

基于双重自编码器和残差注意力机制的异常检测方法

伍晓洁, 刘 强

国防科技大学 计算机学院 长沙 中国 410073

摘要 随着具有复杂关联关系的网络攻击行为层出不穷,传统基于流式数据的网络异常检测方法面临着严峻的安全挑战。为了应对这一挑战,基于图深度学习技术的网络异常检测技术研究方兴未艾。然而,现有的研究工作在深层次图结构特征学习方面还有改进空间,有限的节点特征学习能力影响了网络异常检测效果。为此,本文提出了一种基于双重自编码器和残差注意力机制的异常检测方法 ResDAE,该方法利用基于残差分析的注意力机制,可抑制来自异常节点的信息在图神经网络中传播,降低了解码器的重构误差,进而提升了异常节点的分类能力。具体来讲,ResDAE方法考虑了双重自编码器在重构过程中产生的残差信息,并在结构自编码器上引入基于残差分析的注意力模块,从而有效地捕获了网络结构与节点属性之间的隐式关系;另一方面,该方法使用了图神经网络来学习节点的嵌入表示,提升了图嵌入表示能力。基于三个公开数据集的对比实验结果表明,ResDAE在检测性能上要优于现有的图异常检测方法。此外,针对僵尸网络检测任务,基于CTU-13子数据集的离线实验结果表明,ResDAE方法获得了 Precision 指标和 Recall 指标得分均超过85%的良好性能,相比其他对比方法,其AUC指标的性能表现最佳,从而进一步验证了本文方法在这一特定检测任务下的优越性能以及残差注意力机制带来的性能增益。

关键词 图深度学习; 异常检测; 双重自编码器; 残差注意力机制

中图法分类号 TP399 DOI号 10.19363/J.cnki.cn10-1380/tn.2025.09.08

An Anomaly Detection Method Based on Dual Auto-Encoder with Residual Attention Mechanism

WU Xiaojie, LIU Qiang

College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China

Abstract In the rapidly developing era of information and data, cyberattacks pose a serious threat to human privacy, work and even the safety of life and property. As an important equipment for human to communicate in daily work, and for entertainment and data storage, the host has become the main target of cyberattacks. Therefore, it is urgent and necessary to study the host attack discovery technology. And as a carrier to record all behaviors in the host, host events have become the focus of research. Attackers' various malicious operations in the host will inevitably be recorded as host events. However, malicious events are hidden in large-scale normal events, that are difficult to detect and filter out, leading to academic research on a series of issues such as how to obtain host events, how to identify and extract malicious events, how to reconstruct the attack process, and how to conduct security protection. This paper has carried out extensive research and detailed summary on the research related to host event based attack discovery technology, and combed its research and development history, then compares with intrusion detection and digital forensics from four aspects: analysis object, analysis method, time of action, and analysis purpose. After, the definition of host event based attack discovery technology was conducted. Subsequently, this paper explains the key concepts involved. Two major problems which are dependency explosion and timeliness are pointed out. Then the research is divided into three categories according to the stages: host event collection, host event processing and host event analysis. The research results and progress of the three categories around the two major problems, which are 12 subcategories in total, are introduced respectively. Finally, the possible research directions in the future are pointed out according to the current research situation, including integrity and credibility of host event records, timeliness of attack discovery, cross-device attack discovery, multi-step attack discovery, and algorithm application.

Key words graph deep learning, anomaly detection, dual auto-encoder, residual attention mechanism

通讯作者: 刘强, 博士, 副教授, Email: qiangliu06@nudt.edu.cn.

本课题得到湖南省自然科学基金项目(No. 2021JJ30779)资助。

收稿日期: 2023-09-26; 修改日期: 2024-02-20; 定稿日期: 2025-08-14

1 引言

随着网络空间成为继陆、海、空、天、电之后的第六维战略疆域,网络安全成为维护国家主权、安全和发展利益的重要基石。僵尸网络、分布式拒绝服务攻击、垃圾邮件、中间人攻击等多样化网络攻击行为层出不穷,给网络安全保障带来了严峻的风险挑战。因此,以网络异常检测为代表的网络安全防御技术持续得到了网络安全学术界的重点关注,相关研究方兴未艾^[1]。另外,伴随人工智能与机器学习技术的蓬勃发展,网络安全与人工智能交叉研究成为当前的一个研究热点^[2],已有研究将聚类分析^[3]、强化学习^[4]、图神经网络^[5]等技术应用到异常检测之中,并取得了良好的效果。然而,大多数基于流式数据的智能检测方法不便于刻画动态变化的网络拓扑结构以及主机之间的通信交互行为。因此,使用图结构数据建模网络流量并引入图机器学习技术开展网络异常检测研究是一个新增长点^[6-12],例如,文献[11]针对僵尸网络检测问题,从僵尸网络流数据中构建一个带有属性特征的通信图,并提出了一种基于属性图的异常网络节点检测方法。当前,结合图神经网络和自编码器(Auto-Encoder, AE)的异常检测方法吸引了许多学者的关注^[13-16],其中, Dominant 方法^[14]结合图卷积网络与深度自编码器来学习节点嵌入表示,并通过结构和属性两个特征维度测量节点的重构误差以发现异常行为; DeepAE 方法^[15]则通过一个图卷积自编码器来捕捉拓扑结构和节点属性内的非线性关系,并应用拉普拉斯锐化过程保留了正常和异常节点之间的差异。考虑到单通道自编码器难以捕捉到重构过程中属性图结构与节点属性的跨模态交互特征, AnomalyDAE 方法^[16]通过引入一个双重自编码器设计了一种深度联合表达学习框架,该方法在计算节点的重构误差时综合考虑了网络结构嵌入表示和节点属性信息,从而具备更佳的异常行为检测性能。

综上所述,基于图神经网络的网络异常检测方法已取得了一些进展,然而,有限的节点特征学习能力影响了网络异常检测性能,具体在深层次图结构特征学习方面还有改进空间。为了解决这一问题,本文提出了一种基于双重自编码器和残差注意力机制的异常检测方法 ResDAE,该方法通过在双重自编码器中引入基于残差分析的注意力机制,更加凸显了正常和异常图节点之间的差异,其中,残差来自

双重自编码器的重构过程。具体来讲, ResDAE 使用一个残差矩阵来设计注意力机制,可抑制异常节点信息在图神经网络中传播,从而提升了双重自编码器的图嵌入表示能力和异常节点分类能力。基于 BlogCatalog、Flickr、ACM 和 CTU-13 四个公开数据集的多组对比实验结果表明,本文方法相比现有异常检测方法具有更好的性能表现。

本文的主要贡献如下:

(1) 提出了一种基于双重自编码器与残差注意力机制的异常检测方法 ResDAE,该方法在结构自编码器上引入基于残差分析的注意力模块,有效地捕获了网络结构与节点属性之间的隐式关系,提升了异常检测能力。

