基于深度级联网络的入侵检测算法研究

郭卫霞，张　伟，杨国玉
（中国大唐集团科学技术研究院，北京 100043）

摘　要： 针对海量多源异构的网络流量数据难以用传统的机器学习算法有效提取特征，分类效果差的问题，提出一种基于深度级联网络的入侵检测算法，利用神经网络自动学习特征的能力，将卷积神经网络和长短期记忆网络结合起来，同时提取流量数据的空间特征和时序特征，并采用 softmax 进行分类，提高模型的检测性能和泛化能力。最后将该算法在 KDDCUP99 数据集上进行验证，实验结果表明，该入侵检测模型相较于SVM、DBN 等算法有更高的检测率，准确率可达 95.39%，误报率仅 0.96%，有效提高了入侵检测分类性能。

关键词： 入侵检测；特征提取；卷积神经网络；长短期记忆网络
中图分类号： TN03；TP393　　**文献标识码：** A　　DOI：10.16157/j.issn.0258-7998.211394

中文引用格式： 郭卫霞，张伟，杨国玉. 基于深度级联网络的入侵检测算法研究[J].电子技术应用，2021，47(11)：68-72.
英文引用格式： Guo Weixia，Zhang Wei，Yang Guoyu. Research on intrusion detection algorithm based on deep cascade network[J]. Application of Electronic Technique，2021，47(11)：68-72.

Research on intrusion detection algorithm based on deep cascade network

Guo Weixia，Zhang Wei，Yang Guoyu
(China Datang Corporation Science and Technology Research Institute，Beijing 100043，China)

Abstract： Aiming at the problem that traditional machine learning algorithms are difficult to effectively extract features from massive multi-source heterogeneous network traffic data，and the classification effect is poor，an intrusion detection algorithm based on deep cascaded network is proposed，which uses the ability of neural network to automatically learn features. Convolutional neural network (CNN) is combined with long short-term memory network(LSTM) to extract the spatial and temporal characteristics of traffic data at the same time. And softmax is used for classification to improve the detection performance and generalization ability of the model. Finally，the algorithm is verified on the KDDCUP99 data set. The experimental results show that the intrusion detection model has a higher detection rate than SVM，DBN and other algorithms，with an accuracy rate of 95.39% and a false alarm rate of only 0.96%，which effectively improves intrusion detection classification performance.

Key words： intrusion detection；feature extraction；convolutional neural network(CNN)；long short-term memory(LSTM)

0 引言

信息技术的高速发展极大地丰富和便利了人们的学习、生活和工作，但与此同时网络攻击导致的网络异常中断、用户个人信息泄露等事件频频发生，互联网所面临的各种安全威胁变得日益严重，因此维护网络安全变得至关重要。网络入侵检测作为一种动态有效的主动检测技术，能够通过分析网络流量数据识别具有攻击行为的信息，在网络受到攻击之前进行及时的拦截和响应，目前已经成为信息安全领域研究的重要内容之一。

入侵检测技术最早于 1980 年由 Anderson[1]提出。1987 年 Denning[2]采纳了 Anderson 技术报告中的检测建议，提出了入侵检测专家系统(Intrusion Detection Expert System，IDES)，后来大量的研究人员提出了各种入侵检测算法来提升检测效果。近些年，机器学习算法被广泛应用在各种入侵检测技术中，文献[3]将支持向量机(Support Vector Machine，SVM)应用于网络异常流量检测中。文献[4]利用 K 近邻(K-Nearest Neighbor，KNN)算法进行网络入侵检测，提高了分类效果。文献[5]基于并行 K-means 聚类算法对异常流量数据进行分簇，降低分类误差。上述算法在一定程度上提高了入侵检测精度，但是基于机器学习的入侵检测算法依赖于人工提取的数据特征，需要人为进行大量复杂的特征工程，并且对于海量多源异构的网络入侵数据没有很好的鲁棒性。

