

基于超图神经网络的恶意流量分类模型

赵文博^{1,2,3}, 马紫彤^{1,2,3}, 杨哲^{1,2,3}

(1. 苏州大学计算机科学与技术学院, 江苏 苏州 215006; 2. 江苏省计算机信息处理技术重点实验室, 江苏 苏州 215006;
3. 江苏省大数据智能工程实验室, 江苏 苏州 215006)

摘要: 随着网络的普及和依赖程度的不断增加, 恶意流量的泛滥已经成为网络安全领域的严重挑战。在这个数字时代, 网络攻击者不断寻找新的方式来侵入系统、窃取数据和破坏网络服务。开发更有效的入侵检测系统, 及时发现并应对恶意流量, 可以应对网络攻击的持续威胁, 极大地减少网络攻击带来的损失。然而现有的恶意流量分类方法存在一些限制, 其中之一是过度依赖对数据特征的选择。为了提高恶意流量分类的效果, 提出了一种创新的方法, 即基于超图神经网络的恶意流量分类模型。这一模型的核心思想是将流量数据表示为超图结构, 并利用超图神经网络(HGNN, hypergraph neural network)来捕获流量的空间特征。HGNN能够更全面地考虑流量数据之间的关系, 从而更准确地表征恶意流量的特征。此外, 为了处理流量数据的时间特征, 引入了循环神经网络(RNN, recurrent neural network), 进一步提高了分类模型的性能。最终, 提取的时空特征被用于进行恶意流量分类, 从而帮助检测网络中的潜在威胁。通过一系列消融实验, 验证了HGNN+RNN模型的有效性, 证明其能够高效提取流量的时空特征, 从而改善了恶意流量的分类性能。在3个广泛使用的开源数据集, 即NSL-KDD、UNSW-NB15和CIC-IDS-2017上, 模型取得了卓越的分类准确率, 分别达到了94%、95.6%和99.08%。这些结果表明, 基于超图神经网络的恶意流量分类模型在提高网络安全水平方面具有潜在的重要意义, 有望帮助网络安全领域更好地应对不断演变的网络威胁。

关键词: 恶意流量; 网络攻击; 超图神经网络; 循环神经网络

中图分类号: TP399

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2023069

Model of the malicious traffic classification based on hypergraph neural network

ZHAO Wenbo^{1,2,3}, MA Zitong^{1,2,3}, YANG Zhe^{1,2,3}

收稿日期: 2022-09-23; 修回日期: 2023-08-18

通信作者: 杨哲, yangzhe@suda.edu.cn

基金项目: 国家自然科学基金(62072321); 教育部产学协同育人项目(220606363154256); 江苏省高校自然科学基金(20KJB520002); 江苏省未来网络科研基金(FNSRFP-2021-YB-38); 江苏高校优势学科建设工程资助项目

Foundation Items: The National Natural Science Foundation of China (62072321), The Project of the Ministry of Education on the Cooperation of Production and Education (220606363154256), The Natural Science Foundation of the Jiangsu Higher Education Institutions of China (20KJB520002), The Future Network Research Foundation of Jiangsu Province (FNSRFP-2021-YB-38), Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions

引用格式: 赵文博, 马紫彤, 杨哲. 基于超图神经网络的恶意流量分类模型[J]. 网络与信息安全学报, 2023, 9(5): 166-177.

Citation Format: ZHAO W B, MA Z T, YANG Z. Model of the malicious traffic classification based on hypergraph neural network[J]. Chinese Journal of Network and Information Security, 2023, 9(5): 166-177.

1. School of Computer Science and Technology, Soochow University, Suzhou 215006, China
2. Provincial Key Laboratory for Computer Information Processing Technology, Suzhou 215006, China
3. Provincial Key Laboratory for Intelligent Engineering in Big Data, Suzhou 215006, China

Abstract: As the use and reliance on networks continue to grow, the prevalence of malicious network traffic poses a significant challenge in the field of network security. Cyber attackers constantly seek new ways to infiltrate systems, steal data, and disrupt network services. To address this ongoing threat, it is crucial to develop more effective intrusion detection systems that can promptly detect and counteract malicious network traffic, thereby minimizing the resulting losses. However, current methods for classifying malicious traffic have limitations, particularly in terms of excessive reliance on data feature selection. To improve the accuracy of malicious traffic classification, a novel malicious traffic classification model based on Hypergraph Neural Networks (HGNN) was proposed. The traffic data was represented as hypergraph structures and HGNN was utilized to capture the spatial features of the traffic. By considering the interrelations among traffic data, HGNN provided a more accurate representation of the characteristics of malicious traffic. Additionally, to handle the temporal features of traffic data, Recurrent Neural Networks (RNN) was introduced to further enhance the model's classification performance. The extracted spatiotemporal features were then used for the classification of malicious traffic, aiding in the detection of potential threats within the network. Through a series of ablative experiments, the effectiveness of the HGNN+RNN method was verified. These experiments demonstrate the model's ability to efficiently extract spatiotemporal features from traffic, resulting in improved classification performance for malicious traffic. The model achieved outstanding classification accuracy across three widely-used open-source datasets: NSL-KDD (94% accuracy), UNSW-NB15 (95.6% accuracy), and CIC-IDS-2017 (99.08% accuracy). These results underscore the potential significance of the malicious traffic classification model based on hypergraph neural networks in enhancing network security and its capacity to better address the evolving landscape of network threats within the domain of network security.

Keywords: malicious traffic, cyberattack, hypergraph neural network, recurrent neural network

0 引言

互联网的飞速发展为人们的生活提供了极大的便利,每天各种联网设备会产生巨大的流量。然而在真实网络环境中,除了正常的网络流量外,恶意流量也无所不在^[1]。

恶意流量是指通过网络创建或接收的任何可疑链接、文件或连接而产生的流量。这些流量通过应用层攻击 API、网站等,以达到获利的目的,也会损害国家、企业及个人的合法权益,严重影响互联网的安全。恶意流量的类型是多样的,即使是同一种类型的恶意流量之间也有很大区别,如网络攻击的流量可以细分为端口扫描、登录破解、漏洞利用、DDoS 攻击等类型^[2]。

恶意流量的分类是入侵检测系统的一个重要功能,自机器学习模型被应用于入侵检测以来,研究人员提出了多种基于传统机器学习的恶意流

量分类模型。此类模型将支持向量机、随机森林、遗传算法等多种机器学习方法与不同的特征选择方法相结合,在相关数据集上取得了比较好的效果^[3-5]。但是随着入侵攻击变得越来越多样化,网络流量的特征随之变得复杂,上述传统机器学习模型在不同数据集上的特征提取方面展现了相当大的局限性。

