

基于图神经网络的暗网流量检测方法及实现

贾慰心¹⁾

¹⁾南京邮电大学 计算机学院 南京 210023

摘 要 随着匿名通信技术的发展和人们隐私保护意识的日益增强,暗网访问趋势正逐渐上升。Tor、I2P、Freenet等暗网匿名通信工具在保护用户隐私的同时,也被犯罪分子所利用。面临暗网中的非法活动威胁,亟需对暗网流量加强网络监管。如何有效地分析和检测暗网流量成为了一个重要的研究方向。本文在前人研究的基础上,尝试利用基于图的机器学习方法E-GraphSAGE M和E-ResGAT来设计并实现暗网流量检测与分析。在数据预处理阶段,本文使用图构造方法将数据集中的源(目的)IP地址和源(目的)端口组合起来作为图的节点,流量记录作为图的边,再将除源地址、目的地址、Flow ID、Timestamp以及两种标签外的其它信息作为边特征,形成一个二分图。通过此图构造方法将暗网流量检测问题转化为一个边分类任务,针对基于GAT的模型则将二分图转化成线图。在流量检测阶段,本文提出了一种基于E-GraphSAGE和GAT算法的改进方案,其核心思想是利用图信息将残差学习集成到图神经网络中。通过添加残差连接能够应对高度类别不平衡问题,能够在保留原始信息的基础上提高少数类的性能。在数据集CIC-Darknet2020上的结果表明,本文提出的方案能够有效处理高度类别不平衡问题,在二分类和多分类结果上各项指标均有明显提升。

关键词 图神经网络;暗网流量检测;残差网络;网络入侵检测;注意力机制
中图法分类号 TP393-08

A Detection Method and Implementation of Dark Web Traffic Based on Graph Neural Network

Jia Wei-Xin¹⁾

¹⁾School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, 210023

Abstract

In response to the issues of poor feature extraction, insufficient consideration of topological features, class imbalance, and lack of interpretability in existing encrypted traffic detection methods, this paper proposes an encrypted traffic detection model that integrates Graph Attention Networks (GAT) and edge feature embedding with residual networks (Edge-Graph Attention Residual Network, E-GA-RNet). First, traffic is preprocessed by combining the source IP address and port number, and the destination IP address and port number as the nodes of the graph, while the remaining flow features are used as edge features, transforming the encrypted traffic data into graph data. To adapt to the GAT algorithm, a new network traffic graph is constructed, where new nodes correspond to the edges of the original graph, and shared vertices in the original graph correspond to edges between two nodes. The traffic detection problem is then transformed into a node classification problem. Next, the GAT algorithm is used to calculate the attention coefficients for each node, aggregating and updating the features. Finally, residual connections are added to the original nodes to improve the performance of the minority class. Experimental results on the CIC-DarkNet dataset show that the proposed method can effectively address the class imbalance issue and significantly improve the performance in both binary and multi-class classification scenarios.

Keywords Graph Neural Network; Darknet Traffic Detection; Residual Network; Network Intrusion Detection; Attention Mechanism

1 绪论

1.1 研究背景与意义

近年来,互联网技术飞速发展,互联网用户不断增加。根据第55次《中国互联网络发展状况统计报告》显示,截至2024年12月,中国网民规模达11.08亿人,较2023年12月增长1608万人;互联网普及率达78.6%,较2023年12月提升1.1个百分点^[1]。互联网已经深入人们的生活中,网络空间已经是人类活动的“第五空间”,互联网已经把世界连接称为“地球村”,深刻改变着人类的生产生活方式。

互联网的快速发展丰富了人们的生活,但是同样也为网络犯罪提供了温床,特别是在暗网中。互联网像是一座冰山。人们大部分上网所处的区域是在互联网的表层,即明网,它是指能够被普通搜索引擎检索到的网络区域,仅占整个互联网的4%左右。而其他96%则是深网,其内容无法被普通搜索引擎检索到,它需要一定的访问权限才能被访问到^[2]。而本课题所涉及的暗网是深网的一个子集,需要通过特定的浏览器、特殊的设置才能访问的网络。典型的暗网架构包括Tor(The Onion Router)、I2P(Invisible Internet Project)、Freenet等,其中Tor使用最为广泛。据最新数据显示,2024年Tor全球平均日活用户已经超过10万人。据Chen等人的研究表明^[3],在自由度较低的国家,Tor用户通常是为了“数字权利”;而在自由度较高的国家中,Tor更多被用于“毒品市场”。不法分子选择暗网作为开展毒品贩卖、枪支交易、用户数据买卖、洗钱等非法犯罪活动的场所,主要是由于暗网具有匿名性强、不可追踪等特性。这些特性导致执法部门很难查清用户的真实身份以及活动轨迹。

网络空间不是“法外之地”。网络空间是虚拟的,但运用网络空间的主体是现实的。近年来,网络安全法、数据安全法、个人信息保护法等重要法律法规相继推出,表明了国家层面对于网络空间安全的重视。同时,近年来也打击了不少基于暗网的违法犯罪。例如:2019年南通、如东两级公安机关共抓获犯罪嫌疑人9名,其从暗网等非法渠道购得350余万条银行开户、手机注册等数据,贩卖公民个人信息数据数千万条,非法牟利70余万元;同年,无锡警方联合腾讯网御、微步在线等网络安全公司,成功告破国内首起暗网平台案。

在这样的背景下,如何有效分析和检测暗网流量成为网络空间安全治理的一大重要方向。吸引了一大批研究人员,为之添砖加瓦。其理论研究历程可以大致分为三个阶段^[4]:第一阶段是传统方法阶

段,主要依赖于基于规则和特征工程的方法,通过定义预先确定的规则或利用统计分析来检测异常。但是由于网络攻击技术的不断进步,传统方法的局限性随之显露,特别是对于新型和隐蔽的攻击模式。第二阶段是机器学习方法阶段,典型的机器学习方法有决策树、随机森林、支持向量机等等,它们能够利用数据进行学习,构建模型识别出异常流量,相较于传统方法有了显著提升。然而,机器学习方法通常需要手动标记流量,对数据的表征能力依赖大,不足以处理大规模数据和复杂任务。第三阶段为深度学习方法阶段,深度学习模型能够自动学习数据的高级表示,能够精准地捕获和识别复杂的异常流量和攻击行为,但其目前需要更多的计算资源并且模型的可解释性较差,且一般的基于深度学习的检测方法受限于神经网络架构,缺乏全局考虑。而基于图神经网络的检测方法能够捕捉流量的全局特征和结构信息,弥补当前方法的不足。

综上所述,基于图神经网络的暗网流量检测的方法与实现具有重要的理论意义和实践意义。本文以现有研究工作为基础,拟实现一种基于图神经网络的暗网流量检测方法,这种方法能够实现从海量网络流量中区分并识别出暗网流量,并进一步区分暗网流量的类别。

1.2 国内外研究现状

暗网基于对等计算机网络构建,结构十分复杂,其通信路由充满不确定性,这就使得暗网的检测工作难以进行。暗网的流量特性也十分复杂,不同类型的暗网流量差异很大,如Tor、I2P、ZeroNet、Freenet等暗网的用户行为流量各具特色。下面分别介绍基于机器学习和基于深度学习的暗网流量检测方法。

1.2.1 基于机器学习的暗网流量检测研究现状

基于传统机器学习的流量检测方法是从大量的流量数据中自动提取流量特征并加以分类。其对于复杂、非线性关系的处理能力强。与原始的基于规则和统计的方法相比,具有更好的自适应性和泛化能力,并且能够处理大规模、高维数据,自动调整检测模型以适应新的行为模式。不过其需要大量标记数据进行训练,而获取准确大量的标记数据困难大成本高^[5]。

近年来,学术界对应用于暗网流量检测方向上的机器学习方法有了很多改进。

由于当前暗网流量分类研究存在诸如大多聚焦单一暗网、未深入用户行为层面、缺乏公开数据集等局限性,Hu等人^[6]提出了一种三层分层分类方法,即将分类的颗粒度从是否为暗网流量,到是哪种暗网流量,最终再细化到用户行为层面。针对这种分类方法的有效性进行了全面阐述,他