近年来，随着深度学习的迅速崛起，卷积神经网络、循环神经网络、深度置信网络等多种深度学习算法逐渐应用到入侵检测领域。文献[6]、[7]基于卷积神经网络进行入侵检测，其漏检率和误检率均大幅低于基于传统机器学习的入侵检测算法。文献[8]、[9]将循环神经网络应用到入侵检测算法中，防止了训练过程中过拟合问题的发生，也提升了分类准确率。文献[10-12]提出基于改进

深度置信网络的入侵检测算法，将深度置信网络与极限学习机结合，提高网络泛化能力，最终得到较好的检测效果。

本文提出一种基于深度级联网络的入侵检测模型，由 CNN 网络和 LSTM 网络串联构成，可以同时自学习流量数据的空间特征和时序特征，避免繁琐的特征工程，提高模型的表达能力和泛化能力。

1 基于深度级联网络的入侵检测模型

1.1 CNN 网络

基于深度级联网络的入侵检测模型的第一层 CNN 网络采用 16 层的 VGGNet[13]来提取流量数据的空间特征。VGGNet 是由牛津大学 Visual Geometry Group 联合 Google DeepMind 公司一起研发的深度卷积神经网络，它构建了很多不同深度(11～19 层)的网络，主要贡献在于探索了卷积神经网络的深度对其最终的检测性能有一定程度的影响。

VGG16 网络总体包含了 5 个卷积层、5 个池化层、3 个全连接层以及 1 个 Softmax 输出层，一共进行了 13 次卷积操作，卷积层数较深，且每层通道数较多，因而能够提取到较为丰富和抽象的高级语义特征；网络结构非常规整，统一采用 3×3 大小的卷积核(stride 为 1)和 2×2 的池化核(stride 为 2)，使得网络收敛速度较快；卷积层与层之间的池化层采用 Max-Pooling，隐层的激活单元全部采用 ReLU 激活函数，结构十分简洁。网络结构如图 1 所示，其中 Conv1～Conv5 表示 5 个卷积层，C 表示卷积操作，P 表示池化操作，FC 为全连接层。

1.2 LSTM 网络

本文采用长短期记忆网络(Long Short-Term Memory Network，LSTM)[14]提取流量数据的时序特征。LSTM 是一种特殊的 RNN 网络，常用于自然语言处理、语音识别等的序列信号处理任务中，能够解决 RNN 网络长时间序列训练中存在的梯度消失问题。LSTM 细胞结构如图 2 所示，通过输入门、遗忘门和输出门 3 个门控单元结构，对于处理长依赖问题有很好的效果。

LSTM 每个门控结构计算操作如下：

(1)遗忘门：决定从细胞状态中丢失上一时刻输入信息的比例，使用当前时刻的输入 x_t 和上一时刻隐藏层的输出 h_{t-1}，通过激活函数映射为[0，1]中的一个值保存至细胞状态 C_{t-1} 中，其中 0 表示完全丢弃，1 表示完全保留。

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f) \tag{1}$$

式中，W_f 和 b_f 分别为遗忘门的循环权重和偏置常量，σ 为 Sigmoid 激活函数，f_t 为丢失比例。

(2)输入门：决定当前输入信息保存至细胞状态的比例，首先由 Sigmod 层确定当前细胞中哪些信息需要更新，然后由 tanh 层生成候选值向量，用来更新当前细胞中的信息。最终的细胞状态是由遗忘门丢弃上一时刻细胞的值和输入门保留当前输入的信息共同决定的，整体操作如下：

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i) \tag{2}$$

$$\widetilde{C}_t=\tanh(W_C\cdot[h_{t-1},x_t]+b_C) \tag{3}$$

$$C_t=f_t\cdot C_{t-1}+i_{t-1}\cdot\widetilde{C}_t \tag{4}$$

(3)输出门：根据当前细胞状态决定最终的输出值，首先由 Sigmod 层确定当前时刻细胞状态中信息的输出比例，然后将细胞状态经由 tanh 层映射为[-1，1]中的一个值，并和 Sigmod 层的输出相乘即为最终的输出值，计算公式如下：