深度学习的强大特征表达能力,可以从训练数据的有限特征中发现新特征,又可以很好地处理复杂的数据集,目前已被广泛应用于图像处理、音频处理、自然语言处理等领域^[6]。而网络流量具有网络拓扑信息,且恶意流量的分布通常与时间相关,网络流量之间存在互相影响的空间特征和时间特征,因此可以使用深度学习直接学习流量数据的时空特征,从而使得流量分类不再依赖特定的特征选择方法。

为了充分学习网络流量的时空特征,本文提出了一种新的恶意流量分类模型,主要工作包含

以下3个方面。

1) 将超图应用于流量分类领域,提出了一种基于超图神经网络的恶意流量分类模型。

2) 为网络流量数据构建超图数据结构,分别使用超图神经网络(HGNN, hypergraph neural network)与循环神经网络(RNN, recurrent neural network)在超图的空间和时间序列角度提取顶点特征,充分使用原有数据的空间和时间序列信息,大大提高了网络流量特征的提取能力。

3) 所提模型在3个代表性的恶意流量数据集(NSL-KDD、UNSW-NB15和CIC-IDS-2017)上取得了较好的检测效果。实验结果表明,所提模型优于传统机器学习模型及相关深度学习模型。

1 相关工作

恶意流量分类是入侵检测系统的重要功能,对网络安全具有重要的现实意义^[7]。研究人员在这方面进行了相当广泛的探索,提出了许多基于传统机器学习或者深度学习的模型。

流量数据的特征较多,且包含部分无关特征与冗余特征,因此对于恶意流量分类这一任务,选择相对较好的特征非常重要^[8]。许多恶意流量分类模型使用特征选择技术对流量特征进行筛选处理后,再进行机器学习模型的训练。

Vijayanand等^[9]提出了一种基于遗传算法的特征选择和支持向量机分类器的入侵检测系统,并在CIC-IDS2017数据集上取得了较高的准确率。Lu等^[10]首先使用信息增益获取对分类结果影响较大的特征,再结合卷积神经网络(CNN, convolutional neural network)与长短期记忆(LSTM, long short-term memory)网络实现对恶意流量的分类,并使用NSL-KDD数据集验证了其方法的有效性。Tang等^[11]使用41个特征中的6个基本网络特征,使用深度神经网络在NSL-KDD数据集上获得了75.75%的准确率。Wang等^[12]在CIC-IDS2017数据集上首先使用决策树生成的特征重要性权重以选择特征,然后通过合成少数过采样技术(SMOTE, synthetic minority over-sampling technique)优化样本类别的分布,再使用LSTM进行数据特征的提取及分类。

特征选择技术虽然能提高分类效果、降低学

习任务的难度,但是不同的数据集之间的特征相差较大,同样的特征选择方式很难在不同的数据集上都取得较好的结果^[13]。为了解决这个问题,研究者尝试使用深度学习提取流量数据的特征,以提高网络流量的分类能力。

Muna等^[14]使用深度前馈神经网络与深度自编码器构建了异常检测系统(ADS, anomaly detection system),使用深度学习模型自动分析原始网络数据以发现其中的异常数据,在UNSW-NB15数据集上取得了不错的效果。Ahmim等^[15]针对CIC-IDS-2017数据集提出了一种由3个分类器组成的混合模型,通过不同分类器提取不同的特征,再综合不同分类器的输出获得最终的预测结果。这两种模型使用深度学习直接提取流量数据的潜在特征,没有考虑流量数据本身所具有的时间与空间特性。

针对时间特征的提取,Yin等^[16]提出了一种基于RNN的入侵检测深度学习模型,在NSL-KDD数据集上的性能上优于其他传统的机器学习模型,表明RNN非常适合入侵检测。这种模型虽然有效提取了流量数据的时间特征,但是忽略了流量数据的空间特征。

为了充分提取流量数据的时间特征与空间特征,Zhang等^[17]提出了使用CNN提取流量数据的空间信息、使用LSTM提取时间特征的模型,在UNSW-NB15数据集上的实验结果表明,其模型有效提高了恶意流量分类的准确率。同样,Liu等^[18]的模型在流量的时间特征提取上也使用了LSTM,不同的是在提取流量的空间特征时使用了简化的图卷积(SGC, simplified graph convolutional)网络,在UNSW-NB15数据集上也取得了较好的结果。以上模型充分考虑了流量时间与空间上的特性,但CNN模型往往要求流量数据长度一致,故不能利用全部的数据,造成部分数据丢失。Liu等^[18]在空间特征的提取方面使用的SGC可以避免数据丢失,但由于图的结构只能表示二元关系,在对多元关系进行建模时往往会造成信息上的损失^[19]。

超图结构与图相比可以尽可能地保存多元关系的高阶信息,因此在过去的几年里,超图学习在许多领域得到了广泛的发展和应用。Zhou等^[19]最早将超图结构应用于机器学习,提出了一种超

图直推学习模型,该模型用来表示超图结构上的标签传播过程,从而可以使用超图进行样本的分类和聚类。

由于卷积神经网络在很多领域取得了巨大的成功,受 Kipf 等^[20]在图结构上提出的图卷积神经网络(GCN)网络的启发,Feng 等^[21]在超图上提出了超图神经网络,通过超边卷积的方法使得超图的顶点可以获得邻居顶点的特征信息,从而提取拓扑图的空间特征。

循环神经网络的结构特点使得其在处理序列信息时具有较大优势^[22],其隐藏层中的循环信息使得其在处理数据时充分使用了当前状态的输入以及上一状态隐藏层的输出。由于 RNN 具有强大的序列信息提取能力,目前已经被广泛使用在输入为序列信息的业务场景,因此可以使用 RNN 提取网络流量中存在的时间特征。在相关的流量分类模型中^[16],基于 RNN 的模型展现出了强大的时间特征的提取能力。

综上所述,本文提出了一种基于超图神经网络的恶意流量分类模型,使用超图神经网络提取网络流量的空间特征,使用 RNN 提取时间特征,分别使用 NSL-KDD、CIC-IDS-2017 与 UNSW-NB15 数据集来评估模型对恶意流量分类的性能,并与其他流量分类模型进行对比。