们首先通过部署暗网数据探针，成功捕获了一个包含Tor、I2P、ZeroNet、Freenet这四种类型的真实暗网数据流量。再使用了六个机器学习算法和两个深度学习算法来评估分层分类器和XGBoost构建的最佳扁平分类器的性能，实验结果表明他们提出的分层分类器在暗网场景下识别效果更好。该研究不仅提供了一种可靠的对暗网流量进行细化分类的方法，还为检测模型提供了宝贵的真实暗网流量数据集。沿着这一研究脉络，Shi等人^[8]专注FreeNet的网络流量提出了一种分层分类方法，使用基于加权K-NN训练分类器，该分类器区分正常流量和FreeNet流量的平均准确率为99.6%，区分五种用户行为的平均准确率为95.8%，与现有的DT、Gaussian NB、K-NN等分类器相比，该分类器的准确率有明显提高。

而Rawat等人^[9]则在暗网流量检测中创新性地引入了TF-IDF (Term Frequency-Inverse Document Frequency) 特征提取算法。TF-IDF是一种用于信息检索与文本挖掘的常用加权技术。他们在暗网流量检测中使用TF-IDF算法是为了从网络流量数据中提取出更具代表性和区分度的特征，从而提高模型效果。然后Rawat等人使用LightGBM模型对暗网流量进行检测，其准确率达98.97%，提高了对暗网流量中犯罪行为的检测能力。

物联网技术快速发展，也成为暗网攻击的对象，但是目前缺乏高效的对于物联网下的暗网流量检测系统，因此Abu等人^[7]提出了一种基于机器学习的算法BAG-DT，并通过使用BAG-DT、ADA-DT、RUS-DT、O-DT、O-KNN、O-DSC这六种监督机器学习方法在CIC-Darknet-2020暗网流量数据集上进行检测分类实验，证实了BAG-DT具有更好的准确率，其分类准确率达99.5%。为物联网领域提供了一种可靠基于机器学习模型的暗网流量监测系统。

钟昱等人^[10]提出了一种合成数据增强的半监督网络异常流量检测方法SEASAND，利用无标记数据辅助模型学习，只需少量标记便可实现较高的准确率，在KDDCup99-10、UNSW-NB15、CICIDS2017数据集上表现出对于少样本、多分类问题有较好的性能。

1.2.2 基于深度学习的暗网流量检测研究现状

相较于传统的机器学习方法，深度学习方法能够自动学习数据特征，并且在处理大规模数据和复杂任务时表现优异，不过它需要更多的计算资源并且模型可解释性较差。近年来，深度学习凭借自动特征提取、泛化能力强等独特优势，受到国内外研究者的广泛探索。

Habibi等人^[11]提出了一种名为DeepImage的方

法，该方法通过特征选择选取最重要的特征来创建灰度图，并将该图通过二维卷积神经网络(2DCNN)来检测暗网流量，其中数据集采用ISCXVPN2016与ISCXTor2017的合并集，对该数据集的暗网流量检测准确率达到86%。Lan等人^[12]提出了一种用于暗网流量分类和应用程序识别的新型自注意力深度学习方法——DarknetSec。该方法利用一维卷积神经网络(1DCNN)和双向长短期记忆网络(Bi-LSTM)的级联模型从数据中提取局部时空特征，同时将自注意力机制集成到上述特征提取网络中，来挖掘先前提取的内容特征之间的内在关系。其在CICDarknet2020数据集上，DarknetSec实现了92.22%的多类准确率和92.10%的宏F1分数。

近年来，深度学习在暗网流量检测领域的应用逐渐从传统的CNN、RNN拓展到了图神经网络(GNN)。GNN因其出色的结构、节点和边缘信息处理能力受到研究者的广泛关注。它能够充分利用现实世界应用领域中遇到的大量数据的固有图结构，如：在计算机网络中，IP地址可以被视为图形的节点，而两个IP地址之间的通信可被视为图形的边。GNN利用图对网络拓扑和流量特征进行建模，提高了对流量之间相关性分析的能力，克服现有流量识别方法的不足。

Sun等人^[13]提出了一种结合流量跟踪图和图卷积网络(GCN)模型中统计特征的方法，该方法在标签较少的情况下能够实现性能提升，但是其准确性仍有待提高。Mo等人^[14]则开发了一种利用GCN进行加密流量分类的方法。此方法通过CNN捕获流量特征，并利用异构图卷积神经网络来识别网络范围内的行为模式，但是这可能会导致部分流级特征的丢失。Zhou等人^[15]提出GCNs用于分析结构信息和检测P2P僵尸网络节点。然而，该模型只学习目标网络的拓扑信息，忽略了节点和边缘特征，也不能检测其它类型攻击。Lo等人^[16]改进了GraphSAGE，开发了E-GraphSAGE算法应用于物联网的入侵检测，尽管其在边分类任务上表现出适用性，但在处理数据不平衡问题以及考虑节点间重要性差异方面仍有待加强。图分析改进了非欧几里德领域中的异常检测^[17]，但其计算成本较高。图嵌入通过将网络中的每个节点转换为低维表示来解决这个问题，但它缺乏泛化到不可见节点的能力。Abou等人^[18]提出，同图分析和图嵌入方法相比，归纳GNN在检测异常方面表现出更好的性能，且复杂度更低。Caville等^[19]提出了Anomal-E，一种基于GNN的自我监督网络入侵检测系统，用于减轻对标记数据的依赖，使用自监督方法进行入侵检测，对于两个现代基准NIDS数据集的实验结果表明，与原始特征和其他基线

算法相比,使用Anomal-E时有显著改进。Chang等人^[20]对GraphSAGE和GAT算法进行了改进,利用可用的图信息将残差学习集成到GNN中,以解决高级类不平衡问题,实现了在预测少数群体时出色的性能。

总而言之,不论是基于传统机器学习的流量检测方法还是基于深度学习的流量检测方法都存在着一些不足,比如:传统的机器学习方法针对从网络流量中提取的特征,其性能依赖于特征提取方法的选择,限制了检测的通用性;一般的基于深度学习的检测方法受限于神经网络架构,缺乏全局考虑。而基于图神经网络的检测方法能够捕捉流量的全局特征和结构信息,弥补当前方法的不足。

1.3 论文的主要工作

基于以上的国内外研究考量,本文将采用基于图神经网络的方法进行暗网流量检测和分析,其中将着重使用E-GraphSAGE M和E-ResGAT算法,它们分别是在现有的E-GraphSAGE和GAT算法基础上,引入了残差学习,将残差学习集成到图神经网络中。添加残差连接主要是为了应对类别不平衡问题,它能够在保留原始信息的基础上提升少数类别的分类性能。而CIC-Darknet2020数据集为我们提供了一个极具代表性的实验样本,能够适用于评估不同算法在暗网流量检测上的性能。所以在后续的实验中,本文重点使用该数据集作为实验流量样本。本文的主要工作内容如下:

(1) 构造图结构,从经过预处理后的数据集中提取出流量的拓扑信息。本文以源/目的地址及端口作为图的节点,流量记录作为图的边,其余信息作为边特征。利用这种拓扑结构构造出一个二分图 $G(S,D,E)$,其中 S 、 D 、 E 分别表示源节点集、目的节点集、边集。这样就将暗网流量检测任务转换为边分类任务。针对基于E-GraphSAGE的模型,该二分图可以直接用于边分类任务。而为了适应基于GAT的模型的点分类特性,我们将二分图进一步转化为线图,从而将边分类任务转换为点分类任务,这一过程进一步简化了暗网流量检测的复杂度。

(2) 在CIC-Darknet2020数据集上进行实验验证。通过与原模型E-GraphSAGE和GAT的对比实验,能够充分体现出本文改进的E-GraphSAGE M和E-ResGAT算法在处理暗网流量上的优势。

(3) 系统设计与实现,设计并实现一种暗网流量检测系统。该系统的功能具体包括数据处理、模型训练、暗网流量检测、检测效果评估,并在前端界面进行展示。

1.4 论文组织结构

本论文的主要结构以及内容安排如下:

第一章为绪论部分。该章首先从当前暗网中存在的一系列问题以及国家对于暗网犯罪的态度入手,展开介绍了暗网流量检测的相关背景和意义;然后分别从机器学习和深度学习在暗网流量检测领域的应用出发,介绍了目前国内外对于暗网流量检测的研究现状;最后对论文的主要工作内容和文章组织结构进行了简要说明。

