计算机应用
Journal of Computer Applications
ISSN 1001-9081,CN 51-1307/TP

《计算机应用》网络首发论文

题目：基于膨胀卷积和门控循环单元组合的入侵检测模型
作者：张全龙，王怀彬
收稿日期：2020-07-23
网络首发日期：2020-10-12
引用格式：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．计算机应用. https://kns.cnki.net/kcms/detail/51.1307.TP.20201011.1418.002.html

Journal of Computer Applications ISSN 1001-9081
计算机应用 CODEN JYIIDU http://www.joca.cn

doi:10.11772/j.issn.1001-9081.2020071082

基于膨胀卷积和门控循环单元组合的入侵检测模型

张全龙，王怀彬*

(天津理工大学 计算机科学与工程学院，天津 300384)

(*通信作者电子邮箱 hbwang@tjut.edu.cn)

摘 要: 基于机器学习的入侵检测模型在网络环境的安全保护中起着至关重要的作用。针对现有的网络入侵检测模型不能够对网络入侵数据特征进行充分学习的问题，本文将深度学习理论应用于入侵检测，提出了一种具有自动特征提取功能的深度网络模型。在该模型中，使用膨胀卷积来增大对信息的感受野，从中提取高级特征，使用门控循环单元(GRU)模型提取保留特征之间的长期依赖关系，再利用深层神经网络对数据特征进行充分学习。与经典的机器学习分类器相比，该模型具有很高的检测率。对著名的 KDD CUP99、NSL-KDD 和 UNSW-NB15 数据集进行的实验表明该模型具有领先的性能。使用 KDD CUP99 数据集的准确率为 99.78%，使用 NSL-KDD 数据集的准确率为 99.53%，使用 UNSW-NB15 数据集的准确率为 93.12%。

关键词: 网络入侵检测模型；深度学习；门控循环单元；膨胀卷积；网络安全

中图分类号:TP393.08 **文献标志码：A**

Intrusion detection model based on dilated convolution and gated recurrent unit

ZHANG Quanlong, WANG Huaibin*

(College of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China)

Abstract: Intrusion detection model based on machine learning plays a vital role in the security protection of the network environment. Aiming at the problem of the existing network intrusion detection model cannot fully learn the data features of network intrusion data, deep learning theory was applied to intrusion detection, and a deep network model with automatic feature extraction function was proposed. In this model, dilated convolution was used to increase the receptive field of information and extract high-level features, a gated recurrent unit (GRU) model was used to extract long-term dependencies between retained features, then deep neural networks was used to fully learn the data features. Compared with the classical machine learning classifier, this model has a high detection rate. Experiments conducted on the famous KDD CUP99, NSL-KDD and UNSW-NB15 data sets show that the model has leading performance. The accuracy using the KDD CUP99 data set is 99.78%, the accuracy using the NSL-KDD data set is 99.53%, and the accuracy using the UNSW-NB15 data set is 93.12%.

Keywords: network intrusion detection model; deep learning; gated recurrent unit; dilated convolution; network security

0 引言

计算机技术在飞速的发展，它已经渗透到人们的工作和日常生活中，给人们带来了极大的便利。但同时网络攻击变得越来越频繁，网络安全已成为人们必须面对的挑战[1-3]。而入侵检测技术是分类问题，需要做的事情是建立入侵检测模型，使其能够有效的对各种网络攻击进行分类识别，以便及时的采取安全防范措施。

当前，广为人知的入侵检测方法之一是使用不同的机器学习技术来降低错误率。例如 K 最近邻(K-Nearest Neighbor, KNN)[4]，神经网络(Neural Network, NN)[5]和支持向量机(Support Vector Machine, SVM)[6]已广泛用于入侵检测。文献[7]中的作者提出了一种基于 KNN 回归的动态多间隔预测模型。文献[8]提出了一种混合的机器学习技术，结合了 K-means 和 SVM 来检测攻击。集成分类器(例如 Adaptive Boosting [9]，Random Forest [10])通常由多个弱分类器构成，避免在训练过程中过度拟合，可以实现更强大的分类功能。Al-Yaseen 等[11]使用 SVM 和改进的 K-means 算法来构建多层混合入侵检测模型。但是，对于 KDD CUP 99 数据集中 U2R(User to Root)和 R2L(Remote to Local)的低频攻击样本，此模型的检测率非常低，远低于其他高频样本的检测率。尽管基于机器学习技术的网络入侵检测模型具有强大的检测能力和适应网络环境变化的自适应能力，但是它们仍然受到不平衡的数据的影响。

