. 本文算法相对于传统 SVM 算法和 K-means 算法, 提高了入侵检测的准确率, 降低了误检率, 可以有效检测出对于电网工控系统的入侵行为.

关键词: 卷积神经网络; 电网工控系统; 入侵检测; SVM; K-means

引用格式: 赵智阳,夏筱筠.基于卷积神经网络的电网工控系统入侵检测算法.计算机系统应用,2020,29(8):179–184. http://www.c-s-a.org.cn/1003-3254/7539.html

Intrusion Detection Algorithm of Power Grid Industrial Control System Based on CNN

ZHAO Zhi-Yang[1,2], XIA Xiao-Jun[2]

[1](University of Chinese Academy of Sciences, Beijing 100049, China)
[2](Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

Abstract: Traditional power grid industrial control systems are mainly isolated from external networks through tools such as firewalls, but with the application of new technologies such as cloud computing and the Internet of Things, the degree of interconnection between networks has continued to deepen, and the difficulty of security protection has greatly increased. How to effectively detect network intrusion behavior has become very important. Compared with traditional intrusion detection technology, convolutional neural networks have a better ability to extract intrusion features. This study proposes a power grid industrial control system intrusion detection algorithm based on convolutional neural networks. The KDD99 dataset is processed for model training, and a cascade convolution layer is added to optimize the network structure. Under the premise of small parameter scale, the real-time requirements of the model are guaranteed. Compared with the traditional SVM algorithm and the k-means algorithm, the intrusion detection accuracy of the proposed algorithm in this study is improved, the false detection rate is reduced, and the intrusion behavior to the power grid industrial control system can be effectively detected.

Key words: Convolutional Neural Network (CNN); grid industrial control system; intrusion detection; SVM; K-means

电网工控系统是控制管理电网的中枢系统, 原本是相对封闭的内部环境, 遭受的网络入侵行为相对较少. 但是随着网络互联在电网系统中的不断应用, 系统智能化加深的同时, 将越来越多的部件暴露在外部网络之中, 面临着越来越多的网络入侵行为. 2010 年, “震网”病毒攻击了伊朗核电站中的计算机控制部件,导致

① 收稿时间: 2020-01-15; 修改时间: 2020-02-13, 2020-02-25; 采用时间: 2020-03-03; csa 在线出版时间: 2020-07-29

离心机转动受到极大影响[1]. 2015 年, 乌克兰遭受两次针对电力系统的网络入侵行为, 造成重大损失[2]. 电网工控系统中控制中心和传感器等设备之间并未采取安全隔离措施, 缺乏入侵检测防护层, 极易造成重大损失[3]. 电网工控系统内部流量主要为内网采集的处于网络边缘位置设备的信息, 系统内部缺乏对网络包的深度解析. 因此, 有效的检测出电网工控系统中的网络入侵行为, 才能保证电网系统的高效安全运转.

入侵检测技术是一种获取并分析网络中的数据, 并判断其为合法访问或网络入侵的技术. 其中, 基于特征的入侵检测将网络行为与已知的攻击行为进行比较, 常用的方法有规则匹配、模式识别等; 基于异常的入侵检测, 首先通过历史数据建立数学模型, 之后将数据包与已建立的数学模型设定的阈值进行比较,若超过安全阈值则视为网络入侵, 常用的方法有支持向量机和机器学习等[4]. 针对电网工控系统的网络入侵主要有: 针对工业网络协议漏洞的攻击; 入侵者通过外联网络进入内网形成入侵, 如未授权访问; 针对环境中广泛采用的无线传输进行攻击等. 本文选取 KDD99 数据集作为实验数据集, 其中入侵数据主要包括 DoS、非法网络探测和未授权访问等 4 种, 其中未授权访问会对电网工控系统造成更大危害.

