基于 1d-MSCNN+GRU 的工业入侵检测方法研究*

宗学军,宋治文,何 戡,连 莲
(沈阳化工大学 信息工程学院,辽宁 沈阳 110142)

摘 要: 针对传统机器学习方法对特征依赖大,以及传统卷积神经网络只通过提取重要的局部特征来完成识别分类,收敛速度慢的问题,提出了一维多尺度卷积神经网络和门控循环单元相结合的入侵检测方法。该方法使用一维多尺度卷积神经网络加强对特征的捕捉能力,加快收敛速度,采用门控循环单元把握空间特征,减少通道数量扩张,降低数据维度。使用 KDD CUP 99 数据集和密西西比州大学的天然气管道的数据集进行仿真实验,结果表明与经典的机器学习分类器相比,该方法具有较高的入侵检测性能和较好的泛化能力。

关键词: 一维多尺度卷积;门控循环单元;入侵检测;深度学习
中图分类号: TP391.9 **文献标识码:** A **DOI:** 10.19358/j.issn.2096-5133.2021.09.005
引用格式: 宗学军,宋治文,何戡,等. 基于 1d-MSCNN+GRU 的工业入侵检测方法研究[J].信息技术与网络安全,2021,40(9):25-31.

Research on industrial intrusion detection method based on 1d-MSCNN+GRU model

Zong Xuejun, Song Zhiwen, He Kan, Lian Lian
(College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: In order to solve the problem that traditional machine learning methods rely heavily on features, and traditional convolutional neural network only extracts important local features to complete recognition and classification, and the convergence speed is slow, an intrusion detection method combining 1-dimensional multiscale convolutional neural network and gated recurrent unit is proposed. In this method, 1-dimensional multiscale convolutional neural network is used to enhance the ability to capture features, speed up the convergence speed, and the gating cycle unit is used to grasp the spatial features, reduce the expansion of the number of channels and reduce the data dimension. The KDD CUP 99 data set and the natural gas pipeline data set of the University of Mississippi are used for simulation experiments. The results show that the method has higher intrusion detection performance and better generalization ability than the classical machine learning classifier.

Key words: 1-dimensional Multiscale Convolutional Neural Networks(1d-MSCNN); Gated Recurrent Unit(GRU); intrusion detection; deep learning

0 引言

随着工业控制网络(ICN)的高速发展,ICN 安全已经是全球性重要问题之一,工业入侵检测作为一种 ICN 安全防护技术已成为研究热点。在全球每年的网络安全事故中,其中有上百起攻击都是针对工业控制系统(Industrial Control System, ICS),虽然所占的比重只是网络安全事件的一小部分,但是所造成的影响对国家而言都是巨大的,最为严重的就是经济损失[1]。因此如何有效地从入侵数据中选择特征进行多分类,并提高数据特征提取的准确度,在整个网络信息安全领域具有重要的研究价值。

机器学习在入侵检测中应用很多,例如支持向量机(Support Vector Machine, SVM)[2-3]、K 均值聚类算法[4]和贝叶斯网络模型[5]。上述算法在处理特征

* 基金项目:2020 年度辽宁省重点研发计划项目(2020JH2/10100035);2019 年度"辽宁省高等学校创新团队及创新人才支持计划"项目(LT2019010)

维度少时拥有较好的检测效果，但却无法满足当今网络安全领域中大量、高维的数据特征分类精度，因此需要开展深度学习的研究。

国内外不少学者将深度学习应用于网络安全领域，文献[6]首先使用不同的降维方法去除了多余的特征，然后将降维后得到的数据传递给卷积神经网络(Convolutional Neural Networks，CNN)，但是其忽略了卷积神经网络的自动提取特征的优势。文献[7]利用遗传算法强大的全局寻优能力来获得最优参数从而优化卷积神经网络。文献[8]提出了一种基于两层神经网络的异常流量检测模型。文献[9]提出了一种基于膨胀卷积和门控循环单元组合的入侵检测模型，并在KDD CUP 99数据集和NSL-KDD数据集进行了仿真，取得了较高的准确率。文献[10]对一维卷积神经网络进行了改进，使用CIC-IDS-2017数据集进行了仿真，结果表明该方法具有较高的检测性能，能更好地保留流量数据的局部特征。