2 模型设计

2.1 模型框架

网络流量存在空间与时间两个维度的特征。

在对恶意流量进行分类时,充分利用这些信息,可以提高模型识别正常和恶意流量的性能。为了充分提取流量的时空特征,本文提出了一种基于超图神经网络和循环神经网络的恶意流量分类模型,使用 HGNN 提取流量的空间特征,使用 RNN 提取流量的时间特征。图 1 为 HGNN+RNN 模型结构示意图,包括输入层、HGNN 层、RNN 层和输出层。输入层首先对原始流量数据进行预处理,然后为网络流量构建超图结构,将得到的超图结构输入 HGNN 层,提取其空间特征,将 HGNN 层的输出输入 RNN 层提取时间特征。在 RNN 层之后,通过输出层输出预测结果。

2.2 输入层

输入层主要对原始流量数据进行数据预处理并构建超图结构。其中,数据预处理流程如图 2 所示,主要分为采样、数据清洗、标签编码、数值归一化、构建超图 5 个步骤。

数据清洗可以清除数据集中的不完整、不准确、有问题的数据和记录,提高数据集的质量。标签编码的作用是将原始数据流量中的非数字特征转化为数字特征,保证所有数据都是数值型,方便模型的学习,同时可以通过标签编码重新划分样本的所属类别,将流量多分类任务转化为区分正常与恶意流量的二分类任务。

流量特征转化为数值类型后,特征的取值大小不同,容易导致特征空间中样本点的聚类效果受个别特征值影响较大,受其他特征值影响较小。为了减小异常值的影响,需要使用归一化将特征

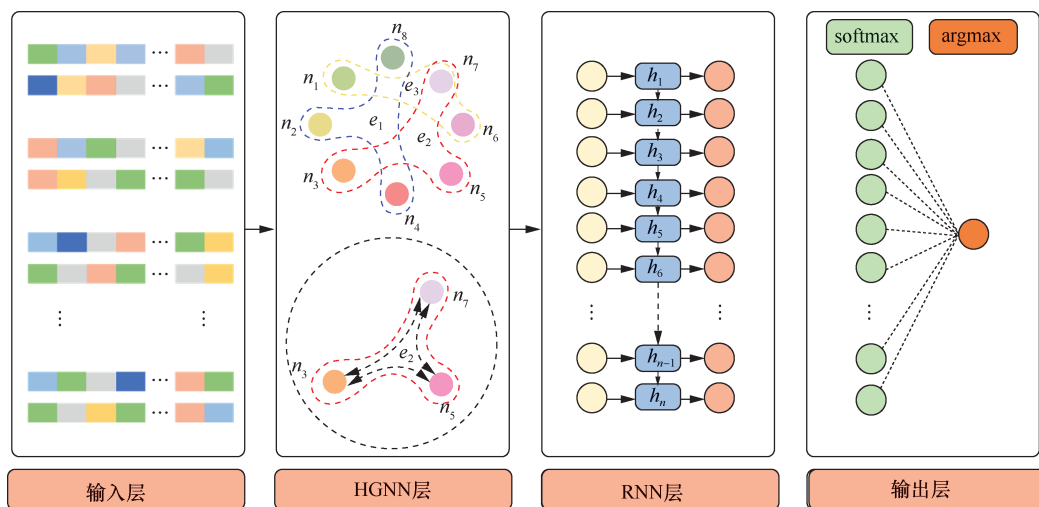


图1 HGNN+RNN 模型结构示意图
Figure 1 Model structure based on HGNN and RNN

值映射在 $(0,1]$ 。本文使用最大最小归一化方法,如式(1)所示。

$$x_i = \frac{x_i - \text{Min}}{\text{Max} - \text{Min}} \quad (1)$$

其中, x_i 表示单个数据的取值, Min 是数据所在列所有数据的最小值, Max 是数据所在列所有数据的最大值。经过归一化后的数据减小了奇异样本导致的不良影响,可用于超图结构的构建。

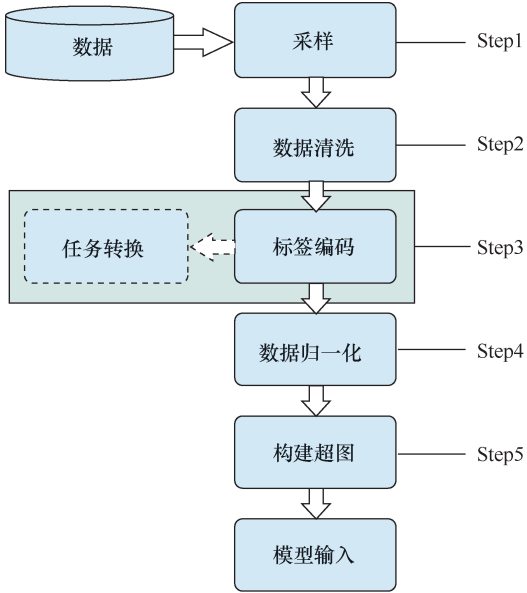


图2 数据预处理流程
Figure 2 Data preprocessing process

2.3 超图神经网络层

超图神经网络层在模型中用于提取流量数据空间特征。流量数据空间特征指的是数据背后的

网络拓扑结构,用超图结构来表示流量数据可以尽可能地保留原始流量的拓扑信息。

超图是普通图的推广,与普通图相比,其最大特点是超图中的一条边可以连接两个以上的顶点,称为超边(hyperedge)^[23],即超边是超图中所有顶点的一个子集。基于此,可以将超图定义为式(2)中的形式。

$$G = (V, E) \quad (2)$$

上式表示具有顶点集(V)和超边集(E)的超图 G 。与普通图的邻接矩阵类似,超图 G 可以用 $|V| \times |E|$ 表示其关联矩阵 H ,其中每个元素 h 可以根据式(3)定义。