卷积神经网络由多组神经元组成, 神经元之间的局部连接, 有效的传递了数据特征, 结合权值共享和局部连接等自身优点, 可以简化人工神经网络的训练, 同时使得程序具有较强的鲁棒性[5]. 电网工控系统与变压器、发电站等实际物理设备相联系, 一方面, 系统与实际生活息息相关, 需要保证控制信息实时传递; 另一方面, 系统不能宕机, 因此需要保证实时性和稳定性. 使用卷积神经网络模型, 在经过良好的数据训练之后, 模型稳定且计算快速, 可以在系统层面达到实用要求. 本文设计卷积层参数 (卷积核大小和卷积核个数), 池化层参数 (池化核大小和池化核个数), 网络层数, 并加入级联卷积层优化网络结构. 使用筛选之后的训练集对网络进行训练. 最终进行实验, 与传统的 K-means 聚类和 SVM 方法对比准确率和误报率, 实验结果表明本文算法提高了入侵检测效果.

1 相关工作

Anderson 于 1980 年首次提出“入侵检测”概念, 将未经授权访问和篡改信息等行为归为网络入侵[6]. 早期的入侵检测手段主要通过专家和技术人员建立入侵特征数据库, 进而将网络数据与数据库数据进行比较. 若特征匹配, 则判定为入侵数据; 否则, 判定为正常访问数据[7].

近年来, 随着机器学习的迅猛发展, 越来越多的研究人员将其应用到入侵检测之中, 提出新型入侵检测技术. Tang 等采用一种基于支持向量机的层次异常入侵检测技术, 可以在较少先验知识的前提下, 用于检测新型网络入侵攻击[8]. Hatim 等基于 K-means 聚类提出一种融合的机器学习技术, 并构建一个低延时入侵检测系统[9]. Su 等将层次分类算法和 SVM 融合, 结合特征选取技术和 SVM 算法特性, 消除不重要数据特征, 从而达到缩短训练时间, 同时提高 DoS 和非法探测两种入侵行为的正确率[10]. Raman 指出目前的入侵检测技术在稳定性和能否检测出新型网络入侵之间需要做出取舍, 提出一种基于 Hypergraph 和概率的神经网络, 可以有效解决网络入侵的分类问题, 在算法稳定性和检测正确率上都有所提升[11]. 杨昆朋采用深度信念网络 (DBN) 进行特征学习, 将其与支持向量机相融合, 验证了深度学习在数据集丰富场景中,具有更高的准确率[12]. 卷积神经网络在多个应用领域都有良好表现, 但是网络层次过深, 网络复杂度太高, 最终则不能在实际应用场景中使用, 只有设计符合特定应用场景的网络结构, 才能真正在实际应用中发挥效果[13].

2 算法和网络模型设计

2.1 算法概述

本文完整算法流程如下所示:

Step 1. 筛选 KDD99 中通信协议为 TCP 和 UDP 数据集, 剔除无效数据集, 增加未授权入侵行为数据比例, 使得四种入侵行为所占比例相同;

Step 2. 将字符型数据转化为多维数值向量, 同时使用 PCA 算法, 降低数据维度;

Step 3. 设计卷积核、池化核和全连接层等卷积神经网络参数, 加入级联卷积层优化网络架构;

Step 4. 使用处理后的训练集和测试集训练网络,比较准确率和误检率指标表现, 调整参数使得网络结构最优;

Step 5. 增加 epoch 数量, 当 epoch 数量增加使得准确率有下降趋势时, 停止训练;

Step 6. 将预测后的数据每隔一段时间加入到数据集中, 用于下一步训练. 流程图如图 1 所示.

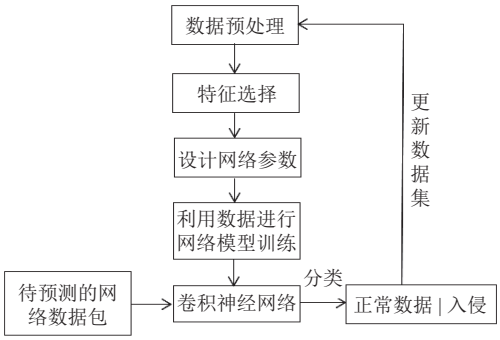

图 1 算法流程

2.2 网络模型设计

神经元是神经网络中基础单元, 数据通过神经元之间的不同权重进行传递, 之后进行输出, 神经元基本结构如图 2 所示.