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o) \tag{5}$$

$$h_t=o_t\cdot\tanh(C_t) \tag{6}$$

网络流量数据通常是按照时间序列采集存储的，LSTM 通过 3 个门控单元对数据进行非线性映射，解决了数据在网络中传输时间越长丢失越严重的问题，能够更好地保留每一时刻流量信息的关键特征。

1.3 深度级联网络

本文把 VGGNet 和 LSTM 网络结合起来，构成一个更深层次的多特征融合的神经网络，同时提取网络流量的空间特征和时序特征，使得网络可以更全面地表达流量

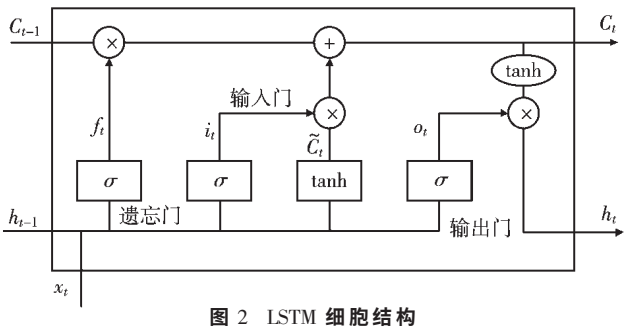

图 2 LSTM 细胞结构

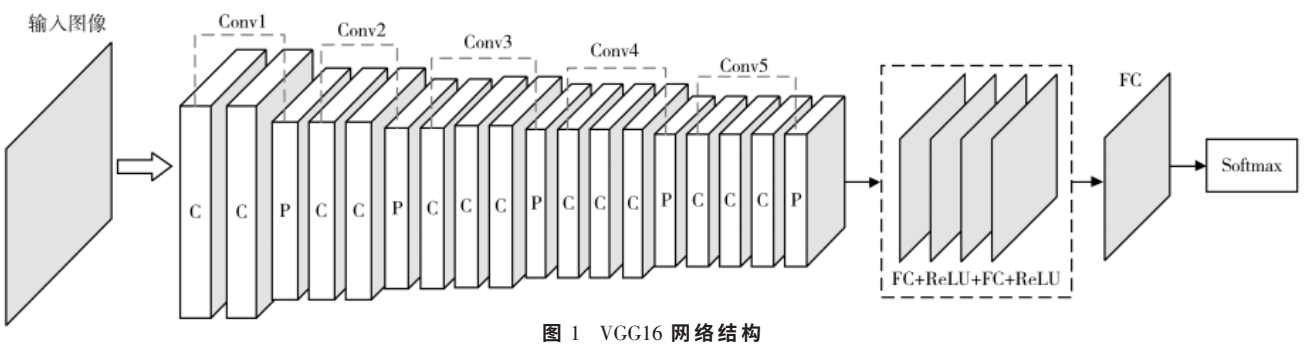

图 1 VGG16 网络结构

的属性。基于深度级联网络的入侵检测模型整体网络结构如图 3 所示，首先将流量数据经过预处理后得到的特征向量转化为 11×11 的灰度图像作为 CNN 网络的输入，先经过 VGG16 网络提取流量的空间特征，然后将全连接层输出的特征图分为 10 个时间步骤送入 LSTM 网络，提取流量的时序特征，本文采用的 LSTM 网络包含 2 个隐藏层，各有 256 个神经元细胞，可以更好地保留流量的特征。LSTM 网络输出的向量送入全连接层，并采用 Softmax 进行分类，输出流量数据属于各类别的概率，概率最大的类别即为网络预测该条流量所属的类别。

2 实验与结果分析

2.1 实验环境及参数设置

本文选择深度学习框架 TensorFlow 1.8.0 作为实验平台，平台软硬件配置如下：操作系统为 Ubuntu 16.04，16 GB 内存，GPU 为 NVIDIA GeForce GTX Titan Xp，GPU 加速库为 CUDA 9.0 和 CUDNN 7.6。

