

文章编号: 2096-1618(2023)06-0668-05

基于卷积神经网络的异常流量鉴别方法

詹鸿辉, 程仲汉

(福建警察学院计算机与信息安全管理系, 福建 福州 350007)

摘要:入侵检测系统是网络安全的重要组成部分。针对已知网络攻击的检测,深度学习和传统机器学习都存在查准率和准确率低,以及对重要特征难以有效提取的问题,提出一种基于卷积神经网络结构的异常流量鉴别方法 CNN-BDF。对入侵数据建立神经网络,在卷积网络后引入批归一化层,并使用 Flatten 函数作用于卷积层到全连接层的过渡,最后在全连接层中间引入 Dropout 层。采用 NSL-KDD 数据集进行模型评估,实验结果表明, CNN-BDF 模型的准确率和查准率分别达到89.01%和84.72%,较基于传统机器学习与深度学习的入侵检测模型具有更好的效果。

关键词:入侵检测;网络安全;机器学习;深度学习;卷积神经网络

中图分类号:TP393.08

文献标志码:A

doi:10.16836/j.cnki.jcui.2023.06.008

随着云计算、大数据和第五代移动通信技术的发展,互联网化需求进一步扩大,网络安全防护内容也随之增加。入侵检测系统(intrusion detection system, IDS),能够在入侵到达计算机系统之前主动进行防御,加强了网络空间安全性,成为计算机安全检测和防御领域的一项重要技术。但是,网络威胁的多样性以及网络安全事件频发,基于传统机器学习的方法已经不适应新的网络安全防护场景。近年来,深度学习在网络安全领域广泛应用,但仍面临数据不平衡和实时检测等问题。如何提高鉴别异常流量的准确率,同时保障用户的安全访问,在网络安全防范领域具有重要研究价值。

1 相关研究

目前,入侵检测系统根据入侵检测的行为分为两种模式:异常检测和误用检测。基于异常的入侵检测领域研究主要有传统机器学习和深度学习。在传统机器学习研究上,陈晨等^[1]利用 PSOGWO 算法提出了一种融合粒子群搜索的灰狼优化算法。付子熾等^[2]结合 SVM 和 K 最近邻近(K-nearest neighbor, KNN)算法,并采用平衡 k 维树作为数据结构提出了 IL-SVM-KNN 分类器。Logeswari 等^[3]提出了一种新颖的 HFS-LGBM IDS,首先应用随机森林递归特征消除(RF-RFE)方法得到最优特征集,然后使用 LightGBM 算法进行检测任务。Duo 等^[4]采用粒子群优化-支持向量机(PSO-SVM)和遗传算法,构建了基于支持向量机的异常检测模型。

以上方法需要人工提取特征,处理高维数据的特征需要消耗大量计算资源,不仅需要大量时间,还会遗漏部分有效特征,导致准确率低。

还有一类是基于深度学习的入侵检测方法。Yin 等^[5]基于深度学习对入侵检测系统进行建模,提出不同的神经元数量和学习速率对模型性能具有影响。董卫宇^[6]采用堆叠含有多个通过残差模块的 Attention(注意力)模块,提出一种基于堆叠卷积注意力(STAC-ON-ATTN)的 DNN 网络流量异常检测模型。曹卫东等^[7]用变分自编码(variational auto-encoder, VAE)处理数据,提出基于深度生成模型的半监督入侵检测模型。连鸿飞等^[8]结合 CNN、双向 LSTM 和注意力机制,提出一种过采样算法与混合神经网络相结合的入侵检测模型。上述方法取得了不错的效果,但是在对已知网络攻击的检测上仍待提高。

本文提出一种基于数据清洗的数据转换。首先,在数据预处理上使用特征值归一化方法,再将一维向量数据转换成二维的图像数据。其次,针对 NSL-KDD 数据集^[9]的不平衡问题,在经典卷积神经网络^[10]基础上将批归一化层应用于卷积层-池化层之后,卷积层过渡到全连接层使用 Flatten 函数。最后,在全连接层中间引入 Dropout 层。此外,运用 Xavier 方法^[11]初始化模型权重和 Adam^[12]优化算法等常用的深度学习技术。由此,提出一种基于卷积神经网络的改进异常流量鉴别方法 CNN-BDF(CNN-BatchNorm_Dropout_Flatten)。在卷积神经网络的基础上,加入了批归一化层、Dropout 层、Flatten 函数三个层面的改进。实验结果表明,本文所提出的入侵检测模型在各项评估指标上具有不错的提升。