传统的卷积神经网络在提取特征时只提取重要的特征，会忽略其他特征，而工业网络中的数据在使用这种方法进行分类时会导致分类准确率大大降低，因此为了解决提取特征不全这一问题，本文使用多尺度一维卷积神经网络(1-dimensional Multiscale Convolution Neural Network，1d-MSCNN)与门控循环单元(Gated Recurrent Unit，GRU)组合模型来对工业网络中的流量数据进行分类。

本文的主要贡献：

(1)提出一种一维多尺度卷积核替代单一尺度卷积核来提取特征，提高数据分类的准确度。

(2)使用GRU模块来获取数据之间的时间关系的特征，以此提高模型训练速度，解决了很多由于深度学习的模型复杂、参数过多而导致的训练难度加大、模型的训练速度变慢的问题。

1 相关工作

1.1 卷积神经网络(CNN)

CNN是深度学习的一种模型[11]，能够通过权值共享的方式快速训练模型，并能有效地减少模型中需要训练的参数。常见的卷积神经网络往往采用二维卷积核。而一维卷积神经网络使用的是一维卷积核，下面介绍一维卷积层与池化层：

(1)一维卷积层：一维卷积层中的每个神经元只与前一层输入神经元中的局部直接相关。卷积核的设置对CNN的性能影响比较大，对于第 l 层(卷积层)，其输出为 x^l，那么对应第 j 个卷积核的输出 x_j^l 为：

$$x_j^l=f\left(\sum_{x_i^{n-1}\in M_j} x_i^{l-1} * W_{ij}^l+b_j^l\right) \tag{1}$$

式中，f 为激活函数，M_j 表示输入数据集合，W_{ij}^l 为卷积核，$*$ 为卷积，b_j^l 为偏置项。本文使用的激活函数为ReLU函数，能大量节省训练时间。

(2)池化层：池化方法有两种，选择相应的池化方法对卷积层的特征图进行池化，其一般形式为：

$$x_j^l=f(\beta_j^l D(x_j^{l-1})+b_j^l)) \tag{2}$$

式中，f 为激活函数，D 为采样函数，β_j^l 为权值，b_j^l 为偏置项。

1.2 门控循环单元(GRU)

GRU模块是循环神经网络(Recurrent Neural Network，RNN)中的一种，能够对数据特征进行充分学习，通过更新门和重置门保存数据在时间维度上的信息[12]，同时也能解决梯度爆炸的问题[13]。

GRU中有两个主门，即更新门和重置门。所有的关系定义如下：

重置门：

$$r_t=\sigma(W_r X_t+W_r H_{t-1}+b_r) \tag{3}$$

更新门：

$$Z_t=\sigma(W_z X_t+W_z H_{t-1}+b_z) \tag{4}$$

其中 W_r、W_z 是权重参数，b_r、b_z 是偏差参数。图1显示了GRU的典型架构。

2 基于1d-MSCNN和GRU的组合模型

2.1 模型设计思路

在处理数据时，若单纯地使用神经网络会发现其对特征的捕捉能力有限，收敛太慢，因此本文加入了改进的一维CNN模型，其中卷积部分提取的特征用于训练分类模型。虽然卷积后特征维度减少但通道数扩张，用reshape函数重新整形为向量后维度还是很高，所以要求后面的FC层仍然很稠密，收敛依旧很慢，而实际上CNN卷积移动过程就已经引入了时序关系，通道数对应RNN的时间步数，加入GRU刚好可以降维。GRU模块可以在时间级别上实现数据特征的提取。为此，本文提出一种1d-MSCNN+GRU的入侵检测模型。