网络载入利用 ImageNet[15]预训练过的 VGG16 网络来初始化特征提取网络的权重，训练过程中选用 SGD 随机梯度下降算法优化网络模型，总迭代次数为 70 000 次，学习率初始值设为 0.001，动量系数为 0.9，5 万次迭代后学习率衰减为 0.000 1。

2.2 数据集预处理

实验采用 KDDCUP99 数据集[16]，该数据集是基于美国国防部高级规划署，通过仿真不同的用户类型、网络流量和攻击手段，采集了 9 周的网络连接和系统审计数据整理而成。KDDCUP99 共有 500 万条记录，每条记录包含 41 个特征和 1 个分类标记(label)，label 分为正常(Normal)和异常(Attack)，异常类型包含 Dos、Probe、R2L 和 U2R 4 种情况。

考虑到训练时间和内存占用情况，本文选用 KDDCUP99 数据集中的“kddcup.data_10_percent”作为训练集，共 494 021 条，“corrected”作为测试集，共 311 029 条，实验数据的类别分布情况如表 1 所示。

表 1 数据集类别分布 (条)

Label	Normal	Attack	Dos	Probe	R2L	U2R
训练集	97 278	396 743	391 458	4 107	1 126	52
测试集	60 593	250 436	229 853	4 166	16 189	228

KDDCUP99 数据集的 41 维特征中有 38 维是数字特征，其余 3 维是符号特征，而 CNN 要求输入数据为数字矩阵，因此在进行模型训练之前需要对数据集进行预处理。

符号特征数字化：将 protocol_type、service 和 flag 这 3 种符号特征转换为二进制向量，如 protocol_type 中的TCP、UDP、ICMP 协议分别表示为[0，0，1]、[0，1，0]、[1，0，0]。同理，flag 特征可转换为 11 维二进制特征，service 特征可转换为 70 维二进制特征，最终数据集中的每一条记录可由 41 维转换为 122 维。

