软件学报 ISSN 1000-9825, CODEN RUXUEW
2026,37(4):1560−1574 [doi: 10.13328/j.cnki.jos.007527] [CSTR: 32375.14.jos.007527]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

全局与局部残差信息联合感知的可泛化图异常检测*

张家强[1,2], 陈松灿[1,2]

[1](南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106)

[2](模式分析与机器智能工业和信息化部重点实验室 (南京航空航天大学), 江苏 南京 211106)

通信作者: 陈松灿, E-mail: s.chen@nuaa.edu.cn

摘　要: 图异常检测作为图数据挖掘中的关键任务, 旨在识别网络中与大多数节点存在显著差异的异常节点. 现有的图异常检测方法普遍采用数据集特定的训练范式, 即为每个数据集单独训练模型. 然而, 该类方法缺乏跨数据集的泛化能力, 且训练成本高昂. 为克服上述局限, 近期研究开始关注残差特征的泛化潜力. 该类特征通过计算节点自身表示与基于邻居传播后的表示之差, 能够在很大程度上抵消特定于数据集的语义信息, 从而保留与异常模式紧密相关的通用性信息. 尽管该方向已取得初步成果, 但残差特征的建模过程仍存在如下关键问题: 首先, 在计算节点基于邻居传播前后的表示差值时, 邻居节点的稀少和潜在的结构噪声会在一定程度上影响结果的可靠性. 其次, 计算时的表示依赖于图神经网络 (graph neural network, GNN) 对局部关系的学习, 这种方式难以建模对异常检测同样有益的全局关系, 从而限制了残差特征的表达能力. 为解决上述问题, 提出一种全局和局部残差信息联合感知的可泛化图异常检测方法 GRAD. 具体地, 该方法在利用 GNN 建模局部节点关系的基础上, 引入线性 Transformer 模块, 在不依赖原始图结构的前提下, 于特征空间中建模节点之间的全局结构相关性, 从而获得具备全局感知能力的节点表示. 随后, GRAD 在全局和局部视角上分别将表示转换为自身与其邻居之间的残差, 并将二者融合, 以构建数据集无关的通用节点表示. 随后在多个不同领域的公开图数据集上进行广泛实验, 验证了 GRAD 的有效性.

关键词: 图异常检测; 图神经网络; 特征表示; 泛化性; 残差特征

中图法分类号: TP18

中文引用格式: 张家强, 陈松灿. 全局与局部残差信息联合感知的可泛化图异常检测. 软件学报, 2026, 37(4): 1560–1574. http://www.jos.org.cn/1000-9825/7527.htm

英文引用格式: Zhang JQ, Chen SC. Generalizable Graph Anomaly Detection via Joint Perception of Global and Local Residual Information. Ruan Jian Xue Bao/Journal of Software, 2026, 37(4): 1560–1574 (in Chinese). http://www.jos.org.cn/1000-9825/7527.htm

Generalizable Graph Anomaly Detection via Joint Perception of Global and Local Residual Information

ZHANG Jia-Qiang[1,2], CHEN Song-Can[1,2]

[1](College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

[2](MIIT Key Laboratory of Pattern Analysis and Machine Intelligence (Nanjing University of Aeronautics and Astronautics), Nanjing 211106, China)

Abstract: Graph anomaly detection, as a critical task in graph data mining, aims to identify anomalous nodes that significantly differ from the majority of nodes in a network. Existing methods for graph anomaly detection typically adopt dataset-specific training paradigms, i.e., training a separate model for each dataset. However, such methods lack generalization capability across datasets and incur high training costs. To overcome these limitations, recent studies have begun to focus on the generalization potential of residual features. Such features are obtained by computing the difference between a node's own representation and the representation after neighborhood propagation,

* 基金项目: 国家自然科学基金 (62376126); 航空发动机及燃气轮机重大专项 (J2019-IV-0018-0086)
本文由"高可信自适应机器学习"专题特约编辑王熙照教授、张敏灵教授、曹付元教授推荐.
收稿时间: 2025-05-12; 修改时间: 2025-06-30, 2025-08-15; 采用时间: 2025-08-20; jos 在线出版时间: 2025-09-02
CNKI 网络首发时间: 2026-01-27

which can largely offset dataset-specific semantic information and thus retain general information closely related to anomalous patterns. Despite initial progress in this direction, the modeling of residual features still faces the following key challenges: First, when computing the difference between a node's representations before and after neighborhood propagation, the sparsity of neighbors and potential structural noise affect the reliability of the results to some extent. Second, the computation of representations relies on graph neural network (GNN) to learn local relationships, which makes it difficult to model global relationships that are also beneficial for anomaly detection, thus limiting the expressive power of residual features. To address these issues, this study proposes GRAD, a generalizable graph anomaly detection method via joint perception of global and local residual information. Specifically, based on GNN for modeling local node relationships, GRAD introduces a linear Transformer module to model global structural correlations among nodes in the feature space without relying on the original graph structure, thus obtaining node representations with global awareness. Then, GRAD transforms the representations into residuals between each node and its neighbors from both global and local perspectives, and integrates them to form dataset-independent general node representations. Extensive experiments on multiple public graph datasets from different domains verify the effectiveness of GRAD.