收稿日期: 2020-07-23; 修回日期:2020-09-05; 录用日期: 2020-09-27。
基金项目: 国家自然科学基金资助项目(61773286)。
作者简介: 张全龙(1994—)，男，安徽宿州人，硕士研究生，CCF 会员，主要研究方向：网络信息安全； 王怀彬(1960—)，男，天津人，教授，博士生导师，主要研究方向：网络信息安全、计算机软件。

卷积神经网络(Convolutional Neural Networks, CNN)[12]是深度学习研究的重点，在计算机视觉，语音识别和自然语言处理方面取得了出色的研究成果。与传统的特征选择算法相比，它可以自动学习更好的特征。文献[13]首先使用不同的降维方法去除了多余的特征，然后将降维数据呈现给 CNN 网络。尽管获得了很好的准确性，但它掩盖了 CNN 的优势：自动提取特征。门控循环单元(Gated Recurrent Unit, GRU)是递归模型，已用于如自然语言处理和情感分析等序列学习。文献[14]实现了网络入侵检测模型的 GRU，MLP(Multi-Layer Perceptron)和 Softmax 模块，他们在 KDD CUP99 和 NSL-KDD 训练数据集上进行了实验。文献[15]进一步提出了通过一维卷积来统一和共享多传感器权重的问题。文献[16]设计结合了 CNN 和长期短期记忆网络(Long Short Term Memory, LSTM)的模型。文献[17]构建了一个新的 DNN(Deep Neural Networks)模型，该模型使用 GRU 和 MLP 提取数据信息，他们的仿真结果表明，GRU 单元在入侵检测方面比 LSTM 单元更有效。

为了提高检测的分类准确度，本文模型通过膨胀卷积来增强感受野的同时，提取增强的特征。并且使用 GRU 模型来挖掘数据样本之间的时间序列信息。本文提出的模型最大优点是可以准确的提取数据的特征，并且检测到以前从未见过的攻击。

本文的模型主要贡献如下：

（a）使用膨胀卷积来增大感受野，以此来提高模型对特征获取的准确度，对数据样本进行充分的学习。

（b）使用 GRU 神经网络来获取数据之间的时间关系的特征，以此来检测未知的攻击。

（c）使用随机梯度下降(Stochastic Gradient Descent, SGD)优化算法用于协助训练模型，并且使用动量法来增加 SGD 更新的稳定性。

1 膨胀卷积

在传统的卷积神经网络中，会使用池化层来保持特征不变性并避免过度拟合，但是会大大降低空间分辨率，会丢失特征图的空间信息。当加深卷积神经网络的层时，网络需要更多的参数，并导致更多的计算资源消耗。Yu 等[18]提出的膨胀卷积很好的解决了这一问题。膨胀卷积是一种卷积算子，它使用不同的膨胀因子在不同范围使用相同的滤波器。膨胀卷积能够更有效地扩展感受野。与传统卷积相反，膨胀卷积的内核中存在孔，孔的大小为膨胀率。一维卷积的公式如下：

$$(f * w)[t] = \sum_{p=-q}^{q} f[t-p]w[p] \qquad (1)$$

其中 f 为输入，w 为卷积核，t 为卷积核的大小，p 为卷积的下限值，q 为卷积的上限值。如果是膨胀卷积，则一维膨胀卷积的公式如下：

$$(f *_l w)[t] = \sum_{p=-q}^{q} f[t-l*p]w[p] \qquad (2)$$

其中 l 是膨胀率。本文对输入数据应用膨胀卷积时，与传统卷积相比，感受野将得到扩展，而不会降低分辨率。能够在不增加参数数量或计算量的情况下增大感受野，换句话说，本文使用的相同层数可以实现更大的感受野，而无需引入比普通卷积更多的操作。膨胀卷积是一个将步幅进行卷积的元素分开的卷积过程，与传统的卷积相比，膨胀卷积是到较宽区域的稀疏连接。本文堆叠 3 个具有不同步幅的膨胀卷积层，尽管膨胀卷积具有与常规卷积相同的过滤器大小，但是通过堆叠它们，可以感知更大的范围。

本文设计的膨胀卷积模型如图 1 所示。该模型具有三个膨胀卷积层，每个卷积层的膨胀率分别为 2、4、8。当膨胀率为 2 时，膨胀卷积过后特征集中神经元数量为 32；当膨胀率为 4 时，膨胀卷积过后特征集中神经元数量为 64；当膨胀率为 8 时，膨胀卷积过后特征集中神经元数量为 128。这样本文提出的模型可以从原始数据中提取尽可能多的特征，并且可以得到神经元数量分别为 32、64、128 的特征集。经过三个膨胀卷积层后，特征集可以获取原始数据包含的所有信息。在每个卷积层之后，都有一个 Relu 激活层，用于为模型添加非线性特征。本文不在每个卷积层之后都使用 max-pooling 层，而是在三个膨胀卷积层之后加入 max-pooling 层来防止过拟合。