2.2 多尺度一维卷积层

当前工业互联网的攻击种类繁多，并且其网络

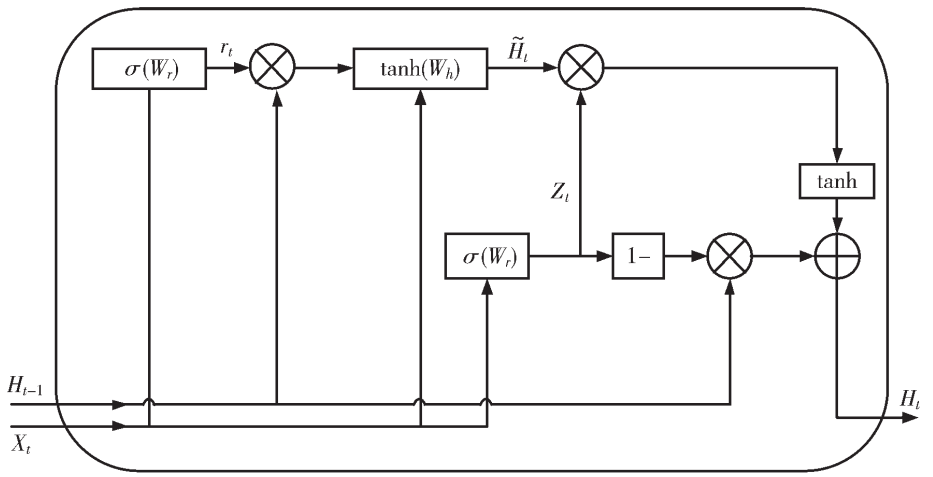

图 1 GRU 模型图

数据具有海量、高维等特征，因此单一尺度卷积神经网络在提取特征时容易造成部分特征遗漏这一问题。本文设计了一维多尺度卷积层(1-dimensional Multiscale Convolution，1d-MC)来取代单一尺度卷积层，其结构如图 2 所示。

该模块由三个分支组成，每个分支所使用的卷积核尺度都和其他卷积核尺度不同，这样就能从上一层输出中提取到多尺度的特征，然后将每个分支卷积得到的特征向量进行拼接作为下一层的输入。其运算过程如式(5)所示，其中 C^l 为上层输出激活值，C^{l+1} 为下一层的激活值，$W^A_{1\times1}$ 和 $b^A_{1\times1}$ 的上标表示所属分支，下标表示卷积核或偏置矩阵的大小。

$$C^{l+1}=\begin{cases} f(C^l W^A_{1\times1}+b^A_{1\times1}) \\ f(f(C^l W^B_{1\times1}+b^B_{1\times1})W^B_{3\times1}+b^B_{3\times1}) \\ f(f(C^l W^C_{1\times1}+b^C_{1\times1})W^C_{5\times1}+b^C_{5\times1}) \end{cases} \tag{5}$$