第二章为理论基础部分。本章详细介绍了为实现暗网流量检测,使用到的理论和技术,重点介绍暗网流量和两种现有的基于图神经网络的流量检测算法。

第三章提出基于图神经网络的暗网流量检测方法。本章分模块对基于图神经网络的暗网流量检测及实现进行阐述,首先介绍数据预处理和图构建,然后着重说明两种基于图神经网络的暗网流量检测方法:E-GraphSAGE M和E-ResGAT的设计和实现过程,再补充说明实验设置,最后进行实验结果的评估与分析。

第四章为了实现并应用暗网流量检测功能,本章设计并实现了一个暗网流量分析的原型系统,提供用户使用接口,然后对系统的主要功能模块进行展示。

2 理论基础

2.1 暗网流量

人们通过传统搜索引擎访问到的只是互联网中的表层网络,与表层网络相对的,被称为深网。深网是指搜索引擎和超链接无法直接访问的资源,包括:需要登录的网站、数据库、付费服务等^[21]。本论文所研究的暗网是深网的一部分,主要被用于进行非法活动。暗网通常无法被普通搜索引擎直接访问,仅能通过电脑上的一系列特殊操作设置或者使用特定的软件才能访问到,如:Tor(The Onion Router)、I2P(Invisible Internet Project)、Freenet等。暗网和深网之间的关系如图2.1所示。

暗网通常架构在匿名通信系统之上,需使用特殊框架才能访问。暗网中的匿名系统主要分为两类:高延迟系统和低延迟系统。典型的高延迟系统,如:第二代洋葱路由器和Mixmaster协议。它们会采用混合、重新排序、修补等技术将数据包的传输顺序和时间特性进行打乱,使得攻击者难以利用数据包的时间信息进行分析,能够更好地保护用户隐私,但是这些技术会产生额外的延迟,使得数据传输速度变慢,影响用户体验。经典的低延迟系统,如:Tor、JAP、I2P、VPN等。它们不会使用额外的技术,传输速度快,能够很好地适应HTTP、SSH等对实时交互性要求高的协议,能

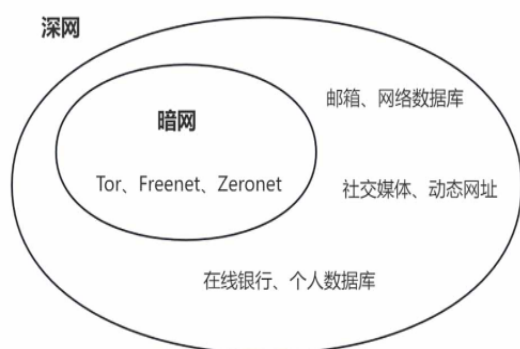


图 2.1 深网和暗网的关系

够更好地适应用户一般地网络访问需求，但是抵御网络检测分析的能力弱。本论文中对于暗网流量的检测与分析，主要是针对低延迟通信系统的。

暗网具有匿名性强、去中心化、路由动态性大等特性^[22]。这些特性使得暗网网络通信行为与传统网络存在明显差异，这种差异直接体现在暗网流量的特性上（以Tor流量为例）：

（1）协议层隐匿性：采用洋葱路由协议，通过多层加密和固定长度数据单元来掩盖流量特征；

（2）拓扑结构不稳定性：入口节点和出口节点间的中继路径不断变化，导致流量行为模式发生动态变化；

（3）元数据模糊性：用户和网站的真实IP和端口信息均被系统性混淆。

这些特性差异使得传统的基于载荷分析的检测方法和基于IP/端口的检测方法在暗网场景下均失效。这就让我考虑到使用基于图神经网络的检测方法。尽管在暗网中真实通信实体的身份会被隐藏，流量记录中仍包含模拟或脱敏的端点标识符，如数据集中虚拟分配的源/目的IP和端口。这些标识符虽然不是真实网络中的IP地址，但是仍然能够反映出流量交互的抽象拓扑关系。因此，本论文采用图神经网络，通过建模流量实体之间的逻辑拓扑关系（以数据集中提供的IP+端口作为图节点，流量记录作为图的边）来进行暗网流量检测与分析。

2.2 图神经网络基础模型

图神经网络（Graph Neural Network, GNN）是指使用神经网络学习图结构数据，提取和发掘图结构数据中的特征和模式^[23]。图神经网络是一种拓展和推广的神经网络，区别于一般神经网络，图神经网络能够直接对非欧氏数据结构进行表示学习和分析预测。而现实中的问题大多都只能抽象成非欧氏数据结构，这就使得GNN受到广泛关注并迅速发展。近年来，各种GNN框架相继被提出，在各

个领域的图结构数据中都发挥出了重要作用。

而在暗网流量检测领域中，图采样和聚合（Graph Sample and aggreGatE, GraphSAGE）和图注意力网络（Graph attention networks, GAT）这两种算法模型被研究者们广泛使用。

2.2.1 GraphSAGE

GraphSAGE模型是由William L. Hamilton等人^[24]于2018年提出，其解决了传统图嵌入方法在归纳学习中的局限性，能够利用节点特征信息为未见数据生成节点嵌入。

该模型的主要特点是学习节点的局部领域特征聚合函数，而非直接为每个节点训练单独的嵌入，这样就使得其能泛化到未见节点或全新子图，能够很好地适应大规模图和动态图。GraphSAGE采样聚合方法可视化过程如图2.2所示。

GraphSAGE算法流程，首先是邻域采样，即以目标节点为中心，按照不同半径（跳数k-hop）对其领域节点进行采样。一般在考虑计算效率的前提下，会指定采样数量，若目标节点邻居数少于采样数量，则采取有放回的采样方法，否则采取无放回的抽样。然后再使用聚合函数对采样到的领域节点的特征信息按层进行聚合，而对于GraphSAGE中的每一层聚合过程可以使用不同的聚合函数，常用的聚合函数有：均值、池化、LSTM（长短期记忆网络）等，最后一层的节点特征就作为目标节点的嵌入。最后再利用目标节点嵌入进行节点分类任务，通常会使用一个softmax层来获取最终的分类结果。

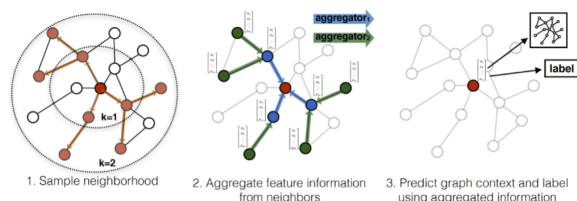


图 2.2 GraphSAGE采样聚合可视化

2.2.2 GAT

GAT^[25]是将传统的图卷积网络与注意力机制相结合的网络模型架构。与传统的图卷积网络相比，GAT不需要使用拉普拉斯矩阵进行节点状态更新，而是在每个节点状态更新时，引入注意力机制计算其邻居节点的重要程度，为每个邻居节点分配不同的权重，重点关注重要程度大的节点，从而提高模型的计算效率。

GAT架构通过堆叠多个图注意力层，逐层提取和学习图中的节点和边的特征表示，每一层都是基

于上一层的输出进行进一步特征提取。最终将经过多层传播得到的节点特征输入到分类器中，如全连接层，然后分类器输出每个节点属于不同类别的概率分布，最后使用softmax得到最终分类结果。

对于单个图注意力层，输入是一组节点特征 $\vec{h} = \{\vec{h}_1, \vec{h}_2, \dots, \vec{h}_N\}$, $\vec{h}_i \in \mathbb{R}^F$ ，其中 N 是节点数， F 是每个节点中的特征数；输出一组新的节点特征 $\vec{h}' = \{\vec{h}'_1, \vec{h}'_2, \dots, \vec{h}'_N\}$, $\vec{h}'_i \in \mathbb{R}^{F'}$ ，具体的计算方法如式2.1：

$$\vec{h}'_i = \sigma \left(\sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{W} \vec{h}_j \right) \quad (\text{式2.1})$$

其中的注意力系数的计算原理图，如图2.3。

为了稳定自注意学习过程，可以采用多头注意力机制，即：使用 K 个独立的注意力机制，并行执行式2.1的变换，然后将它们的输出特征进行连接，如式2.2所示，而在最终的预测层上则采用平均而非连接。

$$\vec{h}'_i = \prod_{k=1}^K \sigma \left(\sum_{j \in \mathcal{N}_i} a_{ij}^k \mathbf{W}^k \vec{h}_j \right) \quad (\text{式2.2})$$