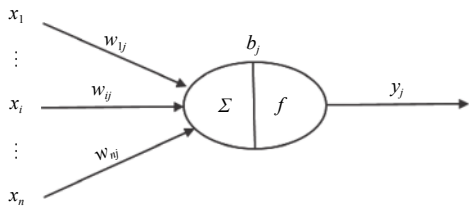

图 2 神经元架构

网络数据流通过神经元被提出数据特征后, 传递到下一个神经元. 经典卷积神经网络架构如图 3 所示.

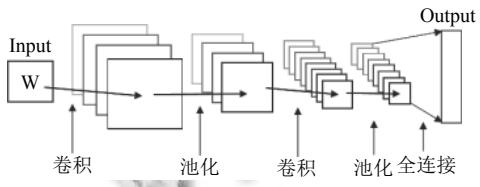

图 3 卷积神经网络

输入 Input 通过中间的卷积层、池化层和全连接层, 最终输出 Output, 本文中 Input 为训练数据, Output 为预测入侵行为和正常行为的概率. 本文神经网络采用的结构为: 4 个卷积层、4 个池化层和 1 个级联卷积层, 另外还有 1 个全连接层, 其中级联卷积层包括 4 个卷积层, 每个卷积层之后相连 1 个激活层, 最终相连 1 个池化层.

输入数据通过卷积层卷积操作之后, 数据不同维度特征将被提取出来, 卷积核一般设置为 $n\times n$ 权值矩阵 (n 一般取值为 3、5 或 7), 本网络采用 3×3 卷积核, 4 层卷积层分别有 4、8、16 和 32 个卷积核. 卷积操作公式如下:

$$x_m^l = b_m^l + \sum_{i=1}^{k} w_{im}^{l-1} * y_i^{l-1} \tag{1}$$

$$y_m^l = f(x_m^l) \tag{2}$$

其中, x_m^l为神经元输入值, b_m^l为偏置值, y_m^l为输出值. 卷积层之后的池化层, 用于对其之前的卷积层特征进行筛选降维, 常用的操作为最大池化和平均池化[14]. 本文神经网络池化核大小为 1×1, 每一层有 8 个池化核. 池化公式如下:

$$x_j^l = f(pool(x_i^{l-1}) + b_j^l) \tag{3}$$

其中, x_j^l为前一层卷积层窗口值, pool 为取样函数. 本文神经网络在 4 个交替的卷积层和池化层之后,连接 1 层级联卷积层. 在数据通道数目固定的情况下, 级联卷积层的参数量更小, 同时感受野保持不变, 减少了运算量, 不会影响整体网络的收敛性. 级联卷积层中加入的激活层, 增加了神经元之间数据的非线性转换, 在不使网络变得复杂的情况下, 优化神经网络效果. 级联卷积层为 4 个卷积层, 每 1 层之后加入激活层, 最后与池化层相连. 结构如图 4 所示.

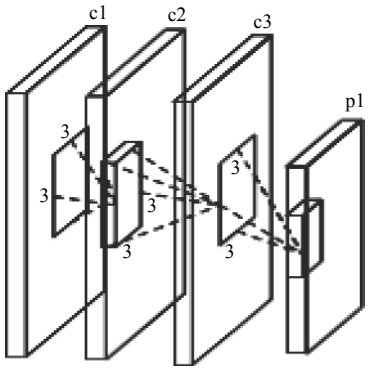

图 4 级联卷积层

数据经过神经网络传递, 最终全连接层用于采集之前网络的局部信息. 为了提升卷积神经网络的性能, 网络中神经单元的激活函数采用 ReLU 函数. 公式如下:

$$f(x) = \max(0, x) \tag{4}$$

ReLU 函数在输入值小于 0 时, 输出值为 0; 当输出值大于 0 时, 输出值与输入值相等. 常用激活函数比较如图 5 所示.

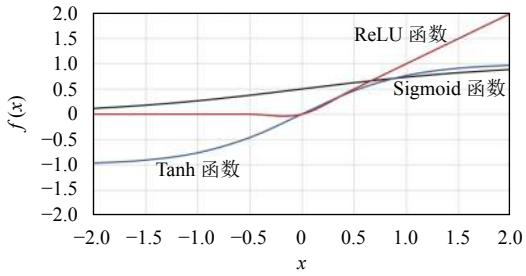

图 5 激活函数对比

输出层采用 Softmax 逻辑回归, 网络输出 5 个概率, 分别对应于 Normal、Dos、Probe、U2R 和 R2L, 选取其中概率最大的标签作为最终的预测结果.