$$h(v, e) = \begin{cases} 1, v \in e \\ 0, v \notin e \end{cases} \quad (3)$$

超图及其关联矩阵如图3所示。已知超图中每个顶点都属于一个特定的类,则根据超图的关联矩阵和已知类别的顶点,就可以预测其他顶点所属的类别。

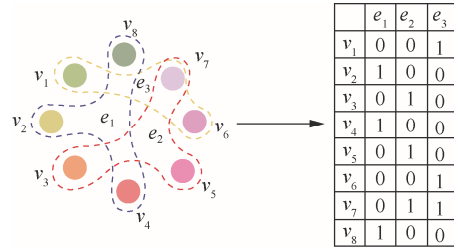


图3 超图及其关联矩阵
Figure 3 Hypergraph and incidence matrix

超边卷积层示意如图4所示, HGNN 通过定

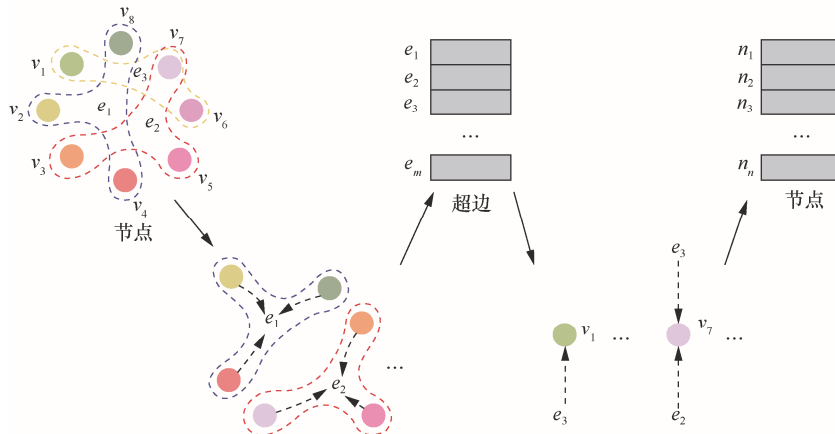


图4 超边卷积层示意
Figure 4 Schematic of the hyperedge convolution layer

义在频谱上的卷积操作，利用顶点-边-顶点的变换有效地提取超图上的高阶相关性。Feng 等^[21]将超边卷积表示为式(4)的形式。

$$Y = D_v^{-1/2} H W D_e^{-1} H^T D_v^{-1/2} X \Theta \quad (4)$$

其中， D_v 为顶点的度对角矩阵，对于顶点 $v \in V$ ，其度定义为 $d(v) = \sum_{e \in E} w(e)h(v, e)$ 。 D_e 为超边的度对角矩阵，对于一个超边 $e \in E$ ，其度的定义如下， $d(e) = \sum_{v \in V} h(v, e)$ 。 W 为超边权重的对角矩阵，即 $W = \text{diag}(w_1, \dots, w_n)$ 。卷积核 $\Theta \in R^{D_1 \times D_2}$ 是训练过程中要学习的参数，其应用于超图中的顶点通过超边卷积提取的特征。经过卷积操作后，各顶点获得相应邻居顶点的信息，得到 $Y \in R^{n \times D_2}$ ，即可用于分类。

超图神经网络利用超边卷积的方法提取超图结构中隐含的高阶信息，可以利用式(5)构建一个超边卷积层。

$$X^{(l+1)} = \sigma(D_v^{-1/2} H W D_e^{-1} H^T D_v^{-1/2} X^{(l)} \Theta^{(l)}) \quad (5)$$

其中， $X^{(l)} \in R^{N \times D}$ 表示第 l 层超图的特征信息， $X^{(0)} = X$ ， σ 表示非线性激活函数。通过式(5)可以构建包含多层卷积的超图神经网络。超图神经网络中顶点-边-顶点的变换过程如图5所示，图中 N 为超图的顶点数量， E 为超图的超边数量。首先，通过可学习的卷积核 $\Theta^{(l)}$ 处理初始顶点 $X^{(l)} \in R^{N \times D_1}$ 的特征以提取到 D_2 维的顶点特征 $X \in R^{N \times D_2}$ 。然后，根据超边聚集顶点的特征，通过 $H^T \in R^{E \times N}$ 的乘法运算形成超边的特征 $R^{E \times D_2}$ ，最后通过与矩阵 H 相乘，得到最终的输出，输出的顶点特征聚合了与其相关的超边特征。式(5)中 D_v 和 D_e 起到了归一化的作用。

由于超图神经网络具有较好的提取空间特征能力，因此使用超图神经网络可以充分地提取流量数据空间特征，从而提升分类的精度。

2.4 循环神经网络层

循环神经网络层在模型中的主要功能是对流量数据进行时序特征的提取。在流量数据中，前一个时间节点的流量特征对于当前时间的数据来说是非常重要的，如在拒绝服务（DoS）攻击的检测中，在较小的时间范围内会出现大量异常流量^[24]。

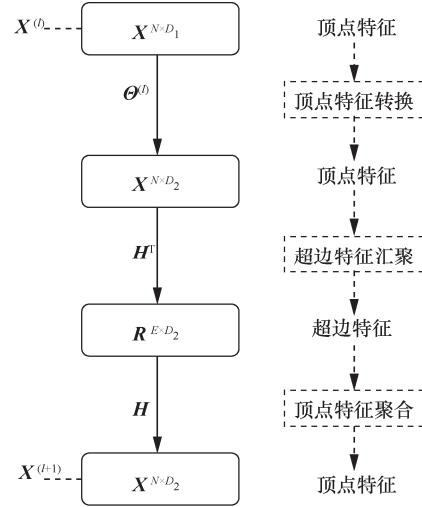


图5 超图神经网络中顶点-边-顶点的转换过程
Figure 5 Process of the node-edge-node transform in HGNN

与传统的前馈神经网络不同，RNN 可以记忆之前的信息并作用于当前的输出。循环神经网络包括输入单元、输出单元和隐藏单元，隐藏单元起到记忆信息的重要工作^[25]。RNN 模型的输出综合了从输入单元到隐藏单元的信息流以及从前一个状态的隐藏单元到当前状态隐藏单元的信息流，RNN 结构示意图如图6所示。

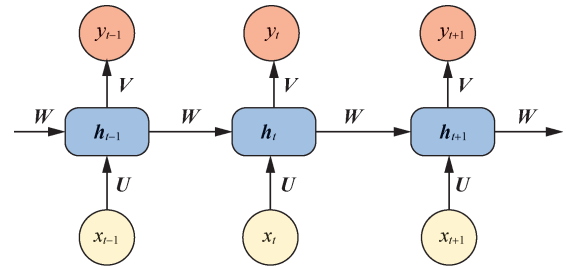


图6 RNN 结构示意图
Figure 6 Schematic of RNN structure

RNN 中隐藏单元的值可以通过式(6)来计算。式中 h_t 为当前状态隐藏单元的值， x_t 为当前状态的输入， h_{t-1} 为上一状态隐藏单元的值。 U 是输入层到隐藏层的权重矩阵， W 是上一隐藏层到前隐藏层的权重矩阵， b 为偏置。 U 、 W 、 b 是整个网络共享的参数， f 表示非线性激活函数。

$$h_t = f(Ux_t + Wh_{t-1} + b) \quad (6)$$

RNN 的输出可以通过式(7)表示，式中 y_t 为当前状态网络的输出， V 为隐藏层到输出层的权重矩阵， c 为偏置。与 U 、 W 、 b 相同， V 与 c 同样是整个网络共享的参数， g 表示非线性激活函数。

$$y_i = g(Vh_i + c) \quad (7)$$

由于隐藏单元的存在,此时网络的输出 y_i 不仅包含当前状态的输入信息,同样获得了上一状态隐藏单元所包含的所有信息。在使用 RNN 处理流量数据时,可以将不同的时间节点表示为不同状态,因此可以通过 RNN 对流量数据进行时序特征的提取。