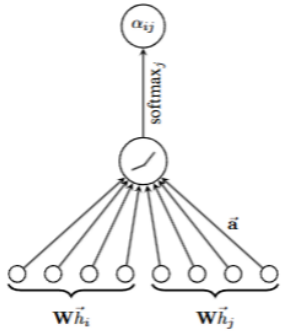


图 2.3 注意力系数的计算原理

在暗网流量检测过程中存在良性流量和恶意流量类别极端不平衡现象，通过引入注意力机制能够为不同邻居分配权重，提高对恶意流量的检测能力。因此GAT算法能够很好的适用于暗网流量检测领域。而本文使用的E-ResGAT算法在GAT基础上，进一步引入残差连接，允许模型在学习过程中保留原始特征信息，防止信息丢失。这样能够更好地发挥GAT算法的优势。

3 基于图神经网络的暗网流量检测

针对暗网流量检测中存在的动态拓扑变化和极端类别不平衡问题，本章基于图神经网络提出

了两种改进模型：E-GraphSAGE M和ResGAT。E-GraphSAGE M是在E-GraphSAGE基础上引入动态邻居采样、边特征聚合以及残差连接，它能够很好地适应拓扑变化和拓展，能够捕捉到流量地交互特性，能够缓解梯度消失；而E-ResGAT是在GAT模型基础上融合残差连接与多头注意力机制，能够缓解深层网络下的特征退化问题并且能很好地兼顾局部拓扑特征与全局关系。

本研究的实验流程分为以下三个阶段：首先对CIC-Darknet2020数据集进行数据预处理并进行图结构的构建；然后进行二分类任务，即暗网流量和正常流量，验证基础检测性能；再进一步进行多分类任务，即识别用户行为类型，评估模型细粒度分析能力。

3.1 图神经网络模型设计

E-GraphSAGE和GAT模型在暗网流量检测中被研究者们广泛使用，能够很好地适应暗网环境。我们使用的改进的E-GraphSAGE M和E-ResGAT模型通过增加残差连接，使得针对暗网的检测效果得到了进一步提高，本节主要介绍E-GraphSAGE M和E-ResGAT模型的具体设计。

3.1.1 E-GraphSAGE M模型

GraphSAGE模型关注节点特征，在节点分类任务中表现出色，但是其无法有效捕捉边的信息，无法充分利用图的拓扑信息，而在暗网流量检测中，网络流量通常表现为图的边，而E-GraphSAGE最大的特点就是支持边缘分类，其核心思想是在边的两个节点上分别执行GraphSAGE算法，然后将产生的两个节点嵌入合并起来作为边表示。在本研究中的E-GraphSAGE M是对E-GraphSAGE的进一步改进，使用残差连接来克服原始边特征潜在的信息丢失问题。E-GraphSAGE M算法在采样方面，它采用均匀采样固定大小邻居集的方法进行采样，具体的采样算法如算法3.1所示。

Algorithm 3.1 E-GraphSAGE 小批量采样

Require: 图 $G(V, E)$; 小批量边 B ; 点集 $V(B) = \{v \mid v \text{ 是边 } uv \in B \text{ 的一个端点}\}$; 深度 K ; 邻域采样函数 $N_v : v \rightarrow 2^V$;

Ensure: 小批量 B^0 的 2 跳邻域

- 1: $B^K \leftarrow B$
- 2: **for** $k = K, \dots, 1$ **do**
- 3: $B^{k-1} \leftarrow B^k$
- 4: **for** $v \in V(B^k)$ **do**
- 5: $B^{k-1} \leftarrow B^{k-1} \cup \{uv \mid \forall u \in N_v\}$
- 6: **end for**
- 7: **end for**

E-GraphSAGE M的聚合过程是一个逐层迭代的操作，主要是为了将邻居边的特征逐步聚合到节点上，来更新节点的特征表示，具体过程如算法3.2所示。对于第k层，这层的输入是边特征 x_{uv} ，其中 $uv \in B^0$ ；算法还会将k-1层的邻居边特征聚合到节点v上，生成 $h_{N_v}^k$ ，见式3.1，其中 AGG_k 表示聚合函数，常见的聚合函数有均值、池化、图卷积或者长短期记忆网络等。在本研究中采用均值聚合函数。在得到 $h_{N_v}^k$ 后，会进行如算法3.2中第5行所示的操作来获取当前层的节点嵌入，其中 W_k 表示可训练的权重矩阵，表示激活函数。由于网络中图的节点初始时是没有特征的，使用全1向量来初始化节点特征。

$$h_{N_v}^k = AGG_k(\{h_{uv}^{k-1} \mid \forall u \in N_v\}) \quad (式3.1)$$

E-GraphSAGE的最终边嵌入是将最后一层的节点嵌入进行连接，而E-GraphSAGE M算法，在E-GraphSAGE最终边嵌入的基础上，再将原始边特征 e_{uv} 也连接进来，这样能够更好地保留原始边特征信息，提高模型性能。

Algorithm 3.2 E-GraphSAGE 小批量聚合

Require: 图 $G(V, E)$ ，边小批量数据 B ，2跳邻域 B^0 ；点集 $V(B) = \{v \mid v \text{ 是边 } uv \in B \text{ 的一个端点}\}$ ；输入边特征 $\{e_{uv} \mid \forall uv \in B\}$ ；输入节点特征 $x_v = 1$ ；深度 K ，非线性激活函数 σ ；权重矩阵 $W^k, \forall k \in \{1, \dots, K\}$ ；不同的聚合函数 AGG_k ；邻居采样函数 $N_v : v \rightarrow 2^V$ ；

Ensure: 向量表示 $z_{uv}, \forall uv \in B$

- 1: $h_{uv}^0 \leftarrow e_{uv}, \forall uv \in B^0$
 - 2: $h_v^0 \leftarrow x_v, \forall v \in V(B^0)$
 - 3: **for** $k = 1$ 到 K **do**
 - 4: **for** $v \in V(B^k)$ **do**
 - 5: $h_v^k = \sigma(W^k[h_v^{k-1} \parallel h_{N_v}^{k-1}])$
 - 6: **end for**
 - 7: **end for**
 - 8: $z_{uv} = \parallel([h_u^K, h_v^K, e_{uv}]), \forall uv \in B$
-

3.1.2 E-ResGAT模型

在E-GraphSAGE M算法中，所有相邻的边都被同等对待，但事实上不同邻居对节点的重要程度有所不同。因此，本研究使用带有残差的图注意力网络（E-ResGAT），让模型能够学习对邻居进行加权聚合。E-ResGAT算法的采样过程与E-GraphSAGE M算法的类似。E-ResGAT算法是使用注意力机制来聚合邻域信息，并在每一层中都连接了原始节点特征的变换，即进行残差学习。具体的

聚合算法见算法3.3。在第k层，基于注意力的残差聚合可以表示为式3.2。其中， a_{uv} 是分配给线图中边 e_{uv} 的注意力系数， W 是跨层共享的线性变换，将输入特征映射到较低维度。注意力系数 a_{uv} 可以通过前馈神经网络 $a[\mathbf{W}h_u \parallel \mathbf{W}h_v]$ 学习到的，其中 a 是指权重向量。然后再通过LeakyReLU激活函数和softmax函数得到所有节点对的注意力系数，如式3.3所示。类似于GAT算法，E-ResGAT也能通过应用多头注意力来增加模型容量，多头E-ResGAT聚合公式为式3.4。其中 M 是注意力头的数量， a_{uv}^m 是第 m 个归一化注意力系数， W^m 是第 m 个权重矩阵。最后将原始节点特征 e_{uv} 与所有 M 个矩阵的平均值 \bar{W}^m 进行连接，如式3.5所示。

$$h_v^k = \sigma\left(\sum_{u \in N_v} a_{uv} \mathbf{W} h_u^{k-1}\right) \parallel (\mathbf{W}' e_v) \quad (式3.2)$$

$$a_{uv} = \frac{\exp(\text{LeakyReLU}(a[\mathbf{W}h_u \parallel \mathbf{W}h_v]))}{\sum_{i \in N_v} \exp(\text{LeakyReLU}(a[\mathbf{W}h_i \parallel \mathbf{W}h_v]))} \quad (式3.3)$$

$$h_{N_v}^k = AGG\left(\left\|\sigma\left(\sum_{u \in N_v} a_{uv}^m \mathbf{W}^m h_u^{k-1}\right)\right\|_{m=1}^M\right) \quad (式3.4)$$

$$h_v^k = h_{N_v}^k \parallel (\mathbf{W}' e_v) \quad (式3.5)$$

Algorithm 3.3 E-ResGAT 小批量聚合

Require: 图 $G'(V', E')$ ；小批量边 B ，2跳邻域 B^0 ；节点特征 e_v ；层数 K ；注意力头数量 M ；权重矩阵 $W^m, \forall m \in \{1, \dots, M\}$ ；非线性激活函数 σ ；节点 v 的邻域 $N_v : v \rightarrow 2^V$ ；