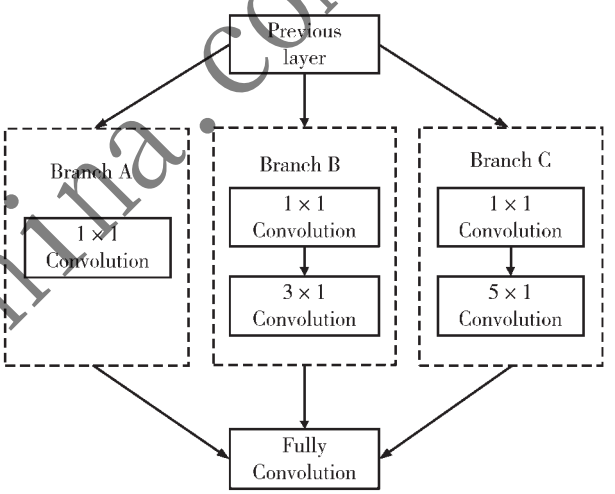

图 2 一维多尺度卷积层

2.3 1d-MSCNN+GRU 模型结构

本文搭建的一维多尺度卷积神经网络与 GRU 混合模型的结构如图 3 所示。

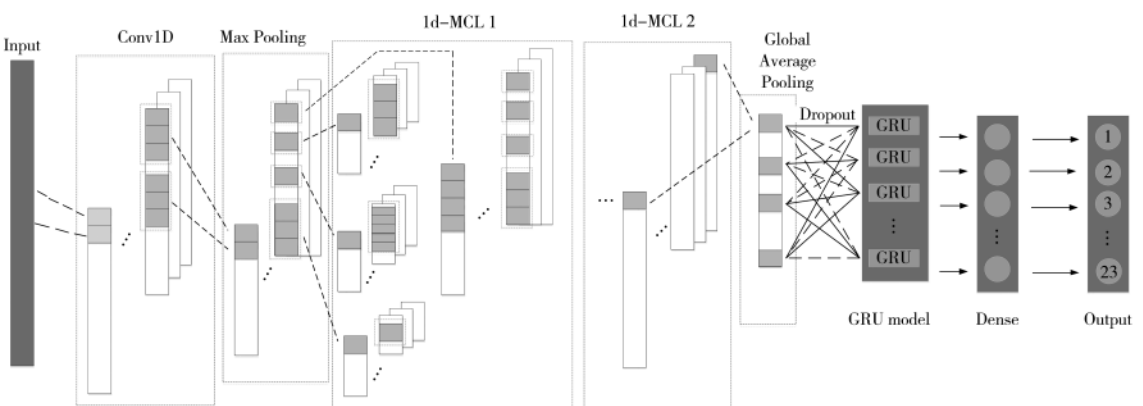

图 3 1d-MSCNN+GRU 组合模型图

1d-MSCNN+GRU 模型由一维多尺度卷积部分、GRU 部分和输出部分组成。模型的第 1 层为输入层，将预处理好的数据传递到下一层；第 2 层和第 3 层为常规的一维卷积层和池化层，能迅速缩短向量长度，同时增加通道数；模型的第 4、5 层为多尺度一维卷积层，该层利用网络层堆叠的方式逐步提取数据集不同粒度的特征；模型的第 6 层为全局平均池化层，该层减少数据通道便于数据输入到后面的 GRU 模块中；第 7 层为 Dropout 层防止模型过拟合；GRU 模块在模型的第 8 层；紧接着是全连接层将各个分类结果输出。模型各层参数如表 1 所示。

表 1 1d-MSCNN+GRU 网络结构与参数

	输入	卷积核			输出大小
		大小	步长	数目	
1	Input				40×1
2	Conv1D	3×1	1×1	16	40×16
3	Max Pooling	3×1	2×1	16	20×16
4	1d-MCL 1	1×1, 3×1, 5×1	2×1, 2×1, 2×1	32, 64, 32	20×128
5	1d-MCL 2	1×1, 3×1, 5×1	2×1, 2×1, 2×1	32, 64, 32	20×128
6	Global Average Pooling				128×1
7	Dropout				64×1
8	GRU				32×1
9	FC/Output				23×1

3 实验与结果

本文实验共分为三个阶段，数据预处理阶段、模型训练阶段和测试分类阶段，总体步骤如图 4 所示。

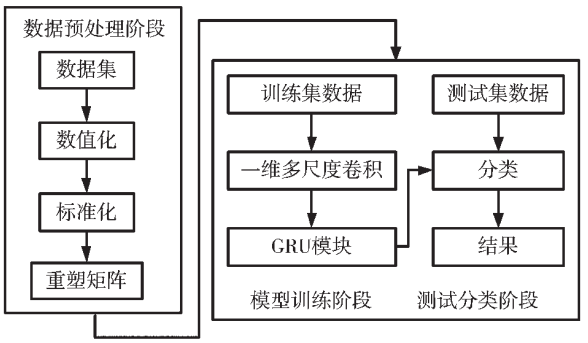

图 4 实验流程图

3.1 数据集描述

本次实验采用两个数据集，分别是 KDD CUP 99 数据集[14]和密西西比州大学的天然气管道数据集[15]。KDD CUP 99 数据集是入侵检测中常用的数据集，本文采用其中 10%的训练子集进行算法测试。样本类别分布如表 2 所示。

表 2 KDD CUP 99 数据集类别

KDD	80%KDD(Train)	20%KDD(Test)
ALL Types	395 216	98 805
Normal	77 822	19 455
DoS	313 166	78 291
Probe	3 285	821
R2L	900	225
U2R	41	10

密西西比州大学的天然气管道数据集包括 1 种正常数据和 7 种攻击类型，共有 26 个特征属性。相比于 KDD CUP 99 数据集，密西西比州采集的数据是工业网络中采集到的数据，其维度更高，攻击类型更多，其数据集标签及描述如表 3 所示。

表 3 密西西比州数据集描述

标签	描述	数量
Normal	正常行为(Instance not part of an attack)	61 156
NMRI	简单恶意响应注入(Naive malicious response injection attack)	2 763
CMRI	复杂恶意响应注入(Complex malicious response injection attack)	15 466
MSCI	恶意状态命令注入(Malicious state command injection attack)	782
MPCI	恶意参数命令注入(Malicious parameter command injection attack)	7 637
MFCI	恶意功能命令注入(Malicious function command injection attack)	573
DoS	拒绝服务攻击(Denial-of-Service attack)	1 837
Recon	侦察攻击(Reconnaissance attack)	6 805