数字特征归一化：为了消除特征之间由于量纲不同导致的差异性，对所有特征采用最大最小归一化法进行均值化处理，将每个特征的取值映射为[0，1]之间，公式如下：

$$y' = \frac{y - M_{\min}}{M_{\max} - M_{\min}} \tag{7}$$

式中，$M_{\max}$ 和 $M_{\min}$ 分别为某一维特征中的最大值和最小值，y 为原始特征值，y' 为归一化之后的值。

2.3 实验评价标准

入侵检测算法的性能通常通过准确率 AC、检测率 DR、误报率 FA 来衡量。AC 是指正确分类的样本数占总测试样本数量的比例，DR 是指正确识别为攻击样本数

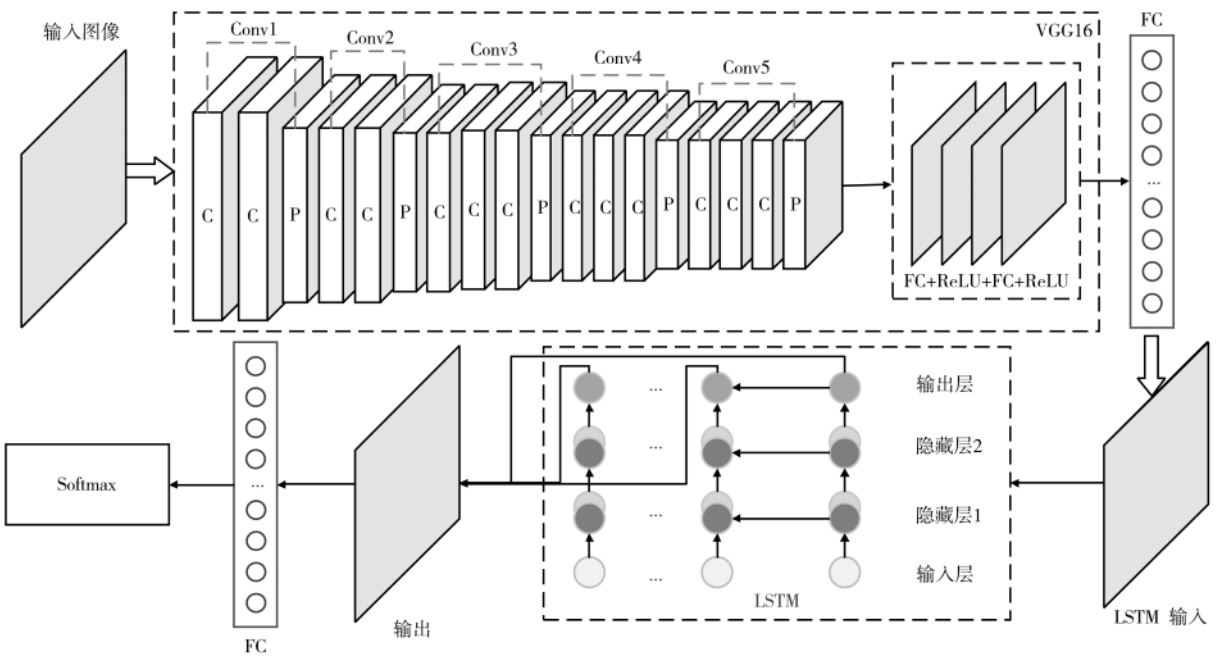

图 3 深度级联网络结构

占总测试攻击样本数量的比例，FA 是指错误识别为攻击的正常样本数占总测试正常样本数量的比例。准确率和检测率越高，误报率越低，说明该入侵检测算法的性能越好。各评价指标计算公式如下：

$$AC=\frac{TP+TN}{TP+FP+FN+TN} \tag{8}$$

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

$$FA=\frac{FP}{FP+TN} \tag{10}$$

式中，TP 为被正确识别为攻击样本的数量，FP 为被错误识别为攻击的正常样本的数量，TN 为被正确识别为正常样本的数量，FN 为被错误识别为正常的攻击样本的数量。

2.4 结果分析

为了验证本文改进算法的有效性，本次实验分别对比了 VGGNet、LSTM 以及改进后的深度级联网络对不同种类攻击的识别情况，检测率对比结果如表 2 所示，可以看出，改进后的深度级联网络在 Normal、Dos、Probe、R2L 和 U2R 这 5 类攻击上检测率均高于 VGGNet 和 LSTM 网络，这说明将 CNN 网络与 LSTM 结合，能够获得更好的检测能力。但对于异常 R2L 和 U2R，训练集样本数量较少，模型学习到的特征也相对较少，检测率略显不足。

表 2 各网络对不同种类攻击的检测率对比

(%)

模型	Normal	Dos	Probe	U2R	R2L
VGGNet	96.79	97.13	84.65	68.8	77.59
LSTM	96.41	98.02	86.25	55.86	73.77
深度级联网络	97.39	98.89	88.74	72.47	81.43

本文还对比了不同大小的训练集对 VGGNet、LSTM 以及本文改进算法的影响，各算法的检测准确率和误报率对比如图 4 和图 5 所示，其中训练集采用不同比例的 kddcup.data_10_percent 数据集，由图可知，当训练集样本数量较少时，CNN 的检测性能存在明显优势，随着训练集样本数量的增多，LSTM 的检测性能优势逐渐显现，深度级联网络的检测效果更佳。由此可以验证在训练样本充足的情况下，采用 CNN 网络和 LSTM 同时学习流量数据的空间特征和时序特征，能够更好地提升入侵检测性能。

本文改进的深度级联网络与常见的机器学习算法的性能比较结果见表 3，由此可见本文提出的深度级联网络相较于传统的机器学习算法有较大的性能提高，对比效果较好的 KNN-RF 算法，深度级联网络的准确率提升 1.03%，误报率下降 1.38%，这也验证了在入侵检测方面，深度神经网络凭借强大的特征学习能力，比传统的机器学习算法更具有优势。

本文与其他基于深度学习的入侵检测算法性能对比见表 4，可以看出，本文提出的深度级联网络的准确率均高于基于 DBN、CNN、RNN 的入侵检测模型，但误报率也相对较高，后期可以进一步优化本文算法以获得更好的整体检测性能。