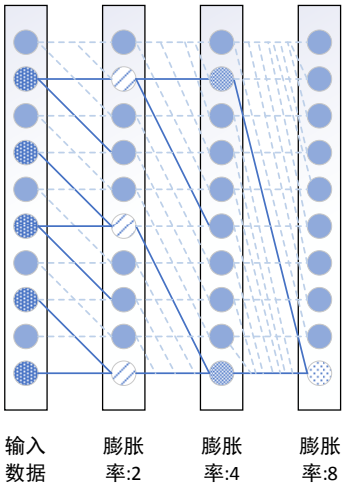

图 1 膨胀卷积模型

Fig. 1 Dilated convolution model

本文使用多级膨胀卷积神经网络捕获数据之间的局部相关性和长期依赖性。具体来说，本文的卷积神经网络是三级膨胀卷积神经网络。它能够以指数形式扩展接受域级别而不增加参数数量，因此，膨胀卷积捕获长期依赖性成为可能。本文使用具有不同膨胀率的多级膨胀卷积，这样做避免了由膨胀导致的重要局部相关性缺失，也能使输入的所有数据都能够参与计算。

2 门控循环单元（GRU）

传统的深层神经网络(DNN)在样本分类和特征提取方面突破了浅层网络的局限性，并且具有强大的非线性拟合能力。然而，DNN 没有考虑分类样本之间的时间关系，导致分类过程中一些信息的丢失。循环神经网络(Recurrent Neural Network, RNN) [19]有效地解决了时序依赖性问题。RNN 引入了隐藏层单元之间的反馈连接，以便网络可以将学习到的信息保留到当前时刻，并确定网络的最终输出结果以及当前时刻的输入。但是，RNN 无法学习导致梯度消失的长期依赖关系[20]。许多用于改善 RNN 的网络结构，其中最广泛使用和有效的结构之一是 LSTM，但是 LSTM 中有很多参数，并且需要花费更多的时间来将模型参数调整为最佳状态。与 LSTM 相比，GRU 的门更少，可以节省更多的训练时间和计算资源。图 2 显示了 GRU 的典型架构。

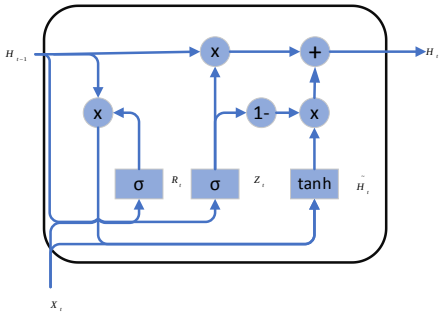

图 2 GRU 模型

Fig. 2 GRU model

GRU 中有两个主门，即更新门和重置门。更新门用于控制将多少先前状态信息带入当前状态。重置门用于控制 GRU 忽略前一时刻的状态信息的程度。所有的关系定义如下：

（a）重置门：

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

（b）更新门：

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

门控循环单元中的重置门和更新门的输入均为当前时间步输入 X_t 与上一时间步隐藏状态 H_{t-1}，输出由激活函数为 sigmoid 函数的全连接层计算得到。其中 W_r，W_z 是权重参数，b_r，b_z 是偏差参数。

（c）候选状态：

$$\tilde{H}_t = \tanh(W_h X_t + W(R_t * H_{t-1}) + b_h) \quad (5)$$

（d）隐藏状态：

$$H_t = (1 - Z_t) * \tilde{H}_t + Z_t * H_{t-1} \quad (6)$$

其中 W_h，W 是权重参数，b_h 是偏差参数。门控循环单元将计算候选状态来辅助稍后的隐藏状态的计算。将当前时间步重置门的输出与上一时间步隐藏状态做按元素乘法。如果重置门中元素值接近 0，那么意味着重置对应隐藏状态元素为 0，即丢弃上一时间步的隐藏状态。如果元素值接近 1 那么表示保留上一时间步的隐藏状态。这个设计可以应对循环神经网络中的梯度衰减问题，并更好地捕捉时间序列中时间步距离较大的依赖关系。