3.2 数据预处理

由于数据集中的字符型特征属性无法直接输入到本文模型中，因此先将 KDD CUP 99 数据集和密西西比州大学的天然气管道数据集中的每一个字符型特征都转成数字以便于处理；接着为了缩小数值之间的差距，对数据进行标准化处理，将每个数据转换为一维矩阵，使其符合一维多尺度卷积模型的输入格式；最后对数据集进行标准化处理，形成重塑矩阵。

3.3 模型评价标准

本文实验中采用分类准确率(Accuracy, ACC)、召回率(Recall, R)、精确率(Precision, P)和综合评价指标(F1-Measure, F)判断模型分类,具体算法如下:

$$ACC=\frac{正确分类的样本数目}{总样本数}\times 100\% \tag{6}$$

$$R=\frac{预测为正确的真实正确样本数目}{真实正确的样本总数}\times 100\% \tag{7}$$

$$P=\frac{预测为正确的真实正确样本数目}{预测为正确的样本总数} \tag{8}$$

$$F=\frac{2\times P\times R}{P+R} \tag{9}$$

3.4 实验环境及参数设置

本实验的仿真是在 Windows 10 操作环境下进行的,CPU 为 Intel Core i7-10700。

为了实现高效计算,本文使用 Adam 优化算法,因为 Adam 的默认参数可以解决绝大部分的问题,所以本文使用默认值。其他实验参数设置如下:学习率设置为 0.01,此时模型的学习状态最佳;Dropout 失活率设置为 0.2,防止模型过拟合;迭代次数由以下仿真实验获得。在实验中改变了训练迭代次数,仿真结果如图 5 所示。

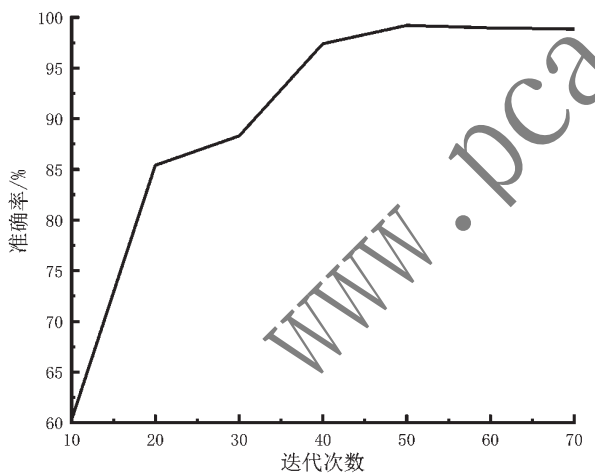

图 5 不同迭代次数准确率对比图

由图 5 看出,当迭代次数选取为 50 次时,准确率达到了最高。

3.5 KDD CUP 99 数据集实验结果分析

(1)入侵检测模型实现

将预处理得到的数据输入到 1d-MSCNN+GRU 模型中,实验结果如表 4 所示。

为了体现 1d-MSCNN+GRU 模型训练时间的优越性,本文将 1d-MSCNN+GRU 模型与传统的 CNN 模型进行了比较,结果如表 5 所示,其中可以看出 1d-MSCNN+GRU 模型的训练时间和测试时间都优于传统的 CNN 模型。

表 4 基于 1d-MSCNN+GRU 模型各子类检测结果

Label	Precision	Recall	F1-Measure	Sample
0	0.96	0.99	0.98	19 409
1	0.00	0.00	0.00	4
2	0.00	0.00	0.00	2
3	0.00	0.00	0.00	1
4	0.99	0.99	0.99	21 462
5	0.99	0.99	0.99	56 158
6	0.00	0.00	0.00	14
7	0.00	0.00	0.00	58
8	0.90	0.96	0.93	213
9	0.91	0.63	0.75	213
10	0.81	0.72	0.76	213
11	0.00	0.00	0.00	227
12	0.00	0.00	0.00	7
13	0.00	0.00	0.00	3
14	0.00	0.00	0.00	2
15	0.95	0.82	0.88	305
16	0.00	0.00	0.00	1
17	0.00	0.00	0.00	47
18	0.00	0.00	0.00	0
19	0.00	0.00	0.00	1
20	0.00	0.00	0.00	198
21	0.00	0.00	0.00	0
22	0.00	0.00	0.00	2