(2) 基于 3 个公开图数据集的多组对比实验结果表明, ResDAE 对比现有的 5 种基准检测方法具有更佳的性能。为了对比不同方法针对僵尸网络检测任务的性能优劣,本文进一步使用了 CTU-13 公开数据集开展对比实验。针对不同的 CTU-13 子数据集, ResDAE 方法的 AUC 性能指标表现均显著优于对比方法,相关结果也验证了本文方法的优越性。 ResDAE 方法的实现代码可公开获取^①。

本文后续内容结构组织如下:第 2 节介绍了相关工作;第 3 节形式化地描述了图自编码器和图异常检测问题;第 4 节详细介绍了本文提出的方法;第 5 节基于多个权威数据集开展对比实验和性能评价;第 6 节总结本文工作并展望下一步工作。

2 相关工作

随着信息技术的发展,互联网深度融合到人们生活的方方面面。然而,互联网容易遭受多样化恶意攻击行为的渗透破坏,进而威胁到网络用户的信息安全和数据隐私。为此,网络入侵检测系统成为当今不可或缺的网络安全防护组件之一^[17,18]。近年来,基于图深度学习开展的网络入侵检测技术及其应用研究工作取得了许多进展,而针对僵尸网络这一致命安全威胁的检测技术研究是其中的一个重要方面^[18-20]。Chowdhury 等人^[8]提出一种基于图节点拓扑特征的僵尸网络检测方法,考虑了节点的入度、出度、入度权重、出度权重、聚类系数、节点介数和特征向量中心性。利用聚类方法建立网络中节点的簇,将僵尸主机节点隔离在正常簇之外。并通过过滤算法删除图中不活跃的节点,以进一步提高算法效率。但该方法只考虑了通信图中的拓扑特征,而忽略

了通信节点之间的流特征,在一定程度上影响着僵尸主机检测的准确性。Daya 等人^[9]提出一个基于图的两阶段检测方法 BotChase, 第一阶段将可推测为良性的主机节点修剪, 第二阶段进行图特征选取与特征归一化, 以克服拓扑效应带来的负面影响, 实现高精度的僵尸主机节点检测。该方法对零日攻击具有鲁棒性, 且适用于大型数据集。但该方法在选取特征时同样忽略了通信图中基于流的特征。Wang W 和 Shang Y 等人^[10]提出一种自动化的僵尸网络检测模型 BotMark, 该方法通过混合分析基于流和基于图的网络流量来检测僵尸网络, 提取了 15 个基于统计流的流量特征和 3 个基于图的特征来构建检测模型。针对基于流的检测, 将流的相似性和稳定性作为检测度量。针对基于图的检测, 使用最小二乘技术和局部离群因子来计算通信图中节点邻域差异的异常分数, 其中异常节点的邻域与正常节点的邻域显著不同。BotMark 集成基于相似性分数、稳定性分数以及异常分数的检测结果来确定异常的僵尸主机。Afna Alharbi A 等人^[11]提出了一种基于机器学习的僵尸网络检测模型, 该模型首先考虑了图特征的重要性, 然后根据所选择的重要特征开发检测僵尸网络的通用模型。该模型从一致性、相关性、信息性等不同理论角度出发, 使用 5 种基于过滤器的特征评价指标来生成通信图最具有代表性的特征集, 以提高算法的性能和可扩展性。

在具有属性特征的通信图上检测僵尸主机节点的问题可以转化为在属性图上进行异常节点检测的问题。因此, 优化基于图的异常节点检测问题的求解过程可提升僵尸网络检测性能。由于自我中心网络^[21]、社区分析^[22]或子空间选择^[23]的方法只能解决针对属性图的浅层学习机制的问题, 容易受到网络稀疏性和数据非线性问题的影响, 并且无法捕获不同信息模态之间的复杂交互, 从而影响异常检测的性能。残差分析^[24-26]作为发现异常节点的另一种方式, 将异常定义为无法从其他节点中近似得到的节点, 但基于线性模型的分析无法捕捉网络的非线性特点, 且难以从输入网络中自适应地建模。由上述分析可知, 需要一个统一的特征空间来描述属性网络中属性特征与结构特征两种模态之间复杂的交互关系, 以进行异常检测。基于自编码器设计的检测方法能够在一定程度上缓解上述问题^[14-16]。例如, Dominant 方法^[14]是一种基于属性网络的深度异常检测方法, 它使用图卷积神经网络作为编码器函数, 将输入的属性网络压缩为简洁的低维嵌入表示, 然后利用相应的解码器函数来重构拓扑结构和节点属性,

基于编码器和解码器之间的节点重构误差来检测属性网络上的异常节点。然而, 单纯地基于 GCN 进行编码, 缺乏对不同节点之间隐含关系的描述。DeepAE 方法^[15]则通过深度自编码器将拓扑结构和节点属性无缝地嵌入到统一的表示中, 充分挖掘了网络的内在信息, 捕获了局部和全局的关系信息, 使得隐藏层中的节点属性的语义信息与原网络保持一致。引入拉普拉斯锐化来保持正常节点和异常节点之间的差异, 通过利用两种信息模态的重构误差来发现异常。然而, 该方法的复杂性取决于图的大小以及卷积层的维度, 在大型稠密图场景计算时面临复杂度较高的问题。AnomalyDAE 方法^[16]提出了一种双自编码器的深度联合表示学习异常检测框架, 通过捕获网络结构和节点属性之间复杂的交互来获得高质量的嵌入。由结构自编码器和属性自编码器组成, 通过分别重构原始网络拓扑和节点属性来联合学习节点和属性的潜在表示。基于结构和属性两个角度测量节点的重构误差来检测网络中的异常。然而, 在编码器中使用多层感知机学习节点表示对高层次隐藏特征的提取具有一定限制。

3 图自编码器与图异常检测问题描述

3.1 面向图学习的自编码器

自编码器是一种无监督学习模型, 可用于获取输入数据样本的低维表示, 它由两个部分组成, 分别是编码器和解码器。在面向图学习的自编码器中, 对于图中的每个节点 $v_i \in V$, 其中 V 为节点集合, 它在邻接矩阵中的对应行 $a_i = A_i$ 作为编码器的输入, 得到其低维表示:

$$z_i = f_{\text{enc}}(a_i; \Theta_{\text{enc}}) \quad (1)$$

其中, f_{enc} 表示编码器(Encoder), 对应参数为 Θ_{enc} 。

之后, z_i 作为解码器的输入以期重构 a_i , 即

$$a'_i = f_{\text{dec}}(z_i; \Theta_{\text{dec}}) \quad (2)$$

其中, f_{dec} 表示解码器(Decoder), Θ_{dec} 是对应的参数。最后, 该自编码器的重构损失定义为

$$\mathcal{L}_{\text{enc}} = \sum_{v_i \in V} \|a_i - a'_i\|_2^2 \quad (3)$$

通过最小化式(3)中的重构损失, 可将节点 v_i 的邻居信息压缩到低维表示 z_i , 并保持原始输入数据的相对距离关系。通常来讲, 邻接矩阵 A 是稀疏的, 直接采用式(3)计算重构损失会使得优化过程中更倾向于重构出元素 0。针对这一问题, 需要对非零元素的重构误差增加惩罚项, 重构损失计算式变为

$$\mathcal{L}_{\text{enc}} = \sum_{v_i \in \mathcal{V}} \|(a_i - \tilde{a}_i) \odot b_i\|_2^2 \quad (4)$$