3 基于膨胀卷积和门控循环单元组合模型

本文通过将膨胀卷积与 GRU 模型结合在一起形成新模型来提取数据的特征，两者的结合构成了一个深层网络，可以实现更优化的结果。其结构如图 3 所示。

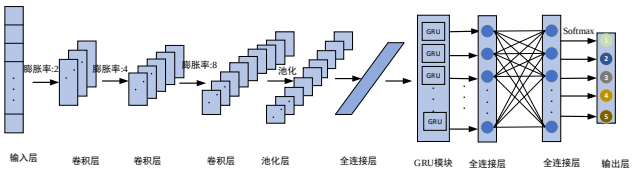

图 3 基于膨胀卷积和 GRU 的组合模型

Fig. 3 Combined model based on dilated convolution and GRU

该模型由膨胀卷积部分，GRU 部分，全连接层部分和输出部分组成。由于膨胀卷积和 GRU 网络结构的输入形式不同，因此提取的空间特征会在膨胀卷积部分的输出处进行调整，保证 GRU 部分输入的大小与膨胀卷积的输出大小一致，以满足 GRU 模型的输入格式。在 GRU 模型的输出层之后连接一个全连接层，对先前提取的特征进行集成，最后一个全连接的层的输出值传递给 Softmax 进行分类。模型各层参数如表 1 所示。

表 1 模型中各层结构参数

Tab. 1 Parameters of each layer in the model

层数	类型	参数
1	Dilated rate =2	32
2	Dilated rate =4	64
3	Dilated rate =8	128
4	Max pooling	128
5	GRU	128
6	Dense	48
7	Softmax	5

膨胀卷积部分提取的特征用于训练分类模型，考虑到特征在不同位置具有局部性，因此在三层膨胀卷积之后使用池化层，在一定程度上汇总不同位置的统计信息，将小邻域中的特征点集成以获得新特征，以减少数据量并避免过度拟合。经过膨胀卷积和合并后，使用 reshape 函数重新整形为向量。然后，可以通过全连接层获得输出，这样就可以得到膨胀卷积提取的空间特征。膨胀卷积的使用可以准确地提取空间特征，但在学习序列相关信息时效果不佳。因此，需要提高仅使用膨胀卷积的网络入侵检测的准确性，本文加入了 GRU 模型。膨胀卷积和 GRU 模型都代表了深度学习算法，膨胀卷积可以提取空间维度中的数据特征，并且增大感受野，GRU 具有可以长时间保存上下文历史信息的特性，并且可以在时间级别上实现数据特征的提取。

4 实验结果和分析

本文实验的总体步骤如图 4 所示。使用提出的模型提取数据的特征，以提高分类的准确性。训练后，获得了具有良好分类性能的模型，并使用该模型对测试集进行分类，以获得优异的分类结果。本文实验使用的 CPU 为 Intel Core i7-7700、GPU 为 GeForce GT 730、操作系统为 Windows 10，内存为 16 GB。

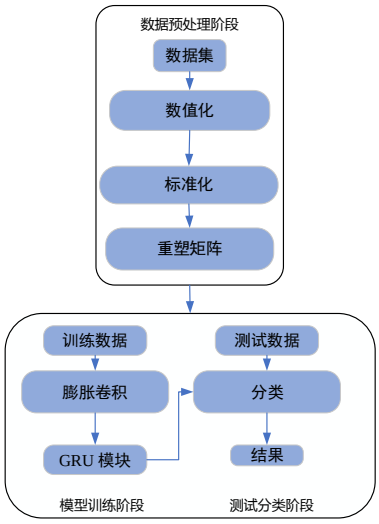

图 4 实验总体步骤

Fig. 4 General steps of the experiment

本文使用随机梯度下降(SGD)优化算法，经过多次和小范围实验训练，实验参数设置如下：

学习率设置为 0.01，此时的模型的学习状态最佳；

权值衰减系数为 0.000001，此时模型的复杂度对损失函数影响最小；

动量(momentum)设置为 0.9，此时 SGD 的稳定性最好；

正则化方法 Dropout 失活率设置为 0.2。

4.1 数据描述

本文使用三个公开可用的入侵检测数据集 KDD CUP99、NSL-KDD 和 UNSW-NB15 数据集。在入侵检测领域，KDD CUP99 和 NSL-KDD 是著名的数据集[21]，两个数据集中，每个入侵记录都具有 42 维特征，标签主要包含普通数据和 4 种攻击数据 Dos(Denial of Service)、Probe、U2R、R2L。UNSW-NB15 数据集包含许多现代网络的新攻击，可以将其分为 1 个正常类和 9 个攻击类。在本文的实验中，KDD CUP99、NSL-KDD、UNSW-NB15 数据集中样本类别分布如表 2 所示。