2 卷积神经网络

2.1 深度学习

深度学习是机器学习新的研究方向,是一种网络层更深的神经网络,能够学习样本数据的内在规律和表示层次^[13]。卷积神经网络(convolutional neural networks, CNN)是一类包含卷积计算且具有深度结构的前馈神经网络(feedforward neural networks),是深度学习的代表算法之一。

2.2 卷积神经网络基本原理

卷积神经网络是一种前馈神经网络,对于图像识别有出色表现。本文先将一维的入侵检测数据转换为二维数据,然后再进行训练。卷积神经网络由具有可学习的权重和偏置常量的神经元组成。每个神经元都接收一部分输入,并进行卷积计算^[14]。卷积神经网络的基本结构由输入层、卷积层、池化层、全连接层和输出层组成。卷积神经网络通常包含以下几层:

(1)卷积层(convolutional layer),卷积神经网络中的卷积层由若干卷积单元构成,各个卷积单元的参数通过反向传播算法优化得来。卷积运算能提取输入数据的不同特征,首层卷积层可能只提取部分低级的特征。例如线条、边缘和角等层级,更深卷积层则能从低级特征中迭代提取更复杂的特征。

(2)激活层(activation),是神经网络中神经元上运行的函数,负责将神经元的输入映射到输出端。其中,线性整流层(rectified linear units layer, ReLU layer)^[15]是神经网络常用的激活函数。公式如下:

$$f(x) = \max(0, x)$$

(3)池化层(pooling layer),在卷积层处理后一般会产生维度较大的特征,该层将特征切分成几个区域,取其最大值或平均值,产生新的、维度更小的特征。其作用是降低数据的空间尺寸,减少网络中参数的数量,计算资源耗费,也能有效控制过拟合。

(4)全连接层(fully-connected layer),将所有局部特征结合转换成全局特征,用于计算每一类的得分。根据计算神经网络的推测结果与真实标签的差距,构造损失函数。将损失函数对各种权重、卷积核参数求导,慢慢优化参数找到损失函数的最小值。这一过程称为梯度下降。经过训练的模型即可用于分类任务。

2.3 网络结构

对异常流量的鉴别实际上是根据数据特征对进行分类的问题。本文采用卷积神经网络对数据进行训练后得出异常流量分类 CNN-BDF 网络。针对实验采用的 NSL-KDD 数据集的不平衡问题,在经典卷积神经网络

基础结构上将批归一化层应用于卷积神经网络。CNN-BDF 还将 Flatten 层应用于卷积层到全连接层的过渡,在全连接层中间引入 Dropout 层。并调整卷积层的关键参数以提高模型准确性。

CNN-BDF 神经网络共有 14 层,结构和参数如图 1 所示。

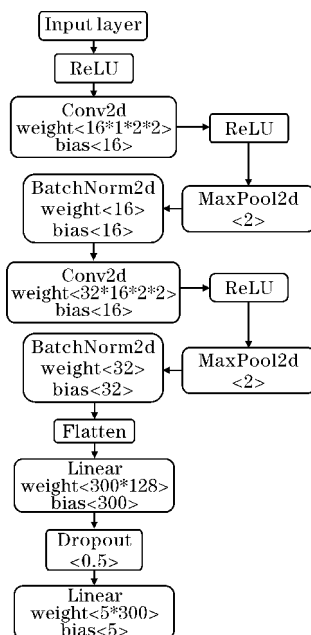


图1 CNN-BDF 模型结构

CNN-BDF 的分层结构描述如下:

(1)第1层为输入层。入侵记录数据是一维数据,经过特征编码和数据特征值归一化处理后,首先去除一列空值数据,将单条数据由一维 1×41 处理为一维 1×40 ,最后将剩余 40 列一维数据转换成二维 $1 \times 5 \times 8$ 大小的图像数据。

(2)第3和第7层是卷积层。卷积层进行卷积运算,对于输入数据,以一定间隔滑动卷积核的窗口并应用。在卷积运算之前需要进行填充处理,以调整输出的大小,填充值设置为 1。卷积核的位置间隔,也称为步幅。在保证网络精度的情况下,减少参数,将卷积核大小都设置为 2×2 ,步幅设置为 1。在一定程度上提高了对数据特征的获取。