Ensure: 向量表示 $z_v, \forall v \in B$

- 1: $h_v^0 \leftarrow e_v, \forall v \in B^0$
 - 2: **for** $k = 1$ 到 K **do**
 - 3: **for** $m = 1$ 到 M **do**
 - 4: $e_{uv} = \text{LeakyReLU}(a[\mathbf{W}^m h_u^{k-1} \parallel \mathbf{W}^m h_v^{k-1}])$
 - 5: $a_{uv} = \text{softmax}(e_{uv}) = \frac{\exp(e_{uv})}{\sum_{i \in N_v} \exp(e_{iv})}$
 - 6: **end for**
 - 7: $h_{N_v}^k = \prod_{m=1}^M \sigma(\sum_{u \in N_v} a_{uv}^m \mathbf{W}^m h_u^{k-1})$
 - 8: $h_v^k = h_{N_v}^k \parallel (\mathbf{W}' e_v)$
 - 9: **end for**
 - 10: $z_v = h_{N_v}^K \parallel \mathbf{W}' e_v, \forall v \in B$
-

3.2 数据集与图结构构建

3.2.1 数据集

本研究实验部分选用公开数据集CIC-Darknet2020进行模型的训练和测试。CIC-Darknet2020数据集是由新不伦瑞克大学加拿大网络安全研究所（CIC）于2020年发布的公开数据集，专门用于暗网流量检测和分析。其核心目标是为加密流量分类、恶意活动检测等研究提供标准化数据支持。

该数据集具体通过Wireshark和tcpdump来捕获流量，生成pcap文件，然后使用ISCXFlowMeter工具读取并解析pcap文件，提取每条流量的统计特征，最终生成csv文件。

该数据集整合了ISCXTor2016和ISCXVPN-2016数据集，覆盖了更全面的暗网场景，共包含141530条样本和85个特征，且这些特征是对当前流量的统计特征。从标签结构上看，该数据集分为两层，第一层为正常流量和暗网流量，其中正常流量包含了NonTor和NonVPN，而Tor和VPN则是暗网流量，流量详情见图3.1；第二层为暗网流量类别的进一步细分，由特定应用程序生成，涵盖八类常见应用场景：Audio-stream、Chat、Browsing、File-transfer、P2P、VIOP、Email、Video-stream，流量详情信息如图3.2所示。该数据集提供了丰富的暗网流量样本，为研究人员开发和验证新的暗网流量检测方法提供了基础，是当前暗网流量检测领域的重要资源。

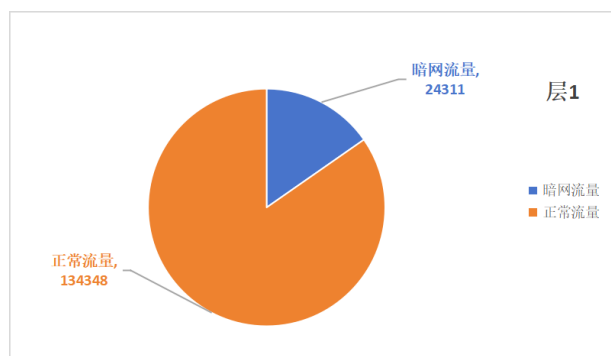


图 3.1 第一层流量类别分布

3.2.2 数据预处理

数据预处理是进行暗网流量检测的关键环节之一，数据预处理的质量会直接影响到后续进行图构建和模型训练的效果。本研究的数据集是公开数据集CIC-Darknet2020，使用Python读取数据集文件。在对数据集内容进行一定了解后，再对数据集进行预处理。

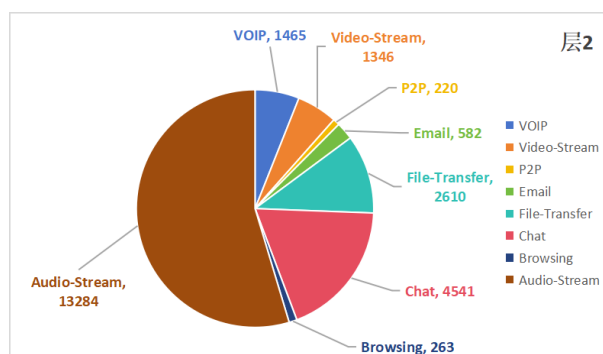


图 3.2 第二层应用场景细分

首先是进行数据清洗。对缺失值和异常值进行处理，原数据集中包含Nan值，我们对这些Nan值进行赋0，以解决可能出现的问题。然后再删除无效特征列，如：时间戳、流ID，减少数据冗余。然后进行地址映射。将源IP地址映射到172.16.0.1到172.31.0.1的地址区间，这样做一方面能够减少线图内存占用，另一方面还能防止源节点为恶意流量提供无意标签，避免潜在的分类偏差，从而提高模型的准确性和鲁棒性。并将源/目的IP地址和源/目的端口使用“:”进行合并，便于后续处理和分析。

再将分类标签进行编码，使用数字来代替具体类别，方便进行分类评估。其中值得注意的是在对多分类标签进行编码时，把二分类标签是NonTor和NonVPN的对应的多分类标签都定为0，然后对暗网流量对应的多分类标签，进行编码，这样就使得多分类标签一共有9种。

最后将处理好的数据保存。将节点（IP地址+端口）信息、节点构成的边以及两种标签进行保存为.npy文件。再将其余的信息作为边特征，进行标准化处理后，也进行保存为.npy文件。

3.2.3 图结构构建

图结构构建就是将网络流量数据转化为图结构，以适配图神经网络模型的输入要求。

首先是进行原始二分图的构建，把网络流量数据中的源地址和目的地址定义为图的节点，因为节点无先验特征，将节点特征初始化为全1向量；将每条流量记录作为图的边，而边特征就是经过标准化后的流量的统计特征。这样就构造出了一个二分图 $G(S, D; E)$ ，其中 S 、 D 、 E 分别表示源节点集、目的节点集、边集。这样就将暗网流量检测问题转化为对图上的边的分类任务。这些数据都经过预处理，单独保留出来了，在具体实现图构建时，可直接使用数据预处理后的信息。

由于E-ResGAT算法是对于节点进行分类的，

为适应其特点，我们需要通过线图转换重构拓扑，将二分图转化为线图。具体就是将二分图中的边映射为线图的节点；若原图中两条边共享同一个节点，则在新图中创建一条边。转化过程可视化如图3.3所示。其中左图为二分图，其中S和D分别是指源节点集和目的节点集。这里的S和D都只包含3个节点，且不相交。边连接不同集合中的两个节点，图中共有7条边。这些边和点就构成了二分图。右侧是线图。线图上的每个节点对应二分图的一条边，在线图中共有7个节点。线图上的边是二分图中对应两条边存在共同的节点，就比如，线图上的节点1a和1b相连，是因为在二分图中，连接1-a和1-b的两条边都有共同节点1。

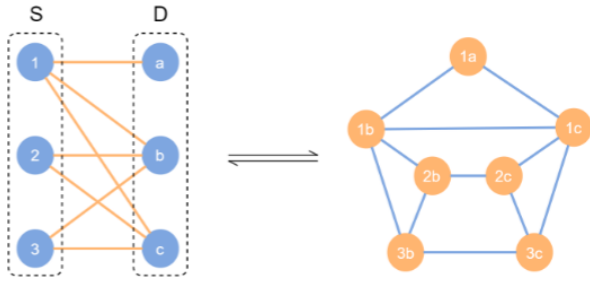


图 3.3 二分图（左图）与线图（右图）的互相转化

3.2.4 数据集划分

基于前文对数据集的分析可知，暗网流量数据分布极不平衡。为确保模型训练和评估结果的真实性和可靠性，本文采用随机采样，严格遵守原始数据的标签分布特征，对数据集进行划分，使训练集、验证集和测试集上的数据分布情况，与现实情况保持一致。具体而言，测试集包含50000条样本，约占总数据量的30%，主要用于评估模型；验证集由5000条样本构成，占比3%，主要用于模型训练过程中的效果检验与超参数调整；其余67%的数据为训练集，用于模型的学习和优化。