本文卷积神经网络通过减小卷积核尺寸、增加级联卷积层和减少网络层数等技术, 保持整体参数规模不大. 采用 3×3 卷积核, 相对于传统 7×7 卷积核, 减少了运算量. 级联卷积层在保持参数量变小的前提下, 保持网络可以获得相同的感受野. 采用较少的网络层数, 减少了计算的复杂度. 结合卷积神经网络权值共享等特性, 本文实验表明网络可以达到实时性要求.

2.3 卷积神经网络训练算法

卷积神经网络经常采用 BP 算法进行训练[15]. 开始训练之前, 使用小的随机数对网络中的参数进行初始化设置, 之后训练过程主要分成两个步骤. 首先, 训练集中带标签的数据传入神经网络, 经过神经网络中连接参数的运算, 输出层结果与数据标签进行对比; 然后, 输出结果与标签的误差经过反向传播, 进而调整网络中的参数. 经过不断训练, 网络中的参数得到不断调整, 使得损失函数结果得到优化, 最终停止训练.

梯度下降法每次使用全部数据进行参数更新, 运算量非常大, 使用训练时间较长. 进一步采用随机梯度下降, 每次选取其中一条数据进行更新, 训练速度大大提升, 但是会陷入局部最优解. 采用批量更新, 每次选取部分样本数据, 在运算时间和效果之间取得平衡. 学习率的选取深刻影响模型训练的时间和效果, 初始采用较大学习率, 训练一段时间后, 降低学习率, 以完成最终的网络训练[16]. 本文采用初始学习率 0.01, 之后降低学习率.

3 实验分析

3.1 数据预处理

电网工控系统中数据主要分为 2 种: 从传感器传入控制中心的数据和从控制中心对各个传感器发送的命令指令, 具有数据格式固定、数据包小而多等特点[17]. 本文使用公开权威数据集 KDD99. 数据集格式如图 6 所示.

```
0,tcp,http,SF,278,397,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,11,47,0.00,0.
00,0.00,0.00,1.00,0.00,0.00,0.11,26,255,1.00,0.00,0.04,0.02,0.00,0.00,0.
00,0.00,normal.
0,udp,private,SF,105,146,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0.00,0
.00,0.00,0.00,1.00,0.00,0.00,0.00,255,254,1.00,0.01,0.00,0.00,0.00,0.00,
0.00,0.00,normal.
0,udp,private,SF,105,146,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,0
.00,0.00,0.00,1.00,0.00,0.00,0.00,255,254,1.00,0.01,0.00,0.00,0.00,0.00,
0.00,0.00,normal.
```

图 6 数据格式

KDD99 数据集共 500 万余条, 每条数据代表一次网络行为, 每条初始数据由 42 维组成, 其中 38 维为数值特征, 4 维为符号特征. 每一个维度之间用逗号分隔, 用来刻画网络协议、字节数、流量统计等特征, 最后一维标签为 normal (正常网络访问) 或者 attack (网络入侵).

根据第二维度代表的通信协议, 选取其中 TCP 和 UDP 协议数据. 考虑只有数值维度可以通过卷积神经网络处理, 将字符型的数据维度通过热编码技术转化为数值. 处理之后数据维度提升至 64 维, 维度过多,影响训练时间. 考虑 U2R 和 R2L 入侵对电网造成危害更大, 提高数据集中 U2R 和 R2L 数据比例, 使得 4 种入侵行为数量均衡. 不同维度数据大小相差过多, 如图 6 中第一行数据, 数值型最大值为 397, 而部分维度为 0.04、0.02. PCA 算法为常见降维算法, 利用方差理论, 根据不同特征维度对结果的影响, 选取影响较大的维度, 将多维数据映射为低维数据. 采用 Min-Max 标准归一化算法, 在不破坏原有映射关系的情况下, 将数据映射到 0 和 1 之间:

$$x = \frac{x - \mathrm{Min}}{\mathrm{Max} - \mathrm{Min}} \tag{5}$$

根据 PCA 算法, 选取其中重要的 52 维数据. 符合条件数据选取训练集样本 95 136 个, 其中正常样本 49 603 个, 入侵样本 45 533 个. 测试集 46 331 个样本, 其中正常样本 27 743 个, 入侵样本 18 588 个. 其中 U2R 和 R2L 在原始集中样本量就非常少, 全部选取, 处理后放入数据集.