(3)第2、4、8层为激活函数层,使用 ReLU 作为激活函数。第5和第9层为最大池化层,其作用是对微小的位置变化具有鲁棒性^[16],并且能减少计算量。

(4)第6和第10层为批归一化层(batch norm, BN)。为了使各层拥有适当的广度,BN 层可以将激活值的分布调整成高斯分布。

(5)第12~14层包含全连接层和 Dropout 层。Dropout 可以简单地实现,在某种程度上能够抑制过拟合,在每一个 batch 的训练中随机减掉一些神经元。这里将 Dropout 值设置为 0.5。

(6)第 11 层为 Flatten 全连接层。Flatten 层将输入“压平”,即把多维的输入一维化,多应用于卷积层到全连接层的过渡。

2.4 超参数设置

在 CNN-BDF 的基础上,采用以下 3 种优化方法:
(1)Xavier 初始化方法。在深度学习中,神经网络的权重初始化方法对模型的收敛速度和性能具有重要的影响。随着网络深度的增加,训练中容易出现梯度消失或梯度爆炸等问题。因此,对权重 W 的初始化至关重要,本文采用正态分布 $N(\text{mean}=0, \text{std}=0.01)$ 的值填充输入张量,将网络中参数 weight 初始化,初始化参数值符合正态分布。参数初始化的目的是为了让神经网络在训练过程中抑制过拟合、提高泛化能力,有利于提升模型的收敛速度和性能表现。

(2)本文模型训练所采用的损失函数为交叉熵损失函数,这是一个平滑函数,其本质是信息理论中的交叉熵在分类问题中的应用。NSL-KDD 数据集的各类标签分布不平衡,交叉熵相比其他方法计算得到梯度更加稳定。

(3)梯度下降是一种通用的优化算法,能为大范围的问题找到最优解。梯度下降的中心思想就是迭代地调整参数从而使成本函数最小化。Adam 优化算法结合 AdaGrad 和 RMSProp 两种优化算法的优点。对梯度的一阶矩估计(即梯度的均值)和二阶矩估计(即梯度的未中心化的方差)进行综合考虑。在深度学习中易于实现,能降低模型训练对计算机资源的需求。这里对 Adam 算法的网络参数学习率、权重分别预设 为 0.001 和 0.0001 开始学习以提升模型性能。

3 实验

3.1 实验方法与环境

为验证模型的有效性,进行实验对比,将原生的卷积神经网络所训练的模型记为 CNN。为验证本文模型具有更好的效果,同时对比 CNN、SVM、RandomForest、lightGBM 方法,选取多分类任务中常用的 3 种指标,分别为准确率、查准率、查全率,以此评估模型。

(1)准确率 (Acc) 是模型正确检测的样本数占总样本数的比值。

(2)查准率 (precision) 是被正确检测的样本数占被检测到样本总数的比值。

(3)查全率 (recall) 是被正确检测的样本数占该类样本总数的比值。

实验环境参数如表 1 所示。

表 1 实验环境参数

实验环境	参数
操作系统	Window 10
CPU	Intel(R) i5-10210U
GPU	NVIDIA GeForce MX250
RAM	16.0 GB
TensorFlow	2.3.0
scikit-learn	1.1.1
torch	1.10.0+cu102

3.2 数据集与数据预处理

使用 2009 年公开的 NSL-KDD 数据集,它是对 KDD CUP99 数据集的改进,解决了 KDD99 的一些固有问题^[17]。NSL-KDD 训练集中没有冗余记录,不会导致分类器频繁的记录。NSL-KDD 测试集没有重复记录,使检测评估更具有准确性。NSL-KDD 中共有数据 148517 条,每条数据有 41 位特征值。其中,训练集有 125973 条数据,测试集有 22544 条数据。数据中的入侵检测攻击类别如表 2 所示。

表 2 NSL-KDD 训练集的攻击类型

攻击类别	子类	数据量
Dos	back (956), land (18), neptune (41214), pod (201), smurf (2646), teardrop (892)	45927
Probe	ipsweep (3599), nmap (1493), portsweep (2931), satan (3633)	11656
U2R	buffer_overflow (30), perl (3), loadmodule (9), rootkit (10)	52
R2L	ftp_write (8), guess_passwd (53), imap (658), multihop (7), phf (4), spy (2), warezclient (890), warezmaster (20)	995