表 2 数据集样本分布

Tab. 2 Date set sample distribution

数据集	攻击类型	训练集	测试集
KDD CUP99	Normal	97278	60593
	Dos	391458	229853
	Probe	4107	4166
	U2R	52	228
	R2L	1126	16189
NSL-KDD	Normal	97278	60593
	Dos	391458	229853
	Probe	4107	4166
	U2R	52	228
	R2L	1126	16189
UNSW-NB15	Normal	56000	37000
	Reconnaissance	10491	3496
	Backdoor	1746	583
	Worms	130	44
	Analysis	2000	677
	Shellcode	1133	378
	Generic	40000	18871
	Fuzzers	18184	6062
	Dos	12264	4089
	Exploits	33393	11132

4.2 数据预处理

本文对数据集中的字符型特征属性进行数字化和标准化，得到一个标准化的数据集，然后将每个数据转换为二维矩阵，使其符合膨胀卷积模型的输入格式。处理后的数据集有训练数据集和测试数据集。训练数据集用来训练网络模型，测试数据集用来验证模型的有效性。由于数据特征的复杂性，数据预处理包括以下三步：

（a）数值化处理

由于模型的输入是数字矩阵，因此使用 one-hot 编码方法将数据集中具有符号特征的数据映射到数字特征向量。将 KDD CUP99 和 NSL-KDD 数据集中正常数据(Normal)和 4 种攻击类型(Dos, Probe, U2R, R2L)这 5 种类标签进行数值化处理，也对 UNSW-NB15 数据集中正常数据(Normal)和 9 种攻击类型(Reconnaissance, Backdoor, Worms, Analysis, Shellcode, Generic, Fuzzers, Dos, Exploits)这 10 种类标签进行数值化处理。

（b）标准化处理

在数据集中，不同类别的数据值大小明显不同，最大值的范围变化很大。为了便于算术处理和消除尺寸，采用归一化处理方法，在[0,1]区间内均匀且线性地映射每个特征的值范围。用以下方程式给出的线性函数将数值数据归一化为[0,1]：

$$x^* = \frac{x - min}{max - min} \tag{7}$$

其中 max 为样本数据的最大值，min 为样本数据的最小值，x 为标准化后的数据。

（c）将标准化数据转换为矩阵

读取数据的每个网络记录都将进行尺寸转换以符合网络模型的格式。为了输入到膨胀卷积神经网络中，将使用 reshape 转换函数将网络数据重塑为矩阵。

4.3 评估指标

在本文中，Accuracy，Precision，Recall 和 F1-measure 被用作评估模型性能的关键指标。这些指标是从混淆矩阵的四个基本属性中得出的，如表 3 所示。

表 3 混淆矩阵

Tab.3 Confusion matrix

真实标签	预测标签	
	Attack	Normal
Attack	True positive(TP)	False Negative(FN)
Normal	False Positive(FP)	True Negative(TN)

（a）True Positive(TP)-攻击数据被正确的分类为攻击。

（b）False Positive(FP)-正常数据被错误的分类为攻击。

（c）True Negative(TN)-正常数据被正确的分类为正常。

（d）False Negative(FN)-攻击数据被错误的分类为正常。

本文将使用以下评估指标来评估本文所提出模型的性能。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$\Pr ecision = \frac{TP}{TP + FP} \quad (9)$$

$$\mathrm{Re}\ call = \frac{TP}{TP + FN} \quad (10)$$

$$F1 - measure = \frac{2 \cdot \Pr ecision \cdot \mathrm{Re}\ call}{\Pr ecision + \mathrm{Re}\ call} \quad (11)$$

4.4 KDD99 数据集的实验结果

实验包括训练和测试两个过程。使用 KDD CUP99 数据集的训练集和测试集来进行实验。使用本文提出的模型用训练集数据进行训练，最后使用测试集对该模型进行测试。

五个标签类的评估指标值通过图 5 条形图可以被更清楚地观察到，低频样本 U2R、R2L 类在本文提出的模型下，Precision、Recall、F1-measure 三个评估指标依然拥有较高的值。