3.2 SVM 算法和 K-means 算法

SVM (Support Vector Machine, 支持向量机) 是一种有监督的机器学习算法, 常用于二分类问题, 在本实验中用于将数据判定为正常数据或者入侵数据. 学习目标为二元变量 1(入侵) 或者–1(正常), 输入数据为电网数据和数据标签. 算法求解过程是给定决策函数求

超平面 ($\omega \times x+b=0$), 转化为数学函数优化问题. SVM 算法有严格的数学推导过程, 在深度学习之前一直是常用的机器学习算法, 但是依赖于计算能力和数据量的提升, 卷积神经网络能达到更好的效果[18].

K-means 算法核心思想是, 选取 k 个对象作为聚类的中心, 进而算出其余对象与中心的距离并划分到最近的中心, 不断更新中心位置直到聚类中心不再变化为止[19]. 实验中对于 K-means 算法的应用使用 sklearn 中相关的库, 将数据分为 5 个中心, 分别对应于 1 个正常数据中心和 4 个入侵行为中心.

3.3 实验与分析

实验评价指标定义如下所示:

$$\begin{cases} 准确率=\dfrac{TP+TN}{TP+TN+FP+FN} \\ 误检率=\dfrac{FP}{TN+FP} \end{cases} \tag{6}$$

其中, TP: 实际为入侵行为, 预测为入侵行为; TN: 实际为正常行为, 预测为正常行为; FP: 实际为正常行为, 预测为入侵行为; FN: 实际为入侵行为, 预测为正常行为. 准确率即算法模型判断数据正确所占的比例, 误检率用于衡量算法模型错误分类的比例, 即将数据通过相应训练之后的模型得出为正常或者入侵行为, 进而与数据标签对比得出此两者指标大小. 两者组合可以反映出模型对于入侵数据的判断准确率, 同时误检率作为参考.

本文使用 Tensorflow 深度学习框架进行神经网络训练, 训练集的全部数据训练一次称为一次 epoch,实验不断增加 epoch 数量, 如图 7 所示. 其中,准确率先上升, 准确率上升的速率先上升后减缓, 达到 450 次之后, 有下降趋势, 此时网络模型达到最优结构, 停止训练.

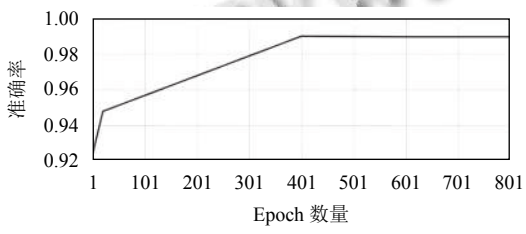

图 7 训练结果

图 8 所示为 SVM 算法和 K-means 算法与本文网络模型对比正确率和误检率实验数据.

图 8 中展示了不同方法的入侵检测准确率和误检率, 其中基于 K-means 算法的实验准确率为 75.63%, 误检率为 9.46%; 基于 SVM 算法的实验准确率为 87.79%, 误检率为 8.33%. 本文基于卷积神经网络的算法准确率为 95.37%, 误检率为 8.09%, 相较于另外两种算法具有更高的准确率和更低的误检率, 其中准确率的提升相较于误检率的下降更为明显. 分析原因如下: 传统 SVM 算法和 K-means 算法相较于深度学习算法, 如本文卷积神经网络算法, 本文算法模型复杂度更高, 因此具有更强的表达能力.