(1)NSL-KDD 内的训练集和测试集中包含的攻击方法不同,在测试集中含有 17 种未在训练集出现的标签类型,共计 3751 条,删除这类样本更易于评价模型对已知网络攻击的检测效果。

(2)本文使用的数据集的 41 列特征内含有字符数据和数值数据,在机器学习中一般使用数值数据。数据含有 protocol_type、service、flag 和 label 4 列字符数据。因此,使用 LabelEncoder (标签编码) 中的 fit_transform 函数进行特征编码将上述 4 列特征转化为数值型特征。为加快本文模型收敛速度,使用 MinMaxScaler (特征值归一化) 方法对所有数据预处理。

3.3 实验方法

由于原始的入侵数据是一维的向量数据,而卷积神经网络一般用于处理二维的图像数据。因此,本文采用数据清洗的方法对一维 41 列数据进行检查后发现数据集中第 20 列全为空值,予以删除。而后将剩余

的 40 列一维数据转换成二维 5×8 大小的图像数据,该方法简单且易于实现。

对于对比模型 SVM,将 C 设置为 100 且选择高斯核函数作为模型的超参数。将 lightGBM 模型的最大深度设置为 3,学习率设置为0.1。

3.4 实验结果分析

将本文的 CNN-BDF 算法与 CNN、SVM、RandomForest、lightGBM 算法进行实验对比,以验证本文方法的有效性。对比结果如表 3 ~5 所示。

表 3 总体指标对比			
模型	Acc	precision	recall
CNN-BDF	0.8901	0.8472	0.9721
CNN	0.8311	0.7775	0.9685
SVM	0.8372	0.7692	0.9832
RandomForest	0.8655	0.8094	0.9749
lightGBM	0.8624	0.8081	0.9733

表 4 查准率对比					
模型	Dos	Probe	U2R	normal	R2L
CNN-BDF	0.9857	0.8096	0.6250	0.8468	0.9556
CNN	0.9809	0.7475	1.0000	0.7735	0.9230
SVM	0.9928	0.8651	0.0454	0.7692	0.9280
RandomForest	0.9873	0.8290	0.6666	0.8094	1.0000
lightGBM	0.9907	0.7943	0.5555	0.8080	0.9952

由表 5 可知,CNN-BDF 模型在 Dos、Probe、normal 类型上的查全率总体上优于 CNN、SVM、lightGBM 模型,虽与 RandomForest 模型相比有微小差距,但在 U2R、R2L 类型的查全率上本文模型远高于 Random-

表 3 是 CNN-BDF 和 CNN、SVM、RandomForest、lightGBM 在总体准确率、查准率、查全率上的对比结果。CNN-BDF 的准确率达到 89.01%,准确率高于 CNN,也高于传统的机器学习算法。CNN-BDF 的查准率达到 84.72%,与 CNN 相比有效提高了查准率、查全率。在高于传统机器学习算法查准率的同时保证了较好的查全率。由表 3 可知,本文提出的 CNN-BDF 模型在数据集的分类效果上高于其他模型。

由表 4 可知,CNN-BDF 模型在 normal 类型的查准率上高于其他 4 种模型。CNN-BDF 模型在 Dos 类型的查准率上略低于其他模型,但在 Probe、R2L 两种类型上分别高于 CNN 模型和 CNN、SVM 模型。SVM 在 U2R 类型的查准率仅有 4.54%,而文中模型达到了 62.5%。总体来说,CNN-BDF 模型在查准率上优于其他 4 种模型。

Forest 模型。本文提出的模型对 U2R、R2L 类型的检测有较好的效果。

综合表 3 ~5 实验结果,本文提出的模型在提高查准率的同时也保证了好的查全率。

表 5 查全率对比					
模型	Dos	Probe	U2R	normal	R2L
CNN-BDF	0.9886	0.9421	0.2702	0.9701	0.2646
CNN	0.9212	0.7631	0.0541	0.9737	0.0109
SVM	0.9205	0.6762	0.0271	0.9862	0.0586
RandomForest	0.9904	0.9954	0	0.9728	0.0150
lightGBM	0.9484	0.9918	0.1315	0.9733	0.0954