其中, $\odot$ 表示阿达玛积, $b_i = \{b_{i,j}\}_{j=1}^V$, 当 $A_{i,j} = 0$ 时, $b_{i,j} = 1$; 当 $A_{i,j} \neq 0$ 时, $b_{i,j} = \beta > 1$ (β 为需要调整的超参数)。此外, 为了使相连的节点具有相似的低维表示, 引入如下正则化损失:

$$\mathcal{L}_{\text{con}} = \sum_{v_i, v_j \in \mathcal{V}} A_{i,j} \cdot \|(z_i - z_j)\|_2^2 \quad (5)$$

进一步地, 引入针对编码器和解码器参数的正则化损失, 即

$$\mathcal{L}_{\text{reg}} = \|\Theta_{\text{enc}}\|_2^2 + \|\Theta_{\text{dec}}\|_2^2 \quad (6)$$

最后, 待优化的目标函数表示为

$$\mathcal{L} = \mathcal{L}_{\text{enc}} + \lambda \cdot \mathcal{L}_{\text{con}} + \eta \cdot \mathcal{L}_{\text{reg}} \quad (7)$$

在上述图自编码器定义基础上, 双重自编码器由两个编码-解码通道构成, 包含了结构自编码器和属性自编码器, 其中, 结构自编码器负责学习与重构节点结构特征, 属性自编码器则负责学习与重构节点属性特征。另外, 为了充分捕捉结构与属性之间跨模态的隐式关系, 属性解码器部分综合使用结构编码器与属性编码器学习到的嵌入表示来生成重构信息。

3.2 基于图学习的异常检测问题定义

3.2.1 属性网络

本文将属性网络定义为 $G = \{V, \varepsilon, A, X\}$, 它是一个无向图, V 、 ε 、 A 和 X 分别表示节点集、边集、邻接矩阵和属性矩阵, 其中 $n = |V|$ 是节点数, $|\varepsilon|$ 表示边的数量, 每个节点 $v_i \in V$ 有一个 $m = |X|$ 维的属性向量 $x_i \in \mathbb{R}^m$ ($i = 1, 2, \dots, n$), $X = \{x_{i,j}\}_{i=1, 2, \dots, n}^{j=1, 2, \dots, m}$ 。属

性网络中节点与边之间的关系通过邻接矩阵 A 来表示, $A = \{a_{i,j}\}_{i,j=1}^n$, 当节点 v_i 与节点 v_j 之间有边相连接, 则 $a_{i,j} = 1$; 否则, $a_{i,j} = 0$ 。

3.2.2 异常检测任务

异常检测旨在检测出那些在结构和属性方面与大多数节点表现明显不一致的节点, 可形式化描述为学习一个评分函数 $f: v_i \rightarrow y_i \in \mathbb{R}$, 该函数能够根据阈值 λ 对节点 v_i 进行分类, 用公式表示为

$$y_i = \begin{cases} 1, & \text{if } f(v_i) \geq \lambda \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

其中, y_i 是节点 v_i 的预测标签, 0 表示正常类, 1 表示异常类。

针对基于双重自编码器的网络异常检测, 给定一个属性网络 $G = \{V, \varepsilon, A, X\}$, 首先学习一个映射 $f: \{A, X\} \rightarrow \mathbb{R}^d, d \ll |A|$, 将网络结构的拓扑关系和节点属性都保留在低维空间的表示中。随后, 根据异常评分函数的计算结果来检测出与多数节点显著不同的异常节点。

4 ResDAE 方法设计

基于双重自动编码器与残差注意力机制的异常检测方法(ResDAE)主要由一个双重自编码器模块和一个注意力机制模块组成, 如图 1 所示, 其中, 属性网络 G 的邻接矩阵 A 和属性矩阵 X 是输入, 在基于残差注意力机制的作用下, 结构编码器利用邻接矩阵和属性矩阵的信息学习属性网络 G 的隐藏层表示 Z^A 。随后, 结构解码器利用隐藏层表示计算出重构

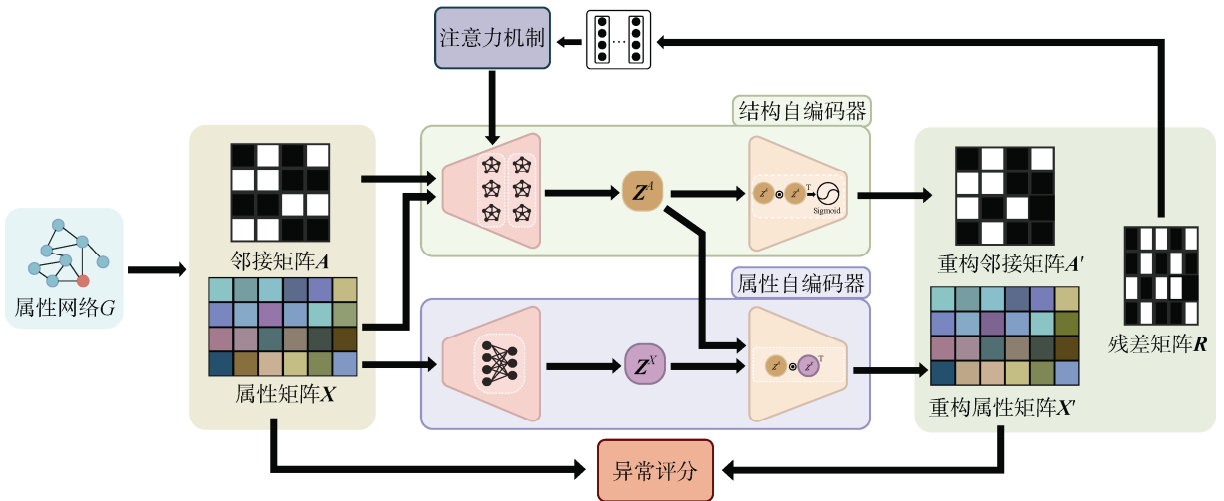


图 1 ResDAE 架构示意图

Figure 1 Illustration of the ResDAE framework

邻接矩阵 $\mathbf{A}'$ 。另一方面, 属性矩阵 $\mathbf{X}$ 作为属性自编码器的输入, 通过多层感知机网络计算得到属性的隐藏层嵌入表示 $\mathbf{Z}^X$ 。随后, 属性解码器结合 $\mathbf{Z}^X$ 与结构编码器学习到的网络嵌入表示 $\mathbf{Z}^A$ 计算出重构属性矩阵 $\mathbf{X}'$ 。ResDAE 通过深度神经网络来捕获异常节点与正常节点之间的显著性差异, 每个节点的残差数据组成残差矩阵 $\mathbf{R}$ 。随后, ResDAE 将 $\mathbf{R}$ 输入至一个全连接层并把处理后的结果输入到注意力机制模块中, 以将残差信息添加至结构编码器的训练过程, 进而优化属性网络的嵌入表示。最后, ResDAE 综合结构重构误差、属性重构误差和残差矩阵的信息来计算节点的异常分数, 据此检测存在异常的节点。