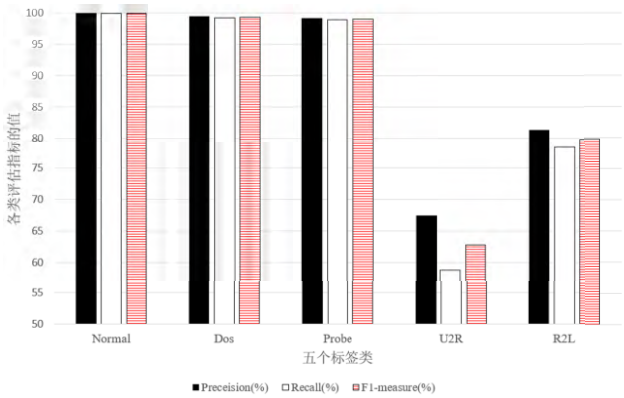

图 5 KDD CUP99 数据集下五个标签类的评估指标值

Fig.5 Evaluation index values of five label classes under KDD CUP99 data set

表 4 KDD CUP99 数据集未知攻击类型的检测结果

Tab.4 KDD CUP99 data set unknown attack type detection results

标签类	未知攻击类	样本数	正确检测数	Recall/%
Dos	apache2	794	761	95.84
	mailbomb	5000	4891	97.82
	processtable	759	728	95.92
	udpstorm	2	0	0
Probe	mscan	1053	1001	95.06
	saint	7367	7198	97.71
U2R	httptunnel	158	87	55.06
	ps	16	7	43.75
	sqlattack	2	0	0
	xterm	13	4	30.77
R2L	named	17	11	64.71
	sendmail	17	8	47.06
	snmpgetattack	7741	4554	58.83
	snmpguess	2406	976	40.57
	worm	2	0	0
	xlock	9	1	11.1
	xsnoop	4	0	0

为了评估提出模型对未知攻击的检测效果，使用了 KDD CUP99 数据集中 17 种未知攻击类型，这 17 种未知攻击类型存在于测试集中，而在训练集中不存在。这 17 种未知攻击的召回率如表 4 所示。这证明了本文提出的模型可以对未知攻击进行检测。

目前，许多机器学习和深度学习算法已应用于网络入侵检测。支持向量机和经典卷积神经网络广泛用于网络入侵检测，因此，将入侵检测中常用的经典分类模型与本文中的模型进行了比较。本文使用 SVM[22]，S-NDAE(Stacked Nonsymmetric Deep Autoencoder)[23]，和 MHCVF(Multilevel Hybrid Classfier with Variant Feature sets)[24]模型和本文提出的模型在 KDD CUP99 数据集上对分类性能进行了比较，如表 5 所示。

表 5 使用 KDD CUP99 数据集的各模型实验结果对比

Tab. 5 Comparison of experimental results of each model using KDD CUP99 data set

模型	Accuracy/%	Recall/%
SVM	94.22	92.99
S-NDAE	97.85	97.85
MHCVF	98.04	95.57
本文模型	**99.78**	**99.33**

从表 5 可以看出，与传统的 SVM、S-NDAE、MNCVF 模型相比，本文提出的模型测试结果最好，准确率达到 99.78%，召回率达到 99.33%。当面对复杂数据时，从分类结果可以看出，本文提出的模型仍然比其他模型获得更高的准确率。

4.5 NSL-KDD 数据集的实验结果

为了进一步验证本文提出的模型，本文还对 NSL-KDD 数据集进行了实验。使用本文提出的模型用训练集数据进行训练，最后使用测试集对该模型进行测试。

各指标的分布通过图 6 条形图可以被更清楚地看到。本文同样选取入侵检测中常用的经典分类模型与本文中提出的模型进行了比较。本文使用 SCDNN(Spectral Clustering Deep Neural Network)[25]，DNN[26]，和 SMOTE+ CANN(Synthetic Minority Oversampling Technique and Cluster Center And Nearest Neighbor)[27]模型和本文提出的模型在 NSL-KDD 数据集上对分类性能进行了比较，比较结果如表 6 所示。

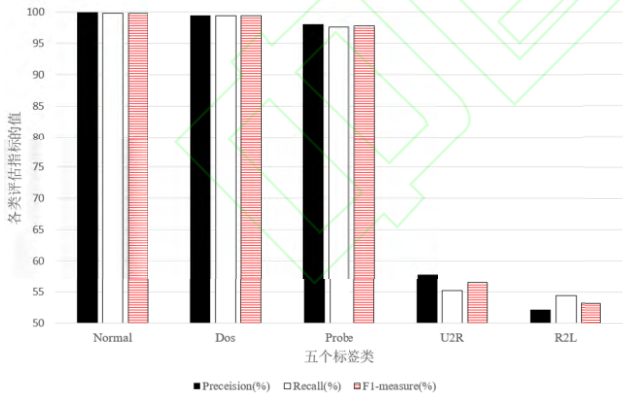

图 6 NSL-KDD 数据集下五个标签类的评估指标值

Fig. 6 Evaluation index values of five label classes under NSL-KDD data set

表 6 使用 NSL-KDD 数据集的各模型实验结果对比

Tab. 6 Comparison of experimental results of each model using NSL-KDD data set

模型	Accuracy/%	Recall/%
SCDNN	92.03	92.23
SMOTE+CANN	98.99	**99.56**
DNN	99.20	99.27
本文模型	**99.53**	99.25