由于恶意流量的特征通常与上一时间的流量状态相关,近些年使用 RNN 处理时序特征的方法被广泛应用于恶意流量分类^[16]。

2.5 输出层

本文所提模型中,RNN 层的输入来源于 HGNN 层的输出,由于在 HGNN 层中进行了空间特征的提取,各顶点已包含邻居顶点的信息。经过 RNN 层之后,又获得了时间序列信息,这种特征处理方法使得网络充分提取了网络流量中潜在的时间、空间信息。

RNN 层处理后的特征即可以用于恶意流量分类模型最终的输出,将此结果输入全连接层,经过 softmax 函数即可获得待分类样本属于各类别的概率,再通过 argmax 操作即可获得待测样本的预测结果。

3 实验

3.1 实验环境与数据

本文实验基于 64 位 Windows 10 操作系统,CPU 为 Intel Core i7-10700 @ 2.9 GHz,内存为 32 GB。算法实现使用 Python 3.6 作为基础语言环境,使用 NumPy 1.19.5 进行数据格式转换和矩阵运算,相关深度学习模型框架的开发基于 pytorch 1.10.0 版本。

为了验证本文所提模型对于不同数据集的泛化能力,本文选用 NSL-KDD、UNSW-NB15 和 CIC-IDS-2017 这 3 个广泛用于流量分类领域的数据集^[21]。模型在这 3 个数据集上分别进行对比实验与消融实验,以验证模型对恶意流量分类的效果,以及添加 RNN 模块对于模型特征提取能力的提升。

NSL-KDD 数据集有效解决了 KDD Cup 1999 数据集的样本冗余问题,同时调整了各样本的比例,使得数据集的类别更加均衡^[26]。由于测试集

中包含部分训练集中不存在的攻击类型,因此在 NSL-KDD 数据集上的实验结果可以很好地体现模型的实际恶意流量检测能力。

UNSW-NB15 是澳大利亚网络安全中心 (ACSC, Australian Cyber Security Centre) 的 Cyber Range 实验室于 2015 年使用 IXIA PerfectStorm 工具生成的新数据集,包含真实的网络正常活动和攻击行为^[27-31]。该数据集由 47 个属性特征和 2 个类别特征组成,包含 9 种攻击类型。UNSW-NB15 数据集的异常行为较为新颖与均衡,适合用于入侵检测研究。

CIC-IDS-2017 数据集是在真实的流量中模拟黑客攻击,并通过监控器收集网络中的数据流量^[32]。该数据集涵盖了非常广泛的流量,包括基于 Web 的攻击、常见渗透攻击、僵尸网络、网络扫描等。CIC-IDS-2017 数据集中的每条数据包含 79 个特征,更容易提高恶意流量检测模型的准确性。

3.2 评价指标

本文采用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall) 以及 $F1$ 值 ($F1$ -score) 作为评价指标,各指标的计算方法如下。

准确率描述了正确预测的样本数与总样本数的比值。

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

精确率描述了预测为正的样本数与实际为正的样本数的比值。

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

召回率描述了预测为正的样本数与所有正样本数的比值。

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

$F1$ 值是由准确率和召回率两个因素平衡构成的综合因子,是评价模型综合检测有效性的有效衡量指标。

$$F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (11)$$

上述 4 个式子中, T (true) 和 F (false) 分别代表正确和错误分类的样本, P (positive) 和 N (negative) 分别表示预测结果是正样本和负样本。

数据集中的所有样本可以分为4类：TP、TN、FP和FN。其中，TP表示预测标签为正且分类正确，TN表示预测标签为负且分类正确，FP表示预测结果为正且分类错误，FN表示预测结果为负且分类错误。

3.3 实验结果和分析

3.3.1 模型超参数选择

在模型的超参数选择上，主要包括学习率、丢弃率、隐藏层大小等。对所有的实验采用相同的超参数设置，首先采用网格搜索来进行超参数调整：学习率 $\eta \in \{0.001, 0.005, 0.01\}$ ，丢弃率 $r \in \{0, 0.25, 0.5\}$ ，隐藏层大小 $d \in \{32, 64\}$ 。使用模型在 CIC-IDS-2017 数据集上的准确率及训练时间作为参考依据，其中不同 HGNN 以及 RNN 层数下的模型性能表现如表 1 所示。

以上实验分别取 HGNN 与 RNN 的 0~3 层搭建模型，考察模型的分类能力及时间。其中，HGNN 层数取 0 时，为模型仅使用不同层数的 RNN 提取序列特征的情形；RNN 层数取 0 时，为仅使用不同层数的 HGNN 提取空间特征的情形。较为特殊的是，HGNN 与 RNN 均为 0 层时，输入数据与输出之间仅有单层全连接层。

表 1 HGNN 以及 RNN 层数下的模型性能表现
Table 1 Model performance with different HGNN and RNN layer number

HGNN 层数	RNN 层数	准确率	时间
0	0	88.90%	14s
0	1	94.85%	20s
0	2	97.45%	57s
0	3	96.60%	1m28s
1	0	96.35%	2m17s
1	1	99.50%	2m30s
1	2	99.45%	2m39s
1	3	99.45%	2m56s
2	0	98.90%	4m35s
2	1	99.25%	4m49s
2	2	99.30%	4m55s
2	3	99.35%	5m13s
3	0	98.95%	6m50s
3	1	99.05%	7m2s
3	2	99.10%	7m13s
3	3	99.35%	7m25s

通过对实验结果的分析可知，随着 HGNN 与 RNN 层数的增加，训练所需的时间也随着模型的复杂程度相应增加，但模型的分类能力并没有随着模型模块层数的堆叠而展现出更好的水平，图 7 展示了在不同参数组合下，CIC-IDS-2017 数据集上模型的准确率分布情况。