更进一步地, ResDAE 方法在结构编码器中使用图卷积神经网络来处理属性网络所具有的稀疏性和非线性。在属性解码器中结合结构编码器与属性编码器生成的嵌入向量来重构属性矩阵, 以此捕获结构特征与属性特征之间跨模态的隐式关系。此外, 异常节点重构前后的残差值较大, 注意力机制模块将赋予异常节点更小的注意力权重, 以降低异常信息在消息聚合时的影响。

4.1 基于残差的注意力机制

本文将残差定义为观测值与预估值之间的差异, 产生于属性矩阵的编码-重构过程, 可形式化地描述为

$$\mathbf{R}^{(l+1)} = f_{\text{ReLU}}(\mathbf{R}^{(l)} \cdot \mathbf{W}^{(l)}), \quad (9)$$

其中, $\mathbf{R}^{(l)}$ 为第 l 层全连接层(FC)的输入, $\mathbf{R}^0 = \mathbf{X}$, $\mathbf{W}^{(l)}$ 是第 l 层的权重矩阵。ResDAE 方法将通过式(9)计算得到的每一个 FC 层的输出作为相应图卷积神经网络(Graph Convolution Network, GCN)层的注意力权重。因此, 学习权重的 FC 层数与结构自编码器中的 GCN 层数相同。

4.2 结构自编码器

ResDAE 在设计结构自编码器时, 使用 GCN 作为结构编码器, 通过式(10)的方式聚合邻居传递的信息

$$\mathbf{H}^{(l+1)} = \sigma\left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(l)} \mathbf{W}^{(l)}\right), \quad (10)$$

其中, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_n$ 为无向图 G 带自环的邻接矩阵, $\mathbf{I}_n$ 为一个大小为 n 的对角矩阵, n 为节点集大小, $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_{ij}$, $\mathbf{W}^{(l)}$ 为第 l 层的可训练权重矩阵, $\sigma(\cdot)$ 表示激活函数, 本文取 $\text{ReLU}(\cdot) = \max(0, \cdot)$, $\mathbf{H}^{(l)}$ 为第 l 层的节点嵌入表示。为了阻止异常节点信息在网络中的传播, 本文利用基于残差信息的注意力机制为邻居节点分配不同的权重, 残差越大的节点权重

越小, 权重可形式化为

$$\gamma_i^{(l)} = \exp(-\xi \cdot \mathbf{R}_i^{(l)}), \quad (11)$$

其中, $\gamma_i^{(l)}$ 为节点 v_i 的第 l 层注意力权重, ξ 为超参数。考虑了注意力权重后, GCN 聚合邻居节点信息时需要计算对应的权重占比, 即

$$\mathbf{H}^{(l+1)} = \sigma\left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} (\mathbf{H}^{(l)} \circ \boldsymbol{\Gamma}) \mathbf{W}^{(l)}\right), \quad (12)$$

其中, $\circ$ 表示矩阵元素项之间的乘法, 且 $\mathbf{H}^{(l)} = \sigma(\tilde{\mathbf{A}} \mathbf{X} \mathbf{W}^{(0)})$ 。最后一层 GCN 的输出为节点嵌入矩阵 $\mathbf{Z}^A$, 则重构后的邻接矩阵可表示为

$$\mathbf{A}' = \text{Sigmoid}(\mathbf{Z}^A (\mathbf{Z}^A)^T) \quad (13)$$

相应地, 邻接矩阵的重构损失定义为

$$\mathcal{L}_A = \|\mathbf{A} - \mathbf{A}'\|_F^2 \quad (14)$$

4.3 属性自编码器

ResDAE 在设计属性自编码器时, 使用了多层感知机模型将原始高维的属性矩阵 $\mathbf{X}$ 映射到低维的隐层表示 $\mathbf{Z}^X$ 中, 即

$$\begin{aligned} \tilde{\mathbf{Z}}^X &= \sigma\left(\mathbf{X}^T \mathbf{W}^{l(1)} + \mathbf{b}^{l(1)}\right), \\ \mathbf{Z}^X &= \tilde{\mathbf{Z}}^X \mathbf{W}^{l(2)} + \mathbf{b}^{l(2)} \end{aligned} \quad (15)$$

其中, $\mathbf{W}^{l(1)}$ 与 $\mathbf{W}^{l(2)}$ 是感知机模型中需要学习的权重参数, $\mathbf{b}^{l(1)}$ 与 $\mathbf{b}^{l(2)}$ 为偏置。此外, ResDAE 在重构属性矩阵时, 综合考虑了结构编码器生成的节点嵌入表示 $\mathbf{Z}^A$ 与属性编码器生成的属性嵌入表示 $\mathbf{Z}^X$, 能够有效地捕捉网络结构与节点属性之间跨模态的隐式关系, 即

$$\mathbf{X}' = \mathbf{Z}^A \cdot (\mathbf{Z}^X)^T \quad (16)$$

4.4 异常检测算法设计

基于前文内容, ResDAE 方法的总体损失函数被定义为

$$\mathcal{L}_{\text{rec}} = \alpha \|\mathbf{A} - \mathbf{A}'\|_F^2 + (1 - \alpha) \|\mathbf{X} - \mathbf{X}' - \delta \mathbf{R}\|_F^2 \quad (17)$$

其中, α 为权衡结构重构损失与属性重构损失占比的参数; δ 为残差参数, 用于控制属性重构误差中使用残差信息的量; $\odot$ 表示阿达玛积, 参数 θ 和 η 分别为

$$\theta = \begin{cases} 1, & \text{if } A_{i,j} = 0 \\ \theta, & \text{otherwise} \end{cases}, \eta = \begin{cases} 1, & \text{if } X_{i,j} = 0 \\ \eta, & \text{otherwise} \end{cases}, \quad (18)$$

并且 $\theta > 1$ 、 $\eta > 1$, 用于对由缺失的边或属性所产生的非零元素重构误差施加更多的惩罚。

随后, 节点 v_i 的异常分数计算公式为

$$\mathcal{S}_{(v_i)} = \alpha \|(\mathbf{A}_i - \mathbf{A}'_i) \odot \boldsymbol{\theta}_i\|_F^2 + (1 - \alpha) \|(\mathbf{X}_i - \mathbf{X}'_i - \delta \mathbf{R}) \odot \boldsymbol{\eta}_i\|_F^2 \quad (19)$$

综上所述, 算法 1 给出了节点异常综合评分算法伪代码。

算法 1. 节点异常综合评分算法.