从表 6 可以看出，与其他 SCDNN、SMOTE+CANN 和 DNN 三种分类器相比，本文模型准确率可以达到 99.53%，召回率达到 99.25%。从图 6 可以看出，本文提出的模型在 Precision，Recall 和 F1-measure 几个评价标准上得到的结果都很高。从分类结果可以看出，该模型是有效的，当面对复杂数据时，本文提出的模型仍然比其他模型获得更好的结果。

4.6 UNSW-NB15 数据集的实验结果

UNSW-NB15 数据集中包含许多现代网络的新攻击，使用本文提出的模型用训练集数据进行训练，最后使用测试集对模型进行测试。

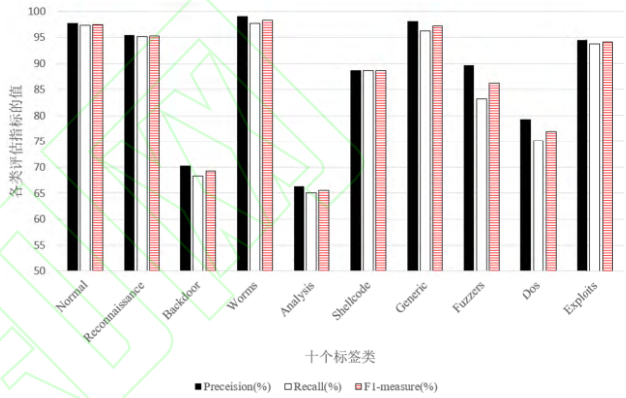

图 7 UNSW-NB15 数据集下十个标签类的评估指标值

Fig. 7 Evaluation index values of ten label classes under UNSW-NB15 data set

各指标的分布通过图 7 条形图可以被更清楚地看到。本文同样选取入侵检测中常用的经典分类模型与本文中提出的模型进行了比较。使用 RF(Random Forests)[28]，SVM[29]，和 MSCNN(Multiscale Convolutional Neural Network)[30]模型与本文提出的模型在 UNSW-NB15 数据集上对分类性能进行了比较，比较结果如表 7 所示。

表 7 使用 UNSW-NB15 数据集的各模型实验结果对比

Tab. 7 Comparison of experimental results of each model using UNSW-NB15 data set

模型	Accuracy/%
RF	80.90
SVM	85.99
MSCNN	91.40
本文模型	**93.12**

从表 7 可以看出，与 RF、SVM、MSCNN 三个模型相比，本文提出的模型有最高的检测率，可以达到 93.12%。与传统的模型相比，在新型数据集 UNSW-NB15 上进行实验时，从分类结果可以看出，本文提出的模型仍然比其他模型获得更高的准确率。

5 结语

本文提出了一种基于膨胀卷积和门控循环单元(GRU)相结合的入侵检测新模型。首先，对数据集进行数值化和标准化处理，这样可以减少模型的训练时间。然后，通过膨胀卷积和 GRU 构建的网络模型对输入数据进行分类。该模型利用深度学习的出色性能，通过重复的多级学习自动提取特征。本文使用 KDD CUP99、NSL-KDD 和 UNSW-NB15 三个入侵数据集来进行实验。根据统计显著性检验，可以得出结论，该模型优于其他分类器。当针对测试集进行验证时，所提出的模型在准确率和召回率方面产生了优异的结果。尤其是在多特征数据集中，并且发现训练数据规模越大，检测性能越好。

参考文献

[1] 孔令智. 基于网络异常的入侵检测算法研究[D]. 北京: 北京交通大学, 2017: 15-16. (KONG L Z. Research on intrusion detection algorithm based on network anomaly[D]. BeiJing: Beijing Jiaotong University, 2017: 15-16.)

[2] 沈学利, 覃淑娟. 基于 SMOTE 和深度信念网络的异常检测[J]. 计算机应用, 2018, 38(7): 1941-1945. (SHEN X L, QIN S J. Anomaly detection based on synthetic minority oversampling technique and deep belief network[J]. Journal of Computer Applications, 2018, 38(7): 1941-1945.)

[3] YIN C, ZHU Y, FEI J, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. IEEE Access, 2017, 5(99): 21954-21961.

[4] LIAO Y, VEMURI V R. Use of K-Nearest Neighbor classifier for intrusion detection[J]. Computer Security, 2002, 21(5): 439-448.

[5] MUKKAMALA S, JANOSKI G, SUNG A. Intrusion detection using neural networks and support vector machines[C]// Proceedings of the 2002 International Joint Conference on Neural Networks. Piscataway: IEEE, 2002: 1702-1707.