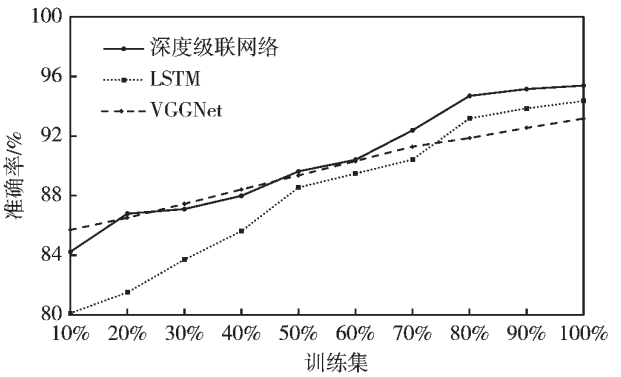

图 4 不同大小训练集下各算法的准确率对比

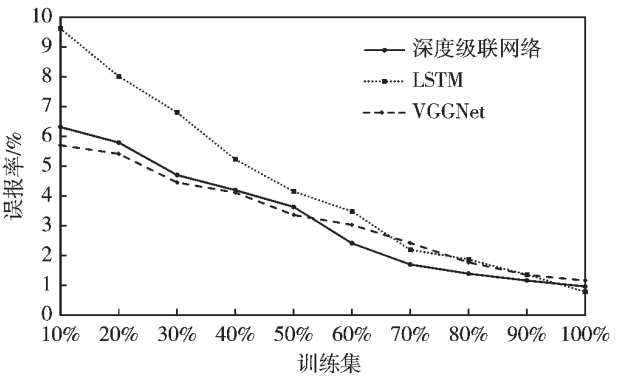

图 5 不同大小训练集下各算法的误报率对比

表 3 本文和机器学习算法性能对比

(%)

模型	AC	DR	FA
RF[17]	92.66	90.96	0.99
KNN-RF[17]	94.36	93.55	2.34
SVM[18]	86.82	86.57	1.96
SOM[18]	90.82	89.57	1.16
深度级联网络	95.39	94.64	0.96

表 4 本文和其他深度学习算法性能对比

(%)

模型	AC	DR	FA
DBN[18]	93.49	92.33	0.76
CNN[19]	92.18	90.95	0.97
RNN[20]	94.36	94.10	0.38
深度级联网络	95.39	94.64	0.96

3 结论

本文针对海量多源异构的网络流量数据难以提取特征的问题，提出一种基于深度级联网络的入侵检测算法，同时利用卷积神经网络和循环神经网络的优势，将 VGG16 网络与 LSTM 级联，分别提取流量数据的空间特征和时序特征，并通过 KDDCUP99 数据集进行训练和测试。实验结果表明，本文改进的深度级联网络能够有效地提高入侵检测准确率，降低误报率。但对于稀疏攻击