输入: 邻接矩阵 $\mathbf{A}$ 、属性矩阵 $\mathbf{X}$ 、epochs

输出: 节点异常评分集合 $\mathcal{S}$

1: Initialize GCN model and MLP model, $t = 0$;

2: WHILE $t \leq \text{epochs}$ DO

3: Calculate residual matrix $\mathbf{R}$ by Eq.(9);

4: $\mathbf{Z}^A \leftarrow \text{GCN}(\mathbf{A}, \mathbf{W}_{\text{GCN}})$;

5: $\mathbf{Z}^X \leftarrow \text{MLP}(\mathbf{X}, \mathbf{W}_{\text{MLP}})$;

6: Update reconstruction loss $\mathcal{L}_{\text{rec}}$ by Eq. (17);

7: END WHILE

8: Calculate score matrix $\mathcal{S}$ by Eq. (19);

5 性能评价

5.1 实验设置

5.1.1 数据集描述

为了方便进行对比实验, 本文使用 BlogCatalog、Flickr 和 ACM 数据集对基于图的异常检测性能展开评估。3 个公开数据集介绍如下。

(1) BlogCatalog: BlogCatalog 是一个博客分享网站, 网站中的博主能够相互关注, 形成社交网络。用户属性由自身描述标签以及博客标签定义。

(2) Flickr: Flickr 是一个图片管理和分享网站, 与 BlogCatalog 类似, 用户可以相互关注, 形成社交网络。用户的节点属性由其相关的兴趣标签定义。

(3) ACM: ACM 是学术领域的一个引文属性网络, 每篇论文被视作网络中的一个节点, 不同论文之间的引用产生网络的边。节点的属性从论文摘要中生成。

由于上述数据集本身不存在异常, 参考文章[14]中的方法, 通过扰动拓扑结构和节点属性为每个数据集进行异常信息注入, 用于评价不同方法的性能差异。表 1 汇总了数据集的相关统计数据。

表 1 社会网络关系数据集描述

Table 1 Social network relationship dataset description

数据集	BlogCatalog	Flickr	ACM
节点数量	5196	7575	16484
边数量	171743	239738	71980
属性数量	8189	12074	8337
异常节点数量	300	450	600

5.1.2 实验环境

为了公平地开展对比实验, 所有对比方法都实现在如表 2 所示的实验环境中。

表 2 对比实验详细环境配置一览表

Table 2 Comparison experiment detailed environment configuration list

环境配置项	详细配置情况
Pycharm 版本	Pycharm 2023.2.4(Professional Edition)
GPU 配置	NVIDIA RTX A4000 16GB
操作系统	Ubuntu 18.04 LTS
Python 环境	Python 3.6
Tensorflow 版本	tensorflow 1.10.0

5.2 性能评价指标

本文使用了如下 2 个量化评价指标来比较不同方法的性能优劣:

(1) ROC-AUC。ROC 曲线(Receiver Operating characteristic Curve)是受试者工作特征曲线, 该曲线的横坐标为假阳率(False Positive Rate, FPR), 即识别为异常的正常节点比率, 纵坐标为真阳率(True Positive Rate, TPR), 即正确识别为异常的真正异常节点的比率。ROC-AUC(Area Under the Receiver Operating Characteristic Curve, 后文简称 AUC)取值在 0.5~1.0, 越接近 1.0 表明被检测方法的性能越高。

(2) Precision@K 和 Recall@K。Precision@K 指标被定义为在前 K 个高异常分数的节点中发现真实异常的比例。Recall@K 指标被定义为前 K 个高异常分数的节点中被检测出真实异常样本的比例。

5.3 对比实验结果与分析

首先, 本文探究嵌入维度对异常检测性能的影响情况, 根据经验将超参数 λ 设为 0.01, 将嵌入维度的值依次设为 8、16、32、64、128、256、512, 统计 AUC 指标得到的实验结果如图 2 所示。从图 2 可知, 当在 BlogCatalog 数据集实验中将嵌入维度设置为 128 和 256 时, AUC 指标均为最大值; 而在 Flickr 和 ACM 数据集实验中, 最佳嵌入维度均为 128。进一步地, 本文采取控制变量法, 将 BlogCatalog、Flickr 和 ACM 数据集实验里的嵌入维度都设置为 128 后, 继续比较不同超参数 λ 的取值对 AUC 指标的影响, 结果如图 3 所示。从图 3 可知, 针对基于 BlogCatalog 和 Flickr 数据集的实验, λ 的最优取值均为 0.05; 针对基于 ACM 数据集的实验, λ 的最优取值是 0.01。基于上述实验结果, 在基于 BlogCatalog、Flickr 和 ACM 数据集的实验中, ResDAE 方法的超参数 $(\alpha, \eta, \theta, \lambda)$ 分别设置为(0.7, 5, 40, 0.05)、(0.9, 8, 90,

0.05)和(0.7, 3, 10, 0.01), 嵌入维度均设置为 128。

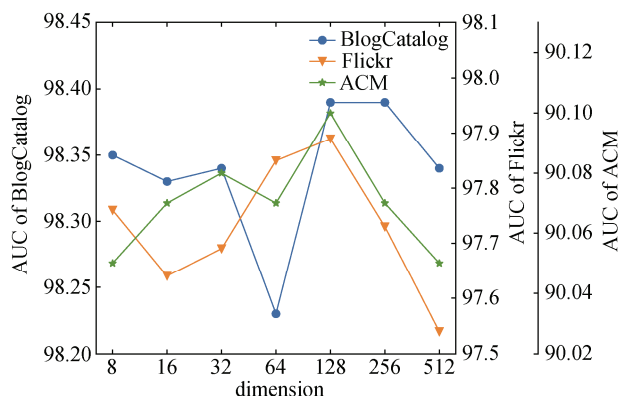


图2 ResDAE 方法在不同数据集上的 AUC 性能随嵌入维度变化情况

Figure 2 The AUC performance of ResDAE versus the embedding dimension over different datasets

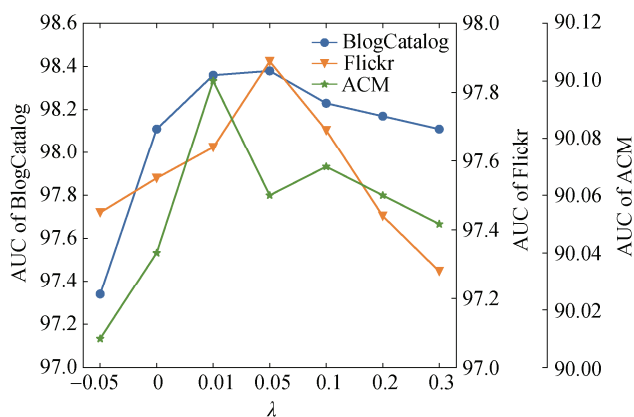


图3 ResDAE 方法在不同数据集上的 AUC 指标随超参数 λ 变化情况

Figure 3 The AUC performance of ResDAE versus the hyperparameter λ over different datasets