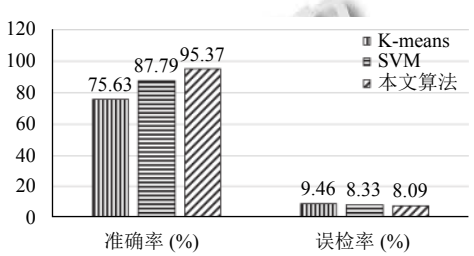

图 8 实验结果对比

4 结论与展望

针对如何在电网工控系统中检测出入侵行为这一问题, 本文提出一种基于卷积神经网络的入侵检测算法. 本文采用 KDD99 数据集作为原始数据集, 根据电网工控系统自身实时性和稳定性的要求, 系统数据流向固定、数据包小等特点, 在数据处理过程剔除不重要的特征维度, 筛选出训练集和测试集. 构建不同的卷积神经网络架构, 选取不同的卷积核、池化核和网络层数等参数进行实验. 加入级联卷积层优化网络结构, 最终与传统 SVM 算法和 K-means 算法进行对比, 在准确率和误检率方面显示本文算法有良好表现.

参考文献

1 Farwell JP, Rohozinski R. Stuxnet and the future of cyber war. Survival, 2011, 53(1): 23–40. [doi: 10.1080/00396338.2011.555586]

2 Li BJ, Liu Y, Li HJ, et al. Enlightenment on the security of cyber information system under smart grid from ukraine blackout. Electric Power, 2017, 50(5): 71–77.

3 王彬笙. 地区供电网调度实时数据网络安全分析. 通讯世界, 2019, 26(11): 181–182. [doi: 10.3969/j.issn.1006-4222.2019.11.120]

4 卿斯汉, 蒋建春, 马恒太, 等. 入侵检测技术研究综述. 通信学报, 2004, 25(7): 19–29. [doi: 10.3321/j.issn:1000-436X.2004.07.003]

5 Gu JX, Wang ZH, Kuen J, *et al*. Recent advances in

convolutional neural networks. Pattern Recognition, 2018, 77: 354–377. [doi: 10.1016/j.patcog.2017.10.013]

6 Kemmerer RA, Vigna G. Intrusion detection: A brief history and overview. Computer, 2002, 35(4): supl27–supl30.

7 Lunt TF. A survey of intrusion detection techniques. Computers & Security, 1993, 12(4): 405–418.

8 Tang CH, Xiang Y, Wang Y, *et al*. Detection and classification of anomaly intrusion using hierarchy clustering and SVM. Security and Communication Networks, 2016, 9(16): 3401–3411. [doi: 10.1002/sec.1547]

9 Mohamad Tahir H, Hasan W, Md Said A, *et al*. Hybrid machine learning technique for intrusion detection system. Proceedings of International Conference on Computing and Informatics. Istanbul, Turkey. 2015. 464–472.

10 Horng SJ, Su MY, Chen YH, *et al*. A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert Systems with Applications, 2011, 38(1): 306–313. [doi: 10.1016/j.eswa.2010.06.066]

11 Gauthama Raman MR, Somu N, Kirthivasan K, *et al*. A hypergraph and arithmetic residue-based probabilistic neural network for classification in intrusion detection systems. Neural Networks, 2017, 92: 89–97. [doi: 10.1016/j.neunet.2017.01.012]

12 杨昆朋. 基于深度学习的入侵检测 [硕士学位论文]. 北京: 北京交通大学, 2015.

13 He KM, Sun J. Convolutional neural networks at constrained time cost. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2014.5353–5360.

14 Boureau YL, Le Roux N, Bach F, *et al*. Ask the locals: Multi-way local pooling for image recognition. Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain. 2011.2651–2658.

15 贾丽会, 张修如. BP 算法分析与改进. 计算机技术与发展, 2006, 16(10): 101–103, 107. [doi: 10.3969/j.issn.1673-629X.2006.10.034]

16 吴立锋, 吴经龙. BP 算法学习率自适应性研究. 大众科技, 2011,(12):16-18.[doi:10.3969/j.issn.1008-1151.2011.12.007]

17 钟志琛. 基于网络流量异常检测的电网工控系统安全监测技术. 电力信息与通信技术, 2017, 15(1): 98–102.

18 池亚平, 凌志婷, 王志强, 等. 基于支持向量机与 Adaboost 的入侵检测系统. 计算机工程, 2019, 45(10): 183–188, 202. [doi: 10.19678/j.issn.1000-3428.0051976]

19 华辉有, 陈启买, 刘海, 等. 一种融合 Kmeans 和 KNN 的网络入侵检测算法. 计算机科学, 2016, 43(3): 158–162. [doi: 10.11896/j.issn.1002-137X.2016.03.030]