3.3 实验设置与结果分析

3.3.1 实验环境

本文实验在Windows11操作系统21H版本下进行，处理器为AMD Ryzen 7 5800H with Radeon Graphics，3.20 GHz频率，16GB内存。采用的编程语言是Python3.7，集成开发环境是Pycharm和Jupyter Notebook，使用anaconda搭建虚拟环境。深度学习框架是Pytorch，版本是1.13.1+cu117。

3.3.2 评估指标

由于本研究实验共有二分类和多分类，两项不同的分类任务，其效果的评估指标也有所不同。

其中对于二分类任务，采用准确率（Accuracy）、精确率（Precision）、召回率（Recall）和F1分数作为评估模型二分类效果的指标。对于多分类任务，考虑到类别分布的不均衡，本文采用加权平均Precision、加权平均F1、加权平均Recall，以及宏观平均Precision、宏观平均F1、宏观平均Recall来对多分类效果进行评估。

在二分类评估中，准确率是指分类正确的样本数占总样本数的比例，用于衡量模型分类的正确程度。精确率是指预测为正的样本中真正为正的样本所占比例，它反映了模型预测出正例的准确性。召回率是指实际为正的样本中预测为正的样本所占比例，用于衡量模型能够正确识别出的正例的能力。F1分数是精确率和召回率的调和平均数，能够更全面地评估模型性能。这些指标的计算方法如下：

$$\text{Acc(Accuracy)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (\text{式3.6})$$

$$P(\text{Precision}) = \frac{TP}{TP + FP} \quad (\text{式3.7})$$

$$R(\text{Recall}) = \frac{TP}{TP + FN} \quad (\text{式3.8})$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (\text{式3.9})$$

其中TP是指预测和实际均为异常的样本个数，TN是指预测和实际都为正常的样本个数，FP是指预测为异常实际为正常的样本个数，FN是预测为正常实际为异常的样本个数。

在多分类任务中，因为不同类别样本的数量差异很大以及不同类别的重要程度也不同，为了更好地评估模型的性能，常常会使用加权平均和宏观平均作为评估指标。本文就选用这两种评估指标来衡量模型表现，其中加权平均指标会以每个类别的样本数量占总样本数量的比例为权重，对每个类别的指标进行加权求和，综合考虑了各类别样本数量差异，能够更客观地反映模型在不同类别上的表现。与之对应的宏观平均指标采取的做法是先分别计算每个类别的评估指标，然后取平均，它平等地看待每一个类别，能够从整体上反映模型在多分类任务的性能。式3.10和式3.11分别是加权平均指标和宏观平均指标的计算方法，其中C是指多分类的类别数， N_i 是该类别样本数， N_{total} 是总样本数。

$$P_w = \sum_{i=1}^C w_i P_i, \quad w_i = \frac{N_i}{N_{total}} \quad (\text{式3.10})$$

$$P_m = \frac{1}{C} \sum_{i=1}^C P_i \quad (\text{式3.11})$$

本研究模型的效率性能是使用训练模型所花费的时间来评估的。该时间是从AMD Ryzen 7 5800H with Radeon Graphics, 3.20 GHz频率, 16GB内存的Windows上获取的。批次训练时间以秒为单位。

3.3.3 实验细节设置

基于E-GraphSAGE的模型的实验设置: 设置模型层数 $K=2$, 即构建两层结构模型; 对于最后的边分类任务, 使用的是softmax分类器, 它将模型的输出转换为各边类别对应的概率值, 然后从中选择概率最大的边类别作为最终分类结果。邻居采样采用2跳8领域采样, 即以2跳邻居为范围进行采样, 且每条采样大小为8。聚合函数使用均值函数, 将邻居节点的特征取平均值, 作为当前节点聚合后的特征表示。使用ReLU作为非线性激活函数, 其能够引入非线性因素, 使得模型能够学习到更复杂的模型和特征, 同时还具备计算简单、收敛速度快等优点。

基于GAT的模型的实验设置: 设置模型层数 $K=3$, 且每一层都使用6头注意力机制。丢弃率设置为0.2, 防止模型过拟合。最后的分类任务同样也使用softmax分类器进行分类。邻居采样选择全邻域, 充分利用邻居信息。选用ELU(指数线性单元)作为非线性激活函数。ELU能够在引入非线性的同时, 使神经元的输出均值接近0, 有助于加快模型收敛速度, 提高模型训练效果。

通用设置: 所有模型均使用Pytorch深度学习框架实现, 该框架提供了灵活的张量计算功能和自动求导功能。使用Adam优化器进行训练, 设置学习率为0.03。Adam优化器是一种自适应学习率的优化算法, 能够在不同参数上分别调整学习率, 有助于模型更快更稳定地收敛。损失函数使用交叉熵损失函数, 该损失函数常用于分类任务, 能够很好地衡量预测概率分布与真实标签分布之间的差异, 从而引导模型优化。对于训练过程的设置, 采用分批次训练方式, 每个批次的大小为500。使用小批量训练能够减少内存占用, 加快训练速度, 特别是对于基于GAT的模型效果显著, 同时又引入了一定的随机性, 能够有助于跳出局部最优解。训练轮次均为两个轮次(epochs)。训练轮次需根据训练集规模、模型复杂程度等因素合理设置, 过少训练可能会欠拟合, 过多则可能过拟合。

3.3.4 二分类暗网流量实验

将CIC-Darknet数据集中的网络流量数据的label列作为二分类标签, 将网络流量分为正常流量和暗网流量这两类。通过训练测试

原模型E-GraphSAGE和GAT及其改进后的模型E-GraphSAGE M和E-ResGAT, 进行对比实验。实验结果如表3.1所示:

表 3.1 基于图神经网络的暗网流量二分类评估结果

方法	评估指标				
	Accuracy	Precision	Recall	F1	AUC
E-GraphSAGE	0.9087	0.7280	0.6462	0.6846	0.9011
E-GraphSAGE_M	0.9139	0.7444	0.6678	0.7040	0.9479
GAT	0.9168	0.7388	0.7072	0.7227	0.9452
E-ResGAT	0.9220	0.7543	0.7285	0.7514	0.9553

可以看到改进后的模型E-GraphSAGE M和E-ResGAT评估指标相比于原模型都有所提升, 其中E-GraphSAGE M的F1分数提升了2%左右, E-ResGAT的F1分数提升了3%左右, 这验证了引入残差连接对边特征的保留效果。而我们对比基于GAT的和基于GraphSAGE的两组模型, 可以明显看出GAT相关模型的效果要优于GraphSAGE相关的模型, 这能够充分体现注意力机制对动态路由的适应性。

3.3.5 多分类暗网用户行为分析实验

将CIC-Darknet数据集中数据预处理后的label.1列作为多分类标签, 共有9种标签, 具体分布情况如表3.2所示。使用原模型E-GraphSAGE和GAT及其改进后的模型E-GraphSAGE M和E-ResGAT分别进行训练和测试, 通过对比验证改进后模型的优越性。首先是整体性能的分析, 使用加权平均和宏观平均作为核心评估指标, 具体实验结果如表3.3所示。实验结果表明, 从总体上看改进后的模型的多分类效果要比原模型效果要好一些, 在大部分指标上来看, 基于GAT模型的多分类效果仍比基于E-GraphSAGE的模型效果要好, 体现出GAT架构在处理多分类任务时的优势。

表 3.2 数据类别分布情况

类别	0	1	2	3	4	5	6	7	8
个数	42337	4188	83	1431	823	183	69	424	462

表 3.3 基于图神经网络的暗网多分类评估结果

方法	加权平均评估指标			宏观平均评估指标		
	Precision	Recall	F1	Precision	Recall	F1
E-GraphSAGE	0.8480	0.8817	0.8546	0.5166	0.3036	0.3240
E-GraphSAGE_M	0.8641	0.8847	0.8613	0.6104	0.3860	0.3888
GAT	0.8598	0.8822	0.8645	0.4794	0.3003	0.3330
E-ResGAT	0.8885	0.8867	0.8826	0.5456	0.4193	0.4335