为了验证 ResDAE 性能优势, 本文选取了其他 5 种代表性的异常检测方法开展对比实验, 具体介绍如下。

(1) Radar^[25]: 该方法用于属性网络的无监督异常检测, 通过表征属性信息的残差和网络信息的一致性来检测与大多数样本行为不同的异常。

(2) Anomalous^[24]: 该方法是一个联合的异常检测框架, 使用矩阵 CUR 分解^[27]和属性网络上的残差分析来优化属性选择和异常检测。

(3) Dominant^[14]: 该方法利用 GCN 来生成输入属性网络的低维嵌入表示, 然后用这些表示重建拓扑结构和节点属性, 通过对重构误差进行排序来选择异常。

(4) AnomalyDAE^[16]: 该方法使用结构自编码器和属性自编码器学习属性网络的低维嵌入表示, 对

重构误差进行排序来选择异常。

(5) DAGAD^[28]: 该方法使用图神经网络编码器学习表示的信息融合, 使用生成的样本对训练集进行数据增强, 通过类平衡损失计算减轻类不平衡问题, 达到区分异常类和正常类分布目的。

表 3 展示了不同方法的 AUC 分数对比结果, 粗体的数据表示对比最佳的结果。从表 3 可知, 利用图神经网络来检测异常节点的方法, 如 Dominant、AnomalyDAE、DAGAD 和 ResDAE, 其 AUC 指标结果均高于仅基于残差来检测异常节点的方法, 如 Radar 和 Anomalous, 这一结果表明, 图神经网络能够更好地编码节点表示以及学习到深层的特征, 从而提升了异常节点检测性能。此外, ResDAE 方法在 3 个数据集上的 AUC 指标分数都高于其他 5 种比较方法的指标值, 这一结果也表明, 通过综合双通道图深度自编码器和残差注意力模块来提升异常检测能力的创新思路是有效且可行的。

表3 不同异常检测方法在不同数据集上的 AUC 性能对比

Table 3 Performance comparison of AUC with respect to different methods over different datasets

对比方法	BlogCatalog	Flickr	ACM
Radar	0.7142	0.7309	0.6963
Anomalous	0.7265	0.7213	0.7188
Dominant	0.7831	0.7500	0.7494
AnomalyDAE	0.9745	0.9680	0.8906
DAGAD	0.8364	0.8312	0.7935
ResDAE	0.9836	0.9789	0.9010

进一步地, 本文还探究了不同方法在 3 个数据集上检测异常节点的 Precision@K 指标性能。由表 1 可知, 数据集 BlogCatalog、Flickr 和 ACM 中包含的异常样本数量分别在 300、450 和 600 左右。因此, 针对上述 3 个数据集实验, 本文将 K 的取值分别设置为 {50, 100, 200, 300}、{100, 200, 300, 400} 和 {100, 250, 400, 550}, 以便于在不同异常样本数量的数据集上进行实验。不同方法在上述数据集上的 Precision@K 性能指标对比结果分别如表 4、表 5 和表 6 所示, 其中, 加粗的数据表示对比最佳的结果。

从表 4 和表 5 可以看到, 在 BlogCatalog 和 Flickr 数据集实验上, ResDAE 的 Precision@K 指标结果都高于其他对比方法。这一实验结果表明, 在这两个数据集实验中, 本文提出的方法相较于其他对比方法来讲具有更为突出的异常节点检测能力。值得注意的是, 表 6 结果显示, 在 ACM 数据集实验中, DAGAD

表 4 不同异常检测方法在 BlogCatalog 数据集上的 Precision@K 性能对比

Table 4 Performance comparison of Precision@K with respect to different methods over the BlogCatalog dataset

K	50	100	200	300
Radar	0.653	0.671	0.552	0.411
Anomalous	0.646	0.655	0.517	0.403
Dominant	0.720	0.710	0.590	0.466
AnomalyDAE	0.780	0.880	0.885	0.780
DAGAD	0.825	0.900	0.880	0.790
ResDAE	0.880	0.930	0.895	0.802

表 5 不同异常检测方法在 Flickr 数据集上的 Precision@K 性能对比

Table 5 Performance comparison of Precision@K with respect to different methods over the Flickr dataset

K	100	200	300	400
Radar	0.730	0.634	0.507	0.490
Anomalous	0.709	0.650	0.505	0.482
Dominant	0.730	0.685	0.593	0.510
AnomalyDAE	0.690	0.690	0.730	0.742
DAGAD	0.714	0.735	0.828	0.786
ResDAE	0.750	0.795	0.830	0.827

表 6 不同异常检测方法在 ACM 数据集上的 Precision@K 性能对比

Table 6 Performance comparison of Precision@K with respect to different methods over the ACM dataset

K	100	250	400	550
Radar	0.550	0.513	0.428	0.330
Anomalous	0.590	0.500	0.395	0.380
Dominant	0.600	0.532	0.445	0.353
AnomalyDAE	0.110	0.145	0.175	0.188
DAGAD	0.717	0.644	0.528	0.440
ResDAE	0.130	0.150	0.177	0.181

方法的 Precision@K 指标值最佳, 原因分析如下: DAGAD 综合了数据增强技术以及在损失函数计算过程中处理类别不平衡情况, 较好地解决了该数据集中异常节点占比很低所导致的检测性能差的问题。

进一步地, 本文在数据集 BlogCatalog、Flickr 和 ACM 实验中, 对比了不同异常检测方法的 Recall@K 性能表现, 如表 7~表 9 所示, 其中, 加粗的数据表示对比最佳的结果。

从表 7 和表 8 可知, 针对 BlogCatalog 和 Flickr 数据集实验, ResDAE 方法的 Recall@K 指标结果都高于其他对比方法。上述结果进一步验证了与前文类似的

结论, 即 ResDAE 在这两个数据集实验中性能表现最佳。从表 9 可知, 针对基于 ACM 数据集的实验, DAGAD 方法的性能表现最佳, 原因同前文分析。

表 7 不同异常检测方法在 BlogCatalog 数据集上的 Recall@K 性能对比

Table 7 Performance comparison of Recall@K with respect to different methods over the BlogCatalog dataset

K	50	100	200	300
Radar	0.105	0.196	0.348	0.407
Anomalous	0.100	0.190	0.352	0.410
Dominant	0.121	0.238	0.395	0.466
AnomalyDAE	0.130	0.295	0.593	0.792
DAGAD	0.136	0.305	0.598	0.700
ResDAE	0.147	0.312	0.601	0.802

表 8 不同异常检测方法在 Flickr 数据集上的 Recall@K 性能对比

Table 8 Performance comparison of Recall@K with respect to different methods over the Flickr dataset

K	100	200	300	400
Radar	0.143	0.278	0.339	0.365
Anomalous	0.148	0.290	0.360	0.415
Dominant	0.164	0.307	0.400	0.458
AnomalyDAE	0.155	0.310	0.492	0.667
DAGAD	0.159	0.342	0.533	0.723
ResDAE	0.168	0.357	0.557	0.739

表 9 不同异常检测方法在 ACM 数据集上的 Recall@K 性能对比

Table 9 Performance comparison of Recall@K with respect to different methods over the ACM dataset