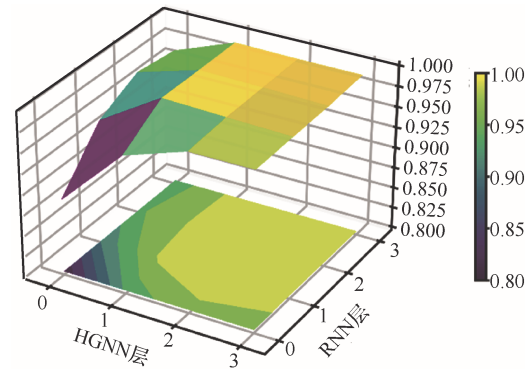


图 7 CIC-IDS-2017 数据集上模型的准确率分布
Figure 7 Model accuracy distribution on CIC-IDS-2017 dataset

由图 7 和表 1 可得，模型使用 1 层 HGNN 与 1 层 RNN 的组合时，获得了最佳的分类效果，更多层的堆叠对于提升分类效果并没有帮助，因此在本文实验中均采用单层 HGNN 与单层 RNN 的参数组合。

3.3.2 对比实验

本文将所提模型与恶意流量分类实验中常用的机器学习经典模型分类回归树（CART，classification and regression tree）、支持向量机（SVM，support vector machine）、K 近邻算法（KNN，K-nearest neighbors）以及其他相关深度学习模型对比，分别考察其在 NSL-KDD、UNSW-NB15、CIC-IDS-2017 数据集上准确率、精确率、召回率、F1 值这 4 个指标的表现。表 2 给出了 NSL-KDD 数据集上不同模型的性能比较，表中 CART、SVM 与 KNN 这 3 种模型的实验结果是在与本文模型相同环境以及实验设置情况下得出，其他模型的实验结果均引用自原论文中提供的结果。

Yin 等^[16]提出的 RNN-IDS 获得了 83.28% 的准确率，Tang 等^[11]使用特征选择以及深度神经网络的模型获得了 92.54% 的准确率。结果表明，本文所提模型在准确率、精确率上大幅优于比较模型，相较表中相应指标的次优结果分别为

94.00%、94.12%，相较表中相应指标的次优结果分别提升了 1.46%、11.54%，虽然召回率指标略低于部分比较模型，但是由于受精确率的影响，*F1* 值在整体上依然优于所有比较模型，达到了 94.06%，相较于 SVM 模型提升了 5.26%。总体而言，本文模型的综合性能优于上述对比模型。

表 2 NSL-KDD 数据集上不同模型的性能比较
Table2 Classification performance comparison of different model on NSL-KDD dataset

模型	准确率	精确率	召回率	<i>F1</i> 值
CART	87.03%	79.56%	96.04%	87.03%
SVM	89.02%	82.58%	96.04%	88.80%
KNN	83.33%	74.04%	97.36%	84.11%
RNN-IDS ^[16]	83.28%	—	—	—
DNN ^[11]	92.54%	—	90.90%	—
本文	94.00%	94.12%	94.00%	94.06%

表 3 给出了 UNSW-NB15 数据集上不同模型的性能比较。SGC-LSTM 指的是 Liu 等^[18]使用的 SGC 结合 LSTM 的模型，在 UNSW-NB15 上的准确率达到 91.67%，召回率达到 89.89%。Muna 等^[14]提出的 ADS 结合了深度前馈神经网络与深度自编码器，在此数据集上的准确率为 92.40%，召回率为 93.00%。本文模型在准确率、精确率、召回率、*F1* 值这 4 个指标上均优于比较模型，相较表中其他模型在相应指标的次优结果准确率提升了 3.2%，精确率提升 2.62%，召回率提升 2.46%，*F1* 值提升 5.32%，分别达到了 95.60%、95.68%、95.46%、95.56%。

表 3 UNSW-NB15 数据集上不同模型的性能比较
Table 3 Classification performance Comparison of different models on UNSW-NB15 dataset

模型	准确率	精确率	召回率	<i>F1</i> 值
CART	92.00%	88.10%	92.50%	90.24%
SVM	88.50%	89.04%	81.25%	84.97%
KNN	91.00%	93.06%	83.75%	88.16%
SGC-LSTM ^[18]	91.67%	—	89.89%	—
ADS ^[14]	92.40%	—	93.00%	—
本文	95.60%	95.68%	95.46%	95.56%

表 4 给出了 CIC-IDS-2017 数据集上不同模型的性能比较，Hybrid IDS 是 Ahmim 等^[15]提出的混合分类模型，在此数据集上达到了 96.67%

的准确率。SMOTE-LSTM 指的是 Wang 等结合少数过采样技术(SMOTE)与 LSTM 的分类模型^[12]，在此数据集上的准确率、精确率、召回率、*F1* 值分别为 98.90%、98.90%、98.90%、98.90%。在该数据集上的多分类的结果表明，不同模型的 4 个指标的表现相对 NSL-KDD、UNSW-NB15 都有较大提升，本文模型在准确率、精确率、召回率、*F1* 值这 4 个指标上领先于比较模型，分别为 99.08%、99.09%、99.08%、99.08%。与其他模型在相应指标的次优结果相比，准确率提升 0.18%，精确率提升 0.13%，召回率提升 0.18%，*F1* 值提升 0.18%。

表 4 CIC-IDS-2017 数据集上不同模型的性能比较
Table 4 Classification performance comparison of on CIC-IDS-2017 dataset

模型	准确率	精确率	召回率	<i>F1</i> 值
CART	98.90%	98.96%	98.90%	98.85%
SVM	97.31%	97.24%	97.31%	97.27%
KNN	98.60%	98.53%	98.60%	98.54%
Hybrid IDS ^[15]	96.67%	—	—	—
SMOTE-LSTM ^[12]	98.90%	98.90%	98.90%	98.90%
本文	99.08%	99.09%	99.08%	99.08%

从 3 个数据集的整体表现来看，本文所提模型在准确率、精确率、召回率、*F1* 值上均能取得较好的结果。与相关模型相比，本文所提模型无须针对特定数据集进行特别处理，更具有实际意义。

3.3.3 消融实验

消融实验分别从准确率、精确率、召回率与 *F1* 值 4 个方面展示了与仅使用 HGNN 提取流量的空间特征以及仅使用 RNN 层提取流量的时间特征用于流量分类相比，使用两种不同的模型同时提取空间特征、时间特征所带来的提升，证明模型从两个维度进行特征提取具有较大优势。