然后是对每个分类的具体细化分析，如表3.4所示。从整体上来看，E-ResGAT算法展现出最为突出的性能，在9个分类类别中，有5类的分类效果最好，E-GraphSAGE M有2个类别上分类效果最好，原始的算法各有1个最好。再结合数据类别分布情况，可以明显看出所有模型对于样本数量较少的类别3、4、5、7、8的分类效果都不是很好，不过同样值得注意的是对于极少数类2、6，E-GraphSAGE M具有最好的分类效果，体现出该模型在处理极少数类的优势。

表 3.4 基于图神经网络多分类F1 对比

方法	各类F1								
	0	1	2	3	4	5	6	7	8
E-GraphSAGE	0.9374	0.6422	0.3448	0.0081	0.0667	0.0000	0.7819	0.0456	0.1909
E-GraphSAGE.M	0.9449	0.6199	0.5688	0.0148	0.2500	0.0000	0.7969	0.1166	0.1875
GAT	0.9484	0.6122	0.3593	0.0892	0.2951	0.0103	0.5396	0.1882	0.1917
E-ResGAT 0.9526	0.6286	0.4600	0.2396	0.2880	0.0909	0.6838	0.1912	0.1942	

最后我们使用t-SNE对所有四个模型学习到的特征表示进行可视化。具体而言，基于E-GraphSAGE的模型选取最后一层的输出作为可视化对象，由于基于GAT的模型的最后一层是全连接层，故选取倒数第二层的输出作为可视化对象。见图3.4可视化结果直观显示带有残差连接的模型相比原模型，在聚类效果上更为出色，各分类之间具有更高的分离度，进一步证实了残差连接有助于提升模型对不同类别特征的区分能力。

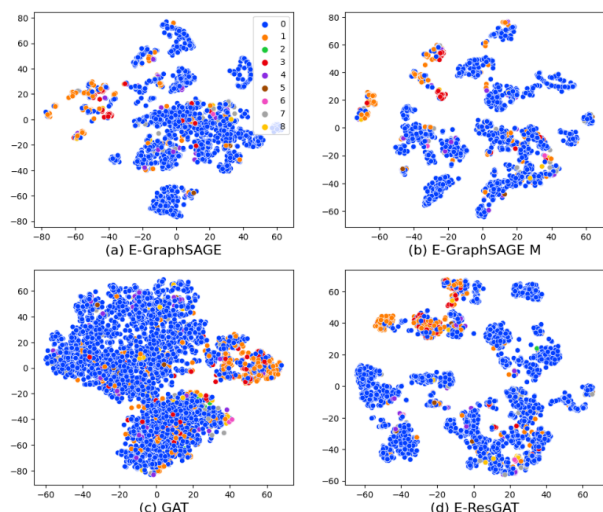


图 3.4 四种模型的t-SNE图

从上面的两个实验来看，基于GAT的模型虽然在评估指标上整体上要优于基于E-GraphSAGE的模型效果，但是如果在从效率性能上来看，后者更具优势。实验结果显示，每个训练批次，基于E-GraphSAGE的训练所需时间是0.25分钟左右

右，而基于GAT的模型训练所需时间是8.6分钟左右。GAT相关模型完成训练所消耗的时间是基于E-GraphSAGE模型的30倍以上。而引入残差连接对训练时间影响几乎可不计。进一步分析时间消耗的构成可以发现，基于GAT的模型训练时大部分时间用于原始二分图到线图的结构转换上。因为基于GAT的模型采取的是全采样，即选择与一批边共享公共端点的所有边，来构建该批边的完整邻域结构，而非基于E-GraphSAGE的局部邻域采样。若排除结构转换所消耗的时间，基于GAT的模型的实际训练时间仅是基于GraphSAGE模型的3到5倍。由此可见，基于E-GraphSAGE的模型凭借更低的时间成本，在实际暗网流量实时检测场景下具有更强的适应性，能够更高效地处理动态流量数据，满足暗网流量检测领域对检测实时性的严格要求。

4 界面开发与设计

为了提高模型训练操作的便捷性，并使训练结果能够可视化的呈现，本研究基于Flask开发框架，设计了可视化交互界面。

Flask是一个轻量级的Python Web框架，以简洁灵活著称，非常适合快速开发小型项目。它遵循“微框架”的设计理念，仅提供核心功能（如路由、请求响应处理），其他功能需要通过拓展实现，开发者拥有很高的自由度。在前后端开发方面，Flask采用前后端分离技术，前端使用HTML、CSS、JavaScript实现用户交互，后端根据Flask提供的RESTful API接口进行数据处理和模型调用，前后端之间通过JSON格式进行数据交换。

本研究的系统界面效果的具体渲染采用Tailwind CSS框架，利用其丰富的类名实现快速样式编写和响应式控制，具体效果如图4.1所示。本系统共包含四个核心功能模块，各模块通过事件监听机制实现参数联动：

（1）算法选择模块：该模块对应于界面中标签为“Algorithm”的部分，采用HTML的下拉选择框来实现。用户可以通过此模块来选择基础模型，共有GAT和E-GraphSAGE两种。

（2）数据集加载模块：即界面中标签为“Dataset”的部分，同样使用HTML的下拉选择框。用户可以选择具体的数据集，目前支持UNSW-NB15、Darknet、CES-CIC、ToN-IoT这四种公开数据集。

（3）任务设置模块：即界面中的“Binary Classification”复选框，用于设置分类任务，当用户勾选该复选框时，系统将执行二分类任务，若未勾选，则执行多分类任务。

(4) 残差连接模块：通过界面中的“Residual”复选框实现，用来设置是否进行残差连接。当用户勾选该复选框时，在模型训练过程就进行残差连接，即论文中的改进后的模型，否则不进行残差连接，即论文中的原始基础模型。

再通过两个按钮“Train Model”和“Test Model”来控制模型训练和测试。

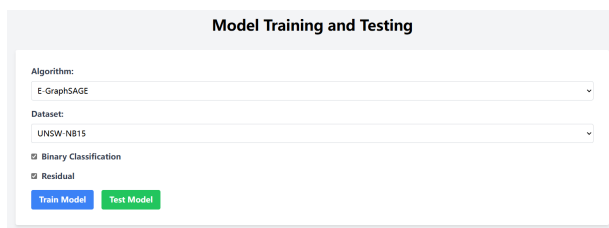



图 4.1 界面设计图

在选择好参数后，点击“Train Model”按钮，可以进行模型训练，最后会输出训练花费的时间，以及训练过程的loss函数，具体如图4.2所示。其中loss曲线的展示的实现主要依赖于后端Python代码中Matplotlib库的调用，训练过程中记录每次迭代的损失值，等到训练结束后在使用Matplotlib绘制loss曲线，并将该图表保存到指定静态文件夹中，最后再使用HTML的标签加载静态图片来展示训练的loss曲线。

为了防止重复训练模型、提高模型测试的效率，系统实现了模型的存储和加载功能。在模型训练好后，系统会将训练好的模型参数保存为.pth文件。在进行模型测试时，系统首先会检查指定路径下是否存在对应的模型文件，若存在则直接加载模型进行测试；若不存在则先执行模型训练，再进行模型测试。

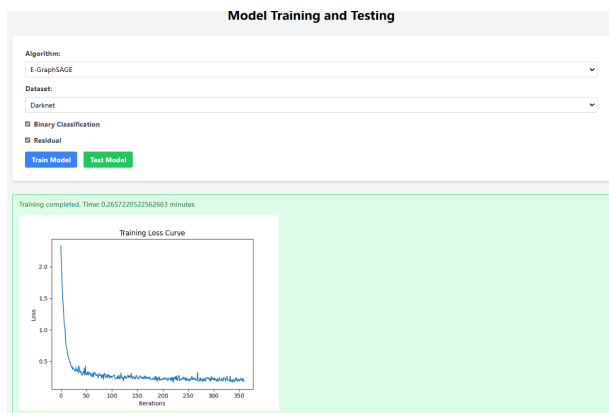


图 4.2 训练模型显示

对于模型测试，同样也需要先进行模型参数选择，选择好具体模型参数后，再点击“Test

Model”按钮进行模型测试。完成好模型测试后，系统会根据所选的分类任务，输出各自的评估指标，二分类任务会展示准确率、精确率、召回率以及F1分数，而多分类任务则会展示对应的宏平均和微平均指标，多分类任务测试的输出如图4.3所示。