K	100	250	400	550
Radar	0.086	0.173	0.225	0.280
Anomalous	0.091	0.177	0.233	0.286
Dominant	0.101	0.222	0.304	0.355
AnomalyDAE	0.018	0.053	0.123	0.157
DAGAD	0.522	0.558	0.630	0.705
ResDAE	0.118	0.127	0.132	0.167

综上所述, ResDAE 在 AUC 性能指标上具有明显的比较优势, 且在 BlogCatalog 和 Flickr 数据集实验中其 Recall@K 与 Precision@K 指标表现均最佳, 总体上验证了本文方法的性能优势。

5.4 基于图的异常检测方法在僵尸网络数据集上的对比结果

5.4.1 僵尸网络数据集描述

鉴于僵尸网络具有极大的危害性, 僵尸网络检

测是网络安全防御工作的一个重要方面。本文进一步通过实验来评价 ResDAE 方法针对僵尸网络的检测性能, 其中, 使用的数据集是僵尸网络基准数据集 CTU-13^[29]。CTU-13 数据集中包含 13 个不同的子数据集, 它们是从 7 个不同的恶意软件中捕获的流量数据, 其中包含了 DDoS 攻击、点击欺诈、垃圾邮件等多种恶意行为流量记录。每个子集中都包含一个独特的通信网络拓扑, 网络结构中包含一定数量的使用不同协议的僵尸主机(Bot)。CTU-13 中的各个特征记录着当前流是来自/流向僵尸主机/正常主机, 以及流量的其他基本特征, 如传输数据包数量、持续时长、连接类型、连接状态等。不失一般性, 本文选取了 CTU-13 数据集中序号为 5、7、11 的子数据集进行对比实验。

考虑到 CTU-13 数据集中僵尸网络流量仅占整个网络流量的 1.5%, 存在严重的类别不平衡情况, 本实验对少数类别样本进行过采样操作, 以克服类别不平衡带来的不利影响。过采样后得到的数据集信息如表 10 所示。

表 10 过采样后的 CTU-13 子数据集
Table 10 CTU-13 subdatasets after oversampling

数据集	CTU-13-5	CTU-13-7	CTU-13-11
节点数量	41658	38204	41931
边数量	130227	123555	168241
异常节点数量	1251	2830	9160

本实验通过网络流数据处理、网络通信图构建和图结构特征提取三个模块的处理^[9], 得到可用于图神经网络处理的图结构数据。随后, 使用不同的方法检测僵尸网络流量并对比方法性能优劣。

5.4.2 对比实验结果分析

表 11 给出了本文提出的 ResDAE 方法在不同

CTU-13 子数据集上的实验结果。从表 11 可知, 针对 CTU-13-5、CTU-13-7 和 CTU-13-11 子数据集, 在分别选择 K 参数值为 2500、3000 和 9000 后, 其 Precision@K 指标和 Recall@K 指标得分都可达 85% 以上, 从而验证了 ResDAE 方法检测僵尸网络的良好性能。

本文进一步基于 CTU-13 的 3 个子数据集开展对比实验, 图 4 展示了本文方法和 Dominant、AnomalyDAE、DAGAD 三种对比方法的 AUC 性能对比结果。从图 4 可知, ResDAE 的 AUC 指标分数最佳, 并显著地优于其他 3 种对比方法, 从而进一步验证了本文方法在僵尸网络检测任务下的优越性能。此外, 由 ResDAE 和 AnomalyDAE 的对比结果可知, 本文方法在不同子数据集上都取得了更高的性能指标结果, 再次证明了残差注意力机制能够带来良好的异常检测性能增益。

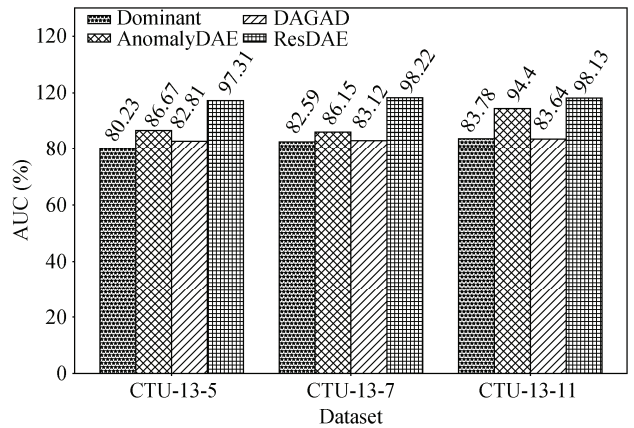


图 4 不同检测方法在 CTU-13 子数据集上的 AUC 性能对比

Figure 4 Performance comparison of different detection methods in terms of AUC over CTU-13 sub-datasets

表 11 ResDAE 在不同 CTU-13 子数据集上的实验结果

Table 11 Experimental results of ResDAE over different CTU-13 sub-datasets

Dataset	CTU-13-5				CTU-13-7				CTU-13-11			
K	1000	1500	2000	2500	2000	2500	3000	4000	3000	4500	6000	9000
Precision@K	0.233	0.466	0.865	0.971	0.682	0.767	0.863	0.938	0.968	0.955	0.960	0.875
Recall@K	0.1880	0.445	0.866	0.970	0.489	0.688	0.862	0.681	0.305	0.422	0.642	0.855

6 总结与下一步工作

本文扩展了 AnomalyDAE 方法中的双重自编码器框架, 提出了一种基于双重自编码器和残差注意力机制的异常检测方法 ResDAE, 它考虑了双重自编码器的重构过程中产生的残差信息, 并在结构自编

码器上引入基于残差分析的注意力模块, 从而有效地捕获了网络结构与节点属性之间的隐式关系; 另外, 该方法使用了图神经网络来学习节点的嵌入表示, 提升了图嵌入表示能力。基于多个公开数据集的对比实验结果表明, ResDAE 方法对比其他主流的图异常检测方法具有更好的性能表现。针对僵尸网络