4 结束语

针对目前入侵检测算法对已知的异常网络流量的检测率低和准确率不高的问题,提出了特征值归一化的预处理方法和基于数据清洗的数据转换方法,将向量数据转换为图像数据。CNN-BDF 算法采用经典卷积神经网络基础结构上加入批归一化层和 Flatten 函数,并在全连接层间引入 Dropout 层。实验结果表明,相比 CNN、SVM、RandomForest、lightGBM,CNN-BDF 模型具有较高准确率和查准率,有效提升了已知的异常

网络流量的检测效果。不过,在未知攻击类型的检测效果上还有待改进,今后将继续研究网络结构,分析特征间的关联性,以改进模型对未知攻击类型的检测效果。此外,将增加时间维度的衡量,提高检测的实时性。

参考文献:

[1] 陈晨,刘曙,王艺菲,等. 基于 PSOGWO-SVM 的网络入侵检测方法[J]. 空军工程大学学报(自然科学版),2022,23(2):97-105.

[2] 付子熾,徐洋,吴招娣,等. 基于增量学习的

- SVM-KNN 网络入侵检测方法[J]. 计算机工程, 2020, 46(4): 115–122.
- [3] Logeswari G, Bose S, Anitha T. An Intrusion Detection System for SDN Using Machine Learning[J]. Intelligent Automation & Soft Computing, 2023, 35(1): 868–880.
- [4] Duo Ruifeng, Nie Xiaobo, Yang Ning, et al. Anomaly Detection and Attack Classification for Train Real-Time Ethernet[J]. IEEE ACCESS, 2021, 9: 22528–22541.
- [5] Yin Chuanlong, Zhu Yuefei, Fei Jinlong, et al. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks[J]. IEEE Access, 2017, 5: 21954–21961.
- [6] 董卫宇, 李海涛, 王瑞敏, 等. 基于堆叠卷积注意力的网络流量异常检测模型[J]. 计算机工程, 2022, 48(9): 12–19.
- [7] 曹卫东, 许志香, 王静. 基于深度生成模型的半监督入侵检测算法[J]. 计算机科学, 2019, 46(3): 197–201.
- [8] 连鸿飞, 张浩, 郭文忠. 一种数据增强与混合神经网络的异常流量检测[J]. 小型微型计算机系统, 2020, 41(4): 786–793.
- [9] Tavallaee M, Bagheri E, Lu W, et al. A detailed analysis of the KDD CUP 99 data set[C]. IEEE symposium on computational intelligence for security and defense applications. Ieee, 2009: 1–6.
- [10] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84–89.
- [11] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]. Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010: 249–256.
- [12] Biggio B, Corona I, Fumera G, et al. Bagging classifiers for fighting poisoning attacks in adversarial classification tasks[C]. International workshop on multiple classifier systems. Springer, Berlin, Heidelberg, 2011: 350–359.
- [13] 张昊, 张小雨, 张振友, 等. 基于深度学习的入侵检测模型综述[J]. 计算机工程与应用, 2022, 58(6): 17–28.
- [14] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016: 90–106.
- [15] Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional networks[C]. Proceedings of the Conference Neural Information Processing Systems (NIPS). 1097: 4875–4884.
- [16] 纪守领, 杜天宇, 邓水光, 等. 深度学习模型鲁棒性研究综述[J]. 计算机学报, 2022, 45(1): 190–206.
- [17] McHugh J. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory[J]. ACM Transactions on Information and System Security (TISSEC), 2000, 3(4): 262–294.

Identification Method of Abnormal Traffic based on Convolution Neural Network

ZHAN Honghui, CHENG Zhonghan

(Department of Computer and Information Security Management, Fujian Police College, Fuzhou 350007, China)

Abstract: Intrusion detection system is an important part of network security. For the detection of known network attacks, both deep learning and traditional machine learning have low precision and accuracy, and it is difficult to effectively extract important features. Aiming at these problems, an abnormal traffic identification method CNN-BDF based on convolutional neural network structure is proposed in this paper. The neural network is established for the intrusion data, the batch normalization layer is introduced after the convolutional network, and the Flatten function is used to act on the transition from the convolutional layer to the fully connected layer. Finally, the Dropout layer is introduced in the middle of the fully connected layer. The NSL-KDD data set is used to evaluate the model. The experimental results show that the accuracy and precision of the CNN-BDF model reach 89.01% and 84.72% respectively, which shows better performance than the intrusion detection model based on traditional machine learning and deep learning.

Keywords: intrusion detection; network security; machine learning; deep learning; convolutional neural network