表 5 两种模型所用的时间对比 (s)

模型	训练时间	测试时间
1d-MSCNN+GRU	134.52	24.34
CNN	351.09	32.86

(2)不同的机器学习入侵检测模型

本文将 1d-MSCNN+GRU 模型与单纯 CNN 模型和 GRU 模型进行了对比,结果如表 6 所示。

由表 6 可以看出,1d-MSCNN+GRU 模型准确率

表 6 不同算法的分类结果 (%)

模型	准确率	召回率	F1 值
1d-MSCNN+GRU	99.16	98.34	98.30
CNN	96.25	95.91	95.13
GRU	94.96	94.31	94.56

为 99.16%，召回率为 98.34%，F1 值为 98.30%，明显高于 CNN 模型和 GRU 模型，其中 1d-MSCNN+GRU 模型的 F1 值比 CNN 模型高出 3.17%，比 GRU 模型高出 3.74%，由此可以证明该方法的有效性。

另外，本文将 1d-MSCNN+GRU 模型与经典机器学习算法进行了比较，基于支持向量机的入侵检测算法准确率为 90.4%，召回率为 87.65%，F1 值为 91.79%。基于神经网络入侵算法的准确率为 94.98%，召回率为 95.60%，F1 值为 96.38%。基于决策树的入侵检测算法准确率为 91.64%，召回率为 96.32%，F1 值为 95.87%。图 6 展示了各个入侵检测方法的对比结果。

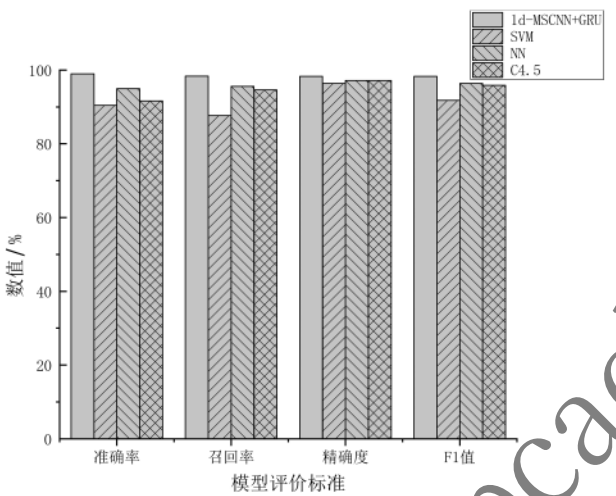

图 6 各类模型对比图

由图 6 可知，本文提出的 1d-MSCNN+GRU 模型准确率明显高于其他传统机器学习、召回率、精确度和 F1 值也优于其他算法，因而可以看出本文所提出的 1d-MSCNN+GRU 模型在收敛性方面表现最好。

3.6 密西西比州数据集实验结果分析

为了进一步验证本文模型的泛化能力，本文又对密西西比州采集的数据进行实验，使用 80%的数据进行训练，用剩余的 20%进行测试，再与其他经典机器学习算法进行比较，结果如图 7 所示。

从图 7 中可以明显看出本文提出的模型在准确率方面仍然比 CNN 和 SVM 算法优越，由此可以看出本文模型的泛化能力强。

4 结论

本文提出了一种基于一维多尺度卷积神经网络和门控循环单元相结合的入侵检测模型。该方法首先通过一维多尺度卷积神经网络有效地提取数据的高级数据特征，弥补局部特征提取不全的问题，达到充分获取输入数据中重要信息的目的；接着使用门控循环单元保存历史输入信息，使前期的输入信息与当前的输入一起映射到当前的输出，增加了精确度。实验结果表明，在使用 KDD CUP 99 数据集仿真实验时，与其他机器学习方法相比，本文模型的准确率较高、训练时间较短。而且 1d-MSCNN+GRU 模型在经过调试后，可以保存为.h5 文件，再次调用后，可以直接使用，不需要再次训练优化，而如果无法满足要求，可以继续训练优化。