[6] SALLAY H, AMMAR A, SAAD M B, et al. A real time adaptive intrusion detection alert classifier for high speed networks [C]// Proceedings of the 2013 IEEE International Symposium on Network Computing & Applications. Piscataway: IEEE, 2013: 73–80.

[7] CHANG H, LEE Y, YOON B, et al. Dynamic near-term traffic flow prediction: System-oriented approach based on past experiences[J]. IET Intelligent Transport Systems, 2012, 6(3): 292-305.

[8] TAHIR H M, HASAN W, SAID A M, et al. Hybrid machine learning technique for intrusion detection system[C]// Proceedings of the 2015 International Conference on computing & informatics. Piscataway: IEEE, 2015: 11-13.

[9] HU W, GAO J, WANG Y, et al. Online adaboost based parameterized methods for dynamic distributed network intrusion detection[J]. IEEE Transactions on Cybernetics, 2013, 44(1): 66-82.

[10] ZHANG J, ZULKERNINE M, HAQUE A. Random-forests-based network intrusion detection systems[J]. IEEE Transactions on Systems, 2008, 38(5): 649-659.

[11] AL-YASEEN W L, OTHMAN Z A, NAZRI M Z A. Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system[J]. Expert Systems with Applications, 2017, 67: 296-303.

[12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[13] YAO Y, WEI Y, GAO F, 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.

[14] HAO Y, SHENG Y, WANG J. Variant Gated Recurrent Units With Encoders to Preprocess Packets for Payload-Aware Intrusion Detection[J]. IEEE Access, 2019, 7: 49985-49998.

[15] YANG J, NGUYEN M N, SAN P P, et al. Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition[C]// Proceedings of the 24th International Joint Conference on Artificial Intelligence. Menlo Park, CA: AAAI, 2015: 3995-4001.

[16] WANG W. HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection [J]. IEEE Access, 2018, 6(99): 1792-1806.

[17] XU C, SHEN J, DU X, et al. An intrusion detection system using a deep neural network with gated recurrent units[J]. IEEE Access, 2018, 6: 48697-48707.

[18] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]// Proceedings of the 2016 International Conference on Learning Representations. Berlin: ResearchGate, 2016: 1-13.

[19] PEARLMUTTER B A. Gradient calculations for dynamic recurrent neural networks: A survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 1995, 6(5): 1212-1228.

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

[21] DHANABAL L,SHANTHARA S P. A study on NSL-KDD data set for intrusion detection system based on classification algorithms[J]. International Journal of Advanced Research in Computer and Communication Engineering, 2015, 4(6): 446-452.

[22] HASAN M A M, NASSER M, Pal B, et al. Support vector machine and random forest modeling for intrusion detection system (IDS)[J]. Journal of Intelligent Learning Systems & Applications, 2014, 6(1): 45-52.

[23] SHONE N, NGOC T N, PHAI V D, et al. A deep learning approach to network intrusion detection[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2018, 2(1): 41-50.

[24] AKYOL A, HACIBEYOGLU M, KARLIK B. Design of multilevel hybrid classifier with variant feature sets for intrusion detection system[J]. IEICE Transaction on Information and Systems, 2016, E99.D(7): 1810-1821.

[25] MA T, WANG F, CHENG J, et al. A hybrid spectral clustering and deep neural network ensemble algorithm for intrusion detection in sensor networks[J]. Sensors, 2016, 16(10): 1701-1724.

[26] REZA M, MIRI S, JAVIDAN R. A hybrid data mining approach for intrusion detection on imbalanced NSL-KDD data set[J]. International Journal of Advanced Computer Science and Applications, 2016, 7(6): 20-25.

[27] DIRO A A, CHILAMKURTI N. Distributed attack detection scheme using deep learning approach for Internet of Tings[J]. Future Generation Computer Systems, 2017, 82(5): 761-768.

[28] JANARTHANAN T, ZARGARI S. Feature selection in UNSW-NB15 and KDDCUP'99 datasets[C]// Proceedings of the 26th International Symposium on Industrial Electronics (ISIE). Piscataway: IEEE, 2017: 1881-1886.

[29] JING D, CHEN H B. SVM based network intrusion detection for the UNSW-NB15 dataset[C]// Proceedings of the 13th International Conference on ASIC (ASICON). Piscataway: IEEE, 2019: 7281-7284.

[30] ZHANG J, LING Y, CHUNG Y, et al. Model of the Intrusion Detection System Based on the Integration of Spatial-Temporal Features[J]. Computers & Security, 2019, 89(3): 167-176.

This work is partially supported by the National Natural Science Foundation of China (61773286).

ZHANG Quanlong, born in 1994, M. S. candidate. His research interests include network information security.

WANG Huaibin, born in1960, professor. His research interests include network information security, computer software.