检测任务, 基于 CTU-13 数据集的对比实验结果也验证了残差注意力机制所带来的检测性能增益。

下一步, 本文将着力解决类别不平衡数据场景下的网络异常检测问题, 以期进一步增强 ResDAE 方法的实用性和对不同应用场景的适用性。

参考文献

- [1] Yuan B, Zou D Q, Jin H. Network Security Visualization: A Survey[J]. *Journal of Cyber Security*, 2016, 1(3): 10-20.
(袁斌, 邹德清, 金海. 网络安全可视化综述[J]. *信息安全学报*, 2016, 1(3): 10-20.)
- [2] Liu R X, Chen H, Guo R Y, et al. Survey on Privacy Attacks and Defenses in Machine Learning[J]. *Journal of Software*, 2020, 31(3): 866-892.
(刘睿琨, 陈红, 郭若杨, 等. 机器学习中的隐私攻击与防御[J]. *软件学报*, 2020, 31(3): 866-892.)
- [3] Zhou M Q, Huang H, Wang Q. A Graph-Based Clustering Algorithm for Anomaly Intrusion Detection[C]. *2012 7th International Conference on Computer Science & Education*, 2012: 1311-1314.
- [4] Ngo Q D, Nguyen H T, Nguyen V D, et al. Adversarial Attack and Defense on Graph-Based IoT Botnet Detection Approach[C]. *2021 International Conference on Electrical, Communication, and Computer Engineering*, 2021: 1-6.
- [5] Li M, Li C F, Li S L, et al. ACGVD: Vulnerability Detection Based on Comprehensive Graph via Graph Neural Network with Attention[C]. *Information and Communications Security*, 2021: 243-259.
- [6] Luo B, Xia J B. A Novel Intrusion Detection System Based on Feature Generation with Visualization Strategy[J]. *Expert Systems with Applications*, 2014, 41(9): 4139-4147.
- [7] Yu X, Tian Z H, Qiu J, et al. An Intrusion Detection Algorithm Based on Feature Graph[J]. *Computers, Materials & Continua*, 2019, 61(1): 255-274.
- [8] Chowdhury S, Khanzadeh M, Akula R, et al. Botnet Detection Using Graph-Based Feature Clustering[J]. *Journal of Big Data*, 2017, 4(1): 14.
- [9] Abou Daya A, Salahuddin M A, Limam N, et al. BotChase: Graph-Based Bot Detection Using Machine Learning[J]. *IEEE Transactions on Network and Service Management*, 2020, 17(1): 15-29.
- [10] Wang W, Shang Y Y, He Y Z, et al. BotMark: Automated Botnet Detection with Hybrid Analysis of Flow-Based and Graph-Based Traffic Behaviors[J]. *Information Sciences*, 2020, 511: 284-296.
- [11] Alharbi A, Alsubhi K. Botnet Detection Approach Using Graph-Based Machine Learning[C]. *IEEE Access*, 2021: 99166-99180.
- [12] Qi M Y, Liu M, Fu Y M. Research on Network Intrusion Detection Using Support Vector Machines Based on Principal Component Analysis[J]. *Netinfo Security*, 2015, 15(2): 15-18.
(戚名钰, 刘铭, 傅彦铭. 基于 PCA 的 SVM 网络入侵检测研究[J]. *信息网络安全*, 2015, 15(2): 15-18.)
- [13] Guo X D, Li X M, Jing R X, et al. Intrusion Detection Based on Improved Sparse Denoising Autoencoder[J]. *Journal of Computer Applications*, 2019, 39(3): 769-773.
(郭旭东, 李小敏, 敬如雪, 等. 基于改进的稀疏去噪自编码器的入侵检测[J]. *计算机应用*, 2019, 39(3): 769-773.)
- [14] Ding K Z, Li J D, Bhanushali R, et al. Deep Anomaly Detection on Attributed Networks[M]. *Proceedings of the 2019 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019: 594-602.
- [15] Zhu D L, Ma Y C, Liu Y L. Anomaly Detection with Deep Graph Autoencoders on Attributed Networks[C]. *2020 IEEE Symposium on Computers and Communications*, 2020: 1-6.
- [16] Fan H Y, Zhang F B, Li Z Y. Anomalydae: Dual Autoencoder for Anomaly Detection on Attributed Networks[C]. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020: 5685-5689.
- [17] Li J Y, Zhao B, Li X, et al. Network Traffic Anomaly Prediction Method Based on Deep Learning[J]. *Computer Engineering and Applications*, 2020, 56(6): 39-50.
(黎佳玥, 赵波, 李想, 等. 基于深度学习的网络流量异常预测方法[J]. *计算机工程与应用*, 2020, 56(6): 39-50.)
- [18] Zhu T, Xia L L, Li P H, et al. Analysis of Botnet Attack Data Based on Log[J]. *Netinfo Security*, 2022, 22(10): 82-90.
(朱涛, 夏玲玲, 李鹏辉, 等. 基于日志的僵尸网络攻击数据分析[J]. *信息网络安全*, 2022, 22(10): 82-90.)
- [19] Xue H T, Wang C. Research on Botnet Traffic Detection Based on GAN[J]. *Electronic Design Engineering*, 2022, 30(17): 146-149.
(薛昊天, 王晨. 基于 GAN 的僵尸网络流量检测研究[J]. *电子设计工程*, 2022, 30(17): 146-149.)
- [20] Zhao D, Traore I, Sayed B, et al. Botnet Detection Based on Traffic Behavior Analysis and Flow Intervals[J]. *Computers & Security*, 2013, 39: 2-16.
- [21] Perozzi B, Akoglu L. Scalable Anomaly Ranking of Attributed Neighborhoods[C]. *The 2016 SIAM International Conference on Data Mining*, 2016: 207-215.
- [22] Gao J, Liang F, Fan W, et al. On Community Outliers and Their Efficient Detection in Information Networks[C]. *The 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2010: 813-822.
- [23] Perozzi B, Akoglu L, Iglesias Sánchez P, et al. Focused Clustering and Outlier Detection in Large Attributed Graphs[C]. *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014: 1346-1355.
- [24] Peng Z, Luo M N, Li J D, et al. ANOMALOUS: A Joint Modeling Approach for Anomaly Detection on Attributed Networks[C]. *The Twenty-Seventh International Joint Conference on Artificial Intelligence*, 2018: 3513-3519.
- [25] Li J D, Dani H, Hu X, et al. Radar: Residual Analysis for Anomaly Detection in Attributed Networks[C]. *The Twenty-Sixth International Joint Conference on Artificial Intelligence*, 2017: 2152-2158.
- [26] Pei Y L, Huang T J, van Ipenburg W, et al. ResGCN: Attention-Based Deep Residual Modeling for Anomaly Detection on Attributed Networks[J]. *Machine Learning*, 2022, 111(2): 519-541.
- [27] Mahoney M W, Drineas P. CUR Matrix Decompositions for Improved Data Analysis[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(3): 697-702.

[28] Liu F Z, Ma X X, Wu J, et al. DAGAD: Data Augmentation for Graph Anomaly Detection[C]. *2022 IEEE International Conference on Data Mining*, 2022: 259-268.



伍晓洁 于 2022 年在华中科技大学计算机科学与技术专业获得学士学位。现在国防科技大学计算机科学与技术专业攻读硕士学位。研究领域为图对抗学习、图深度学习。研究兴趣包括: 图对抗攻击、后门攻击。Email: wuxiaojie1015@nudt.edu.cn

[29] García S, Grill M, Stiborek J, et al. An Empirical Comparison of Botnet Detection Methods[J]. *Computers & Security*, 2014, 45: 100-123.



刘强 于 2014 年在国防科技大学计算机科学与技术专业获得博士学位。现任国防科技大学计算机学院副教授, 硕士生导师。中国计算机学会 CCF 会员。研究领域为网络安全、人工智能安全。研究兴趣包括: 网络安全、深度学习对抗样本、图学习安全。Email: qiangliu06@nudt.edu.cn