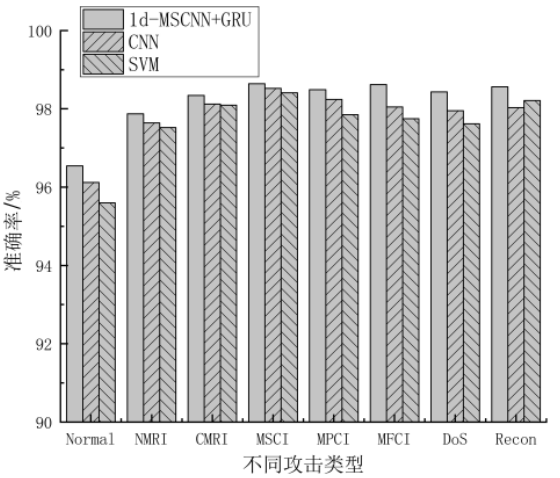

图 7 不同算法的准确率对比图

本文模型 1d-MSCNN+GRU 仅在当前两个数据集上验证了模型的检测效果，下一步将会把本文模型应用到多种数据集上，并继续优化模型，进一步提高检测效果。

参考文献

[1] 赖英旭，刘增辉，蔡晓田，等.工业控制系统入侵检测研究综述[J].通信学报，2017，38(2)：143-156.

[2] BHAVSAR Y B，WAGHMARE K C.Intrusion detection system using data mining technique：Support vector machine[J].International Journal of Emerging Technology and Advanced Engineering，2013，3(3)：581-586.

[3] KUANG F J，XU W H，ZHANG S Y.A novel hybrid KPCA and SVM with GA model for intrusion detection[J].Applied Soft Computing Journal，2014，18(5)：178-184.

[4] KUMAR V，CHAUHAN H，PANWAR D.K-means clustering approach to analyze NSL-KDD intrusion detection dataset[J].International Journal of Soft Computing and Engineering(IJSCE)，2013，3(4).

[5] 赵会群,刘金銮.基于贝叶斯网络的复杂事件大数据处理系统测试数据生成方法研究[J].计算机应用研究,2018,35(8):155-158,162.

[6] YAO Y,WEI Y,GAO F L,et al. Anomaly intrusion detection approach using hybrid MLP/CNN neural network[C]//Proceedings of the 6th International Conference on Intelligent Systems Design and Applications. Piscataway:IEEE,2006:1095-1102.

[7] 谭敏生,彭敏,丁琳,等.基于遗传的CNN优化方法在入侵检测中的应用[J].计算机仿真,2021,38(2):416-421.

[8] ZHANG Y,CHEN X,JIN L,et al.Network intrusion detection:based on deep hierarchical network and original flow data[J].IEEE Access,2019,7:37004-37016.

[9] 张全龙,王怀彬.基于膨胀卷积和门控循环单元组合的入侵检测模型[J].计算机应用,2021,41(5):1372-1377.

[10] 杭梦鑫,陈伟,张仁杰.基于改进的一维卷积神经网络的异常流量检测[J].计算机应用,2021,41(2):433-440.

[11] GUO S,YANG T,GAO W,et al.A novel fault diagnosis method for rotating machinery based on a convolutional neural network[J].Sensors,2018,18(5):1-16.

[12] 陈土生.基于Adam优化GRU神经网络的SCADA系统入侵检测方法[J].现代计算机,2019(15):13-19.

[13] BENGIO Y,SIMARD P,FRASCONI P.Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactionson Neural Networks and Learning Systems,2002,5(2):157-166.

[14] STOLFO S J,FAN W,LEE W,et al.Cost-based modeling for fraud and intrusion detection:results from the JAM project[C]//Proceedings DARPA Information Survivability Conference and Exposition, DISCEX′00.IEEE,2000,2:130-144.

[15] MORRIS T,WEI G.Industrial control system traffic data sets for intrusion detection research[C]//International Conference on Critical Infrastructure Protection,2014.

(收稿日期:2021-04-15)

作者简介:

宗学军(1970-),男,硕士,教授,主要研究方向:工业过程控制、工业信息安全等。

宋治文(1997-),通信作者,男,硕士研究生,主要研究方向:工业信息安全。E-mail:1500270440@qq.com。

何戡(1978-),男,硕士,副教授,主要研究方向:工业过程控制、机器学习等。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部

中国电子信息产业集团有限公司第六研究所