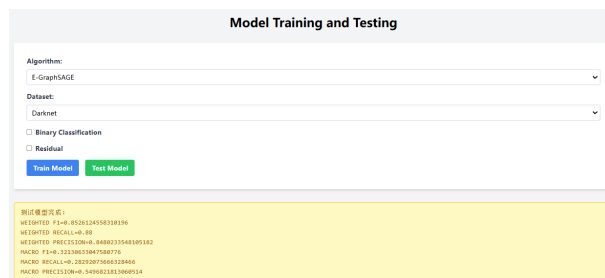


图 4.3 训练模型显示

5 结束语

随着匿名通信技术的不断发展和人们隐私保护意识的日益增强，访问匿名网络的用户数量呈逐年递增趋势。匿名通信技术为普通用户带来更加安全的网络环境的同时，也为不法分子所利用，滋生出了一系列的违法犯罪活动，如非法倒卖个人信息、毒品枪支交易等等。面临这些问题，对暗网流量进行检测和分析，对于维护国家安全和保障公民个人权益尤为重要。国内外针对暗网犯罪已制定了一系列法律法规，但是由于暗网流量具有极强的匿名性和动态路由特性，犯罪信息获取尤为困难。在此背景下，提升暗网流量检测能力成为打击违法犯罪的关键突破口。

本论文针对暗网流量检测中动态路由混淆和极端类别不平衡两大挑战，基于现有图神经网络模型E-GraphSAGE和GAT，创新性地引入残差连接，实现了改进模型E-GraphSAGE M和E-ResGAT，为提高暗网流量检测效能提供了新的思路。本文在暗网数据集CIC-Darknet2020上进行实验，以此评估模型的效能。利用数据集中流量的拓扑结构来构建图数据结构，并针对GAT模型特性将二分图转换为线图，分别进行了二分类和多分类任务的实验。实验结果表明，通过引入残差连接的两个模型的各项评估指标都优于原模型，特别是E-ResGAT模型综合性能最佳，这充分体现了残差连接和注意力机制的有效性。并且引入残差连接对模型时间效率的影响微乎其微。

本论文为暗网流量检测提供了可解释性的图表示学习框架，证明了拓扑关系建模对加密流量分析的重要性。但是研究仍存在一定的局限性，如：动态图构建的计算开销较大，对于新型暗网协议的泛

化能力不足, 以及多分类实验显示模型对部分特定类型的用户应用访问分析不够敏感。所以对于暗网流量的检测工作, 还需要更多的工作。未来的研究可从优化图构建方法、结合数据增强算法、使用多源数据融化提高模型泛化能力等方向开展, 比如开发批量邻域构建算法, 降低E-ResGAT的时间复杂度; 结合生成对抗网络合成少数类样本, 来缓解数据不平衡问题^[26]。

通过本次毕业设计, 我深刻认识到个人能力的局限性与技术发展的无限性, 切实体会到网络安全领域面临的复杂挑战。未来将持续学习, 提升自身专业能力, 致力于在网络安全领域贡献出自己的一份力。

参考文献

- [1] 中国互联网络信息中心. 第55次《中国互联网络发展状况统计报告》[R/OL]. (2025-01-17) [2025-05-08]. <https://www2.cnnic.cn/>
- [2] 罗军舟, 杨明, 凌振等. 匿名通信与暗网研究综述[J]. 计算机研究与发展, 2019, 56(01): 103-130.
- [3] Chen Z, Jardine E, Fan Liu X, et al. Seeking anonymity on the Internet: The knowledge accumulation process and global usage of the Tor network[J]. new media & society, 2024, 26(2): 1074-1095.
- [4] 陈吉祥. 基于数据增强与深度学习的暗网流量检测与分析方法研究[D]. 华南理工大学, 2022.
- [5] 李海龙, 崔治安, 沈燮阳. 网络流量特征的异常分析与检测方法综述[J]. 信息网络安全, 2025, 25(02): 194-214.
- [6] Hu Y, Zou F, Li L, et al. Traffic classification of user behaviors in tor, i2p, zeronet, freenet[C]//2020 IEEE 19th international conference on trust, security and privacy in computing and communications (TrustCom). IEEE, 2020: 418-424.
- [7] Abu Al-Haija Q, Krichen M, Abu Elhaija W. Machine-learning-based darknet traffic detection system for IoT applications[J]. Electronics, 2022, 11(4): 556.
- [8] Shi, KaiChao, et al. "Layered classification method for darknet traffic based on Weighted K-NN." 2022 International Conference on Networking and Network Applications (NaNA). IEEE, 2022.
- [9] Rawat R, Mahor V, Chirgaiya S, et al. Analysis of darknet traffic for criminal activities detection using TF-IDF and light gradient boosted machine learning algorithm[C]//Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2021. Springer Singapore, 2021: 671-681.
- [10] 钟昱, 黄振南, 谢惠超, 陈宁江. 一种基于半监督学习的网络异常流量检测方法[J]. 广西大学学报(自然科学版), 2024, 49(3): 563-574
- [11] Habibi Lashkari A, Kaur G, Rahali A. Didarknet: A contemporary approach to detect and characterize the darknet traffic using deep image learning[C]//Proceedings of the 2020 10th International Conference on Communication and Network Security. 2020: 1-13.
- [12] Lan J, Liu X, Li B, et al. DarknetSec: A novel self-attentive deep learning method for darknet traffic classification and application identification[J]. Computers & Security, 2022, 116: 102663.
- [13] Sun B, Yang W, Yan M, et al. An encrypted traffic classification method combining graph convolutional network and autoencoder[C]//2020 IEEE 39th International Performance Computing and Communications Conference (IPCCC). IEEE, 2020: 1-8.
- [14] Mo S, Wang Y, Xiao D, et al. Encrypted traffic classification using graph convolutional networks[C]//Advanced Data Mining and Applications: 16th International Conference, ADMA 2020, Foshan, China, November 12-14, 2020, Proceedings 16. Springer International Publishing, 2020: 207-219.
- [15] Zhou J, Xu Z, Rush A M, et al. Automating botnet detection with graph neural networks[J]. arXiv preprint arXiv:2003.06344, 2020.
- [16] Lo W W, Layeghy S, Sarhan M, et al. E-graphsage: A graph neural network based intrusion detection system for IoT[C]//NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium. IEEE, 2022: 1-9.
- [17] Jiang, Weiwei. "Graph-based deep learning for communication networks: A survey." Computer Communications 185 (2022): 40-54.
- [18] Abou Rida A, Amhaz R, Parrend P. Evaluation of anomaly detection for cybersecurity using inductive node embedding with convolutional graph neural networks[C]//Complex Networks & Their Applications X: Volume 2, Proceedings of the Tenth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2021 10. Springer International Publishing, 2022: 563-574.
- [19] Caville E, Lo W W, Layeghy S, et al. Anomal-E: A self-supervised network intrusion detection system based on graph neural networks[J]. Knowledge-Based Systems, 2022, 258: 110030.
- [20] Chang L, Branco P. Graph-based solutions with residuals for intrusion detection: The modified e-graphsage and e-resgat algorithms[J]. arXiv preprint arXiv:2111.13597, 2021.
- [21] 陆雨楠. 针对Tor的匿名流量识别[D]. 中国人民公安大学, 2024.
- [22] Saleem J, Islam R, Islam M Z. Darknet traffic analysis: A systematic literature review[J]. IEEE Access, 2024, 12: 42423-42452.
- [23] 吴博, 梁循, 张树森, 等. 图神经网络前沿进展与应用[J]. 计算机学报, 2022, 45(01): 35-68.
- [24] Hamilton W, Ying Z, Leskovec J. Inductive representation learning on large graphs[J]. Advances in neural information processing systems, 2017, 30.
- [25] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J]. arXiv preprint arXiv:1710.10903, 2017.

-
- [26] Xie R, Wang Y, Cao J, et al. Rosetta: Enabling robust tls encrypted traffic classification in diverse network environments with tcp-aware traffic augmentation[C]//Proceedings of the ACM Turing Award Celebration Conference-China 2023. 2023: 131-132