HGNN 模型和 RNN 模型与同时使用 HGNN+RNN 模型在 NSL-KDD、UNSW-NB15、CIC-IDS-2017 数据集上的准确率对比如图 8 所示。HGNN 模型准确率分别为 91%、92.40%、96.35%。RNN 模型准确率分别为 86.85%、94.45%、94.85%。本文所提模型在 3 个数据集上的准确率表现均有所提升，准确率分别为 94.0%、95.6%、99.08%，

相较于 HGNN 模型分别提升了 3%、3.2%、2.73%，相较于 RNN 模型分别提升了 7.15%、1.15%、4.23%。

不同数据集上的精确率对比如图 9 所示。本文所使用的 HGNN+RNN 模型在 NSL-KDD、UNSW-NB15 与 CIC-IDS-2017 数据集上的精确率分别为 94.12%、95.68%、99.09%。HGNN 模型在 3 个数据集上的精确率分别为 94.19%、92.56%、96.35%。RNN 模型在 3 个数据集上的精确率分别为 88.71%、94.46%、93.94%，与 HGNN 模型相比，HGNN+RNN 模型的精确率分别提升了 2.63%、3.12%、2.74%，与 RNN 相比，HGNN+RNN 模型的精确率分别提升了 5.41%、1.22%、5.15%。

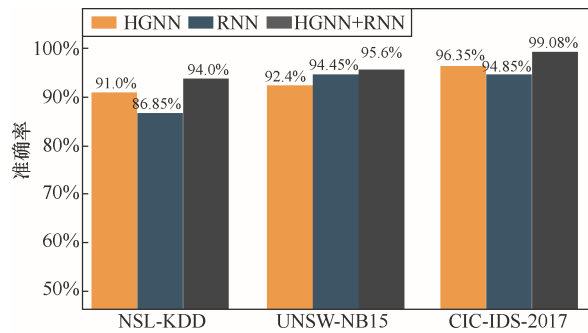


图 8 不同数据集上的准确率对比
Figure 8 Accuracy comparison of different datasets

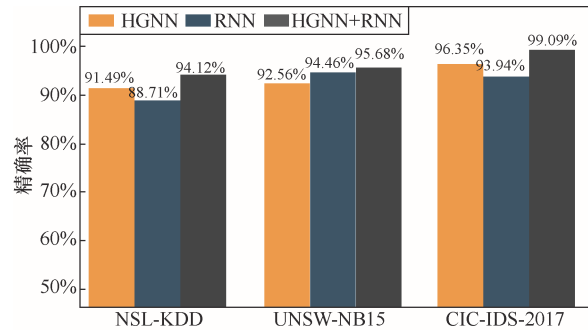


图 9 不同数据集上的精确率对比
Figure 9 Precision comparison of different data sets

在召回率方面，在 NSL-KDD、UNSW-NB15 与 CIC-IDS-2017 数据集上，本文模型相较 HGNN 模型与 RNN 模型更具优势。不同数据集上的召回率对比如图 10 所示，HGNN 模型在这 3 个数据集上的召回率分别为 90.48%、92.4%、96.35%，RNN 模型在这 3 个数据集上的召回率分别为 86.85%、94.45%、94.85%。本文模型的召回率分别为 94.0%、95.46%、99.08%，相较于 HGNN 模型

分别提升了 3.52%、3.06%、2.73%，相较于 RNN 模型分别提升了 7.15%、1.01%、4.23%。

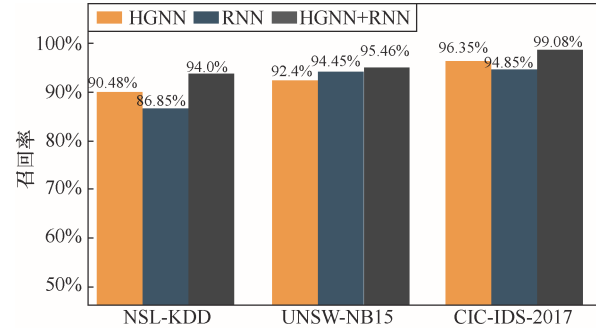


图 10 不同数据集上的召回率对比
Figure 10 Recall comparison of different data sets

$F1$ 值描述了准确率和召回率之间调和平均值的大小，由于在 NSL-KDD、UNSW-NB15 与 CIC-IDS-2017 数据集上，本文所提模型在精确率与召回率上较 HGNN 模型以及 RNN 模型均有所提升，因此 $F1$ 值高于以上两种模型。不同数据集上的 $F1$ 值对比如图 11 所示，在这 3 个数据集上，HGNN 模型的 $F1$ 值分别为 90.98%、92.48%、96.28%，RNN 模型的 $F1$ 值分别为 86.84%、94.45%、93.83%，本文 HGNN+RNN 模型的 $F1$ 值分别为 94.06%、95.56%、99.08%，相较于 HGNN 模型分别提升了 3.08%、3.08%、2.8%，相较于 RNN 模型分别提升了 7.22%、1.11%、5.25%。

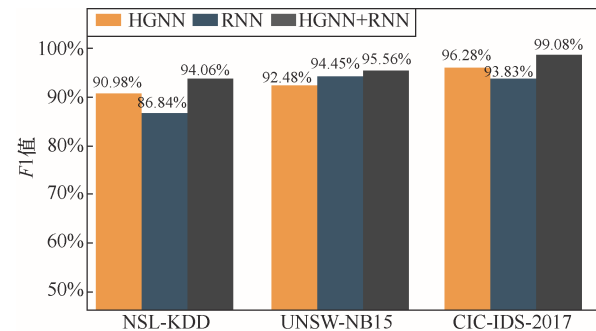


图 11 不同数据集上的 $F1$ 值对比
Figure 11 $F1$ -score comparison of different data sets

从整体的实验结果来看，RNN 模型与 HGNN 模型在不同数据集上的表现各有所长，这与各个数据集的本身情况不同有关。HGNN 模型与 RNN 模型的结合在很大程度上弥补了这一问题，同时在 3 个数据集的所有指标上的性能均有所提升，可以得出 HGNN 模型结合 RNN 模

型对于强化特征表达能力的作用是巨大的,充分证明了本文所设计的使用 HGNN 提取空间特征、使用 RNN 提取时间特征的模型的合理性以及实际意义。

4 结束语

为了有效进行恶意流量分类,本文引入了超图神经网络进行特征提取。在本文模型中,对数据从空间、时间两个角度分别使用 HGNN 与 RNN 进行提取特征,经过与其他现有模型比较,证明将超图神经网络用于恶意流量分类是一种有效的模型。