U2R 和 R2L，本文模型仍需要进行优化改进。

参考文献

[1] ANDERSON J P.Computer security threat monitoring and surveillance[R].Fort Washington, Pennsylvania, 1980.
[2] DENNING D E.An INTRUSION-DETECTION MODel[C]//IEEE Symposium on Security and Privacy, 1986: 118-131.
[3] BALABINE I, VELEDNITSKY A.Method and system for confident anomaly detection in computer network traffic: USA, 14/627, 963[P].2015-02-20.
[4] SIDDIQUI S, KHAN M S, FERENS K, et al.Detecting advanced persistent threats using fractal dimension based machine learning classification[C]//Proceedings of the 2016 ACM on International Workshop on Security And Privacy Analytics, 2016: 64-69.
[5] VELEA R, CIOBANU C, MARGARIT L, et al.Network traffic anomaly detection using shallow packet inspection and parallel K-means data clustering[J].Studies in Informatics and Control, 2017, 26(4): 387-395.
[6] LIU Y C, LIU S L, ZHAO X.Intrusion detection algorithm based on convolutional neural network[C]//Proceedings of 2017 4th International Conference on Engineering Technology and Application, 2017.
[7] ZHOU H Y, WANG Y, LEI X C, et al.A method of improved CNN traffic classification[C]//Proceedings of 2017 13th International Conference on Computational Intelligence and Security, 2017: 177-181.
[8] JIHYUN K, HOWON K.Applying recurrent neural network to intrusion detection with hessian free optimization[C]//roceedings of 16th International Workshop on Information Security Applications, 2015: 357-369.
[9] 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: 954-961.
[10] WANG H P, XU L, GU G F, et al.FloodGuard: a DoS attack prevention extension in software-defined networks[C]//Proceedings of IEEE/IFIP International Conference on Dependable Systems and Networks, 2015.
[11] 汪洋，伍忠东，火忠彩.基于 DBN-KELM 的入侵检测算法[J].计算机工程，2019，45(10)：171-175，182.
[12] 汪洋，伍忠东，朱婧.基于深度序列加权核极限学习的入侵检测算法[J].计算机应用研究，2020，37(3)：829-832.
[13] SIMONYAN K, ZISSERMAN A.Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014.
[14] HOCHREITER S, SCHMIDHUBER J.Long short-term memory[J].Neural Computation, 1997, 9(8): 1735-1780.
[15] Li Feifei.ImageNet: crowdsourcing, benchmarking & other cool things[C]//CMU VASC Seminar, 2010, 16: 18-25.
[16] Information and Computer Science, University of California.KDD Cup 1999 Data[DB/OL].(1999-10-28)[2021-02-09].http://kdd.ics.uci.edu/databases/kddcup99.
[17] 任家东，刘新倩，王倩，等.基于 KNN 离群点检测和随机森林的多层入侵检测方法[J].计算机研究与发展，2019，56(3)：566-575.
[18] GAO N, GAO L, GAO Q L, et al.An intrusion detection model based on deep belief networks[C]//Proceedings of the IEEE Conference on Advanced Cloud and Big Data, 2014: 247-252.
[19] 贾凡，孔令智.基于卷积神经网络的入侵检测算法[J].北京理工大学学报，2017，37(12)：1271-1275.
[20] SHEIKHAN M, JADIDI Z, FARROKHI A.Intrusion detection using redecued-size RNN based on feature grouping[J].Neural Computing and Applications, 2012, 21(6): 1185-1190.

（收稿日期：2021-02-09）

作者简介：

郭卫霞(1994-)，女，硕士研究生，助理工程师，主要研究方向：深度学习、信息安全。

张伟(1976-)，男，硕士研究生，高级工程师，主要研究方向：网络安全。

杨国玉(1980-)，男，硕士研究生，高级经济师，主要研究方向：信息化与网络安全管理。

扫码下载电子文档

（上接第 67 页）

[12] KAJIHARA K, IZUMI S, YOSHIDA S, et al.Hardware implementation of autoregressive model estimation using Burg's method for low-energy spectral analysis[C]//2018 IEEE International Workshop on Signal Processing Systems (SiPS).IEEE, 2018: 199-204.
[13] 苏丽.Matlab 定点仿真在 FPGA 验证平台中的应用[J].电子科技，2013，26(5)：71-73.
[14] 张林生.数字信号处理系统的定点化技术研究[D].哈尔滨：哈尔滨工业大学，2010.
[15] 钱莹晶，周群.Burg 频谱估计算法的硬件加速方法研究[J].电子测量与仪器学报，2015，29(9)：1382-1390.

（收稿日期：2021-02-20）

作者简介：

郭鸣晗(1996-)，女，硕士，主要研究方向：数字信号处理、数字集成电路设计。

陈立平(1983-)，男，硕士，高级工程师，主要研究方向：生物信号处理、SoC 设计。

张浩(1975-)，通信作者，男，博士，副研究员，主要研究方向：生物信号处理、SoC 设计，E-mail：haozhang@ime.ac.cn。

扫码下载电子文档

版权声明