本文通过消融实验证明了在包含时序特征的流量数据集上,利用 HGNN 强大的空间特征提取能力与 RNN 在处理时序数据方面的优势,可以大大提高网络流量分类的能力,具有较好的实际意义与应用价值。下一步工作将聚焦于真实的网络环境,多方位考虑可能影响恶意流量分类效果的其他网络环境因素,进一步提高入侵检测成功率,为网络安全提供更加全面的保障。

参考文献:

- [1] ROBERTS E. Bad bot report 2020: bad bots strike back[J]. Imperva Blog, 2020.
- [2] BHUYAN M H, BHATTACHARYA D K, KALITA J K. Network anomaly detection: methods, systems and tools[J]. IEEE Communications Surveys & Tutorials, 2013, 16(1): 303-336.
- [3] WIDIPUTRA H D. Clustering based intrusion detection for network profiling using k -means ecm and k -nearest neighbor[J]. Konferensi Nasional Sistem Dan Informatika, 2009: 247-251.
- [4] ZANERO S, SAVARESI S M. Unsupervised learning techniques for an intrusion detection system[C]//Proceedings of the 2004 ACM Symposium on Applied Computing. 2004: 412-419.
- [5] ALI A, SALEH A, RAMDAN T. Multilayer perceptrons networks for an intelligent adaptive intrusion detection system[J]. International Journal of Computer Science and Network Security, 2010, 10(2): 275.
- [6] KWON D, KIM H, KIM J, et al. A survey of deep learning-based network anomaly detection[J]. Cluster Computing, 2019, 22(1): 949-961.
- [7] STALLINGS W. Network security essentials: applications and standards[M]. USA: Pearson, 2017.
- [8] ZHANG F, WANG D. An effective feature selection approach for network intrusion detection[C]//2013 IEEE Eighth International Conference on Networking, Architecture and Storage. 2013: 307-311.
- [9] VIJAYANAND R, DEVARAJ D, KANNAPIRAN B. Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection[J]. Computers & Security, 2018, 77: 304-314.
- [10] LU X, LIU P, LIN J. Network traffic anomaly detection based on information gain and deep learning[C]//Proceedings of the 2019 3rd International Conference on Information System and Data Mining. 2019: 11-15.
- [11] TANG T A, MHAMDI L, MC-LERNON D, et al. Deep learning approach for network intrusion detection in software defined networking[C]//2016 International Conference on Wireless Networks and Mobile Communications (WINCOM). 2016: 258-263.
- [12] WANG J H, SEPTIAN T W. Combining oversampling with recurrent neural networks for intrusion detection[C]//International Conference on Database Systems for Advanced Applications. 2021: 305-320.
- [13] HUBBALLI N, SURYANARAYANAN V. False alarm minimization techniques in signature-based intrusion detection systems: a survey[J]. Computer Communications, 2014, 49: 1-17.
- [14] MUNA A L H, MOUSTAFA N, SITNIKOVA E. Identification of malicious activities in industrial internet of things based on deep learning models[J]. Journal of Information Security and Applications, 2018, 41: 1-11.
- [15] AHMIM A, MAGLARAS L, FERRAG M A, et al. A novel hierarchical intrusion detection system based on decision tree and rules-based models[C]//2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS). 2019: 228-233.
- [16] YIN C, ZHU Y, FEI J, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. IEEE Access, 2017, 5: 21954-21961.
- [17] ZHANG J, LING Y, FU X, et al. Model of the intrusion detection system based on the integration of spatial-temporal features[J]. Computers & Security, 2020, 89: 101681.
- [18] LIU X, LIU J. Malicious traffic detection combined deep neural network with hierarchical attention mechanism[J]. Scientific Reports, 2021, 11(1): 1-15.
- [19] ZHOU D, HUANG J, SCHÖLKOPF B. Learning with hypergraphs: clustering, classification, and embedding[C]//Twentieth Annual Conference on Neural Information Processing Systems(NIPS 2006). 2007: 1601-1608.
- [20] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv Preprint arXiv:1609.02907, 2016.
- [21] FENG Y, YOU H, ZHANG Z, et al. Hypergraph neural networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019: 3558-3565.
- [22] KARPATHY A. The unreasonable effectiveness of recurrent neural networks[J]. Andrej Karpathy Blog, 2015, 21: 23.
- [23] BERGE C. Hypergraphs: combinatorics of finite sets[M]. Elsevier, 1984.
- [24] GASTI P, TSUDIK G, UZUN E, et al. DoS and DDoS in named data networking[C]//2013 22nd International Conference on Computer Communication and Networks (ICCCN). 2013: 1-7.
- [25] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural

- network regularization[J]. arXiv Preprint arXiv:1409.2329, 2014.
- [26] TAVALLAEI M, BAGHERI E, LU W, et al. A detailed analysis of the KDD CUP 99 data set[C]//2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. 2009: 1-6.
- [27] MOUSTAFA N, SLAY J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)[C]//2015 Military Communications and Information Systems Conference (MilCIS). 2015: 1-6.
- [28] MOUSTAFA N, SLAY J. The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set[J]. Information Security Journal: A Global Perspective, 2016, 25(1-3): 18-31.
- [29] MOUSTAFA N, SLAY J, CREECH G. Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks[J]. IEEE Transactions on Big Data, 2017, 5(4): 481-494.
- [30] MOUSTAFA N, CREECH G, SLAY J. Big data analytics for intrusion detection system: statistical decision-making using finite dirichlet mixture models[M]//Data Analytics and Decision Support for Cybersecurity. 2017: 127-156.
- [31] SARHAN M, LAYEGHY S, MOUSTAFA N, et al. Netflow datasets for machine learning-based network intrusion detection systems[M]//Big Data Technologies and Applications. 2020: 117-135.
- [32] SHARAFALDIN I, LASHKARI A H, GHORBANI A A. Toward generating a new intrusion detection dataset and intrusion traffic characterization[C]//ICISSP. 2018: 108-116.

[作者简介]



赵文博（1998—），男，安徽淮北人，苏州大学硕士生，主要研究方向为图机器学习、网络安全、深度学习、推荐算法。



马紫彤（1999—），女，江西赣州人，苏州大学硕士生，主要研究方向为图机器学习、网络安全、深度学习、推荐算法。



杨哲（1978—），男，江苏苏州人，博士，苏州大学副教授，主要研究方向为网络与信息安全、深度学习、智能算法。