Key words: graph anomaly detection; graph neural network (GNN); feature representation; generalization; residual feature

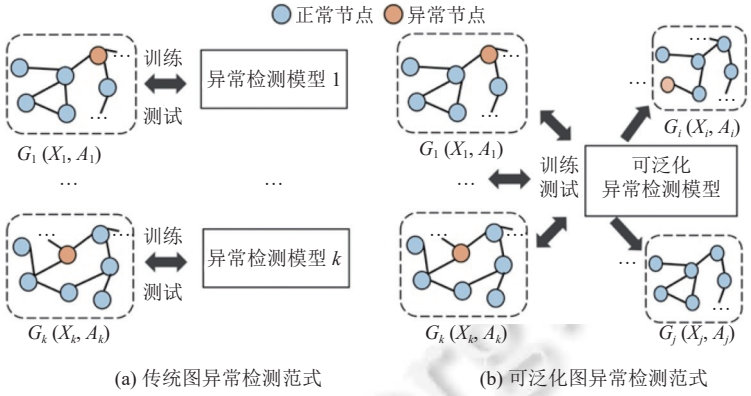
图在现实世界中无处不在, 可用来建模节点之间的复杂关系[1]. 图异常检测是基于图数据的一类关键任务, 旨在识别网络结构中显著偏离多数节点行为的异常节点. 该任务在各种场景中具有广泛的应用, 如社交垃圾邮件检测、金融欺诈检测等[2,3]. 早期的异常检测技术依赖于领域专家构建的特征工程, 如矩阵分解[4]和 OC-SVM[5]. 近期, 随着深度学习技术的飞速发展, 特别是图神经网络 (graph neural network, GNN)[6] 的广泛应用, 图学习任务 (如链路预测[7]、节点分类[8]等) 取得了显著进展. 受此启发, 研究者开始将 GNN 引入图异常检测任务, 以学习异常相关的节点表示. 现有基于 GNN 的方法可大致分为两类[9]: 一类将图异常检测看作是二分类问题, 通过调整神经网络架构或训练策略等, 在标签信息的引导下学习出可区分正常与异常节点的表示. 另一类方法不依赖于异常标签, 通过结合半监督或无监督学习范式, 设计代理学习任务 (如特征重构、邻居预测等) 以模拟潜在的异常模式, 进而使得模型具有异常检测的能力.

尽管这些方法展现了出色的检测性能, 如图 1(a) 所示, 它们都遵循在同一图上进行训练和推理的范式, 即需针对每一个图数据集单独训练一个模型. 这种做法存在显著的实际限制[10,11]: 其一, 随着数据的不断涌现, 基于每个数据集重新训练模型会带来较大的资源消耗; 其二, 在如医疗、金融等对隐私保护要求极高的应用场景中, 往往难以获取目标图的全部数据进行训练. 为了突破这一局限性, 研究者提出了一种新的范式, 即设计一个可泛化的异常检测模型[12]. 如图 1(b) 所示, 在此模型中, 可接纳不同的数据集进行训练. 同时, 在进行对目标数据集的检测时, 无需任何重新训练或微调, 仅需要极少量的正常样本进行目标知识获取. 该范式可显著降低计算成本, 具备更强的泛化能力, 但随之而来的是非凡的挑战, 即, 不同数据集之间的正常和异常模式存在差异. 因此, 该范式的关键之处在于当特征与结构变化时, 建模出节点中与数据集无关且与异常相关的表示. 受以往研究启发, 即大多数异常节点与其邻居的特征相似性较低, 近期的工作, 如 ARC[10]、AnomalyGFM[11]提出了残差特征这一概念, 即节点基于邻居传播前后的表示差值. 与显著变化的初始特征相比, 在通过减法得到的残差特征中, 与数据集相关的成分在一定程度上被相互消除, 保留了异常模式相关的通用信息.

但是这样的做法也存在着以下问题. 首先, 残差特征的计算依赖于基于节点邻居的传播, 然而现实网络中节点的度通常服从幂律分布[13]. 如图 2 所示, 我们绘制了常见的 3 个数据集中节点度的分布情况. 可以发现, 相当比例的节点具有极少的邻居. 同时, 在现实场景中, 会遇到孤立节点或噪声边的情况. 因此, 在这些情况下获取的残差特征会影响检测性能. 其次, 这些方法广泛依赖于 GNN 进行初步的节点表示学习, 但 GNN 本质上基于局部信息聚合机制, 难以建模全局上下文信息. 而已有研究表明, 全局信息对于识别异常节点具有重要价值[14,15]. 虽然通过堆叠多层的图神经网络 (GNN) 可以扩展感受野以捕获长距离依赖信息, 但随之而来的是节点表示过平滑的副作用[1]. GNN 作为一种低通滤波器, 多层的消息传递也会逐渐将异常信号磨灭[16].

为了解决上述问题, 本文提出了一种全局和局部残差信息联合感知的可泛化图异常检测方法 GRAD. 该方法无需在目标图数据上进行重新训练或微调, 便可实现对数据中异常节点的有效检测. 特别地, GRAD 首先在 GNN 进行局部信息感知的基础上, 引入线性 Transformer, 以高效建模节点在特征空间中的全局表示. 通过与原始特征相减, 得到全局残差表征. 这一过程本质上也可视为是基于节点特征的图结构学习, 有效规避了对真实结构信息的

依赖[17,18]. 继而, 将局部和全局残差特征进行融合, 在极少量正常样本的基础上进行上下文重建学习. 其中的重建距离可作为在测试过程中每个节点的异常分数. 最后, 本文在包含不同领域的多个基线数据集上进行了实验, 实验结果验证了所提方法的有效性.



(a) 传统图异常检测范式 (b) 可泛化图异常检测范式

图 1 传统与可泛化图异常检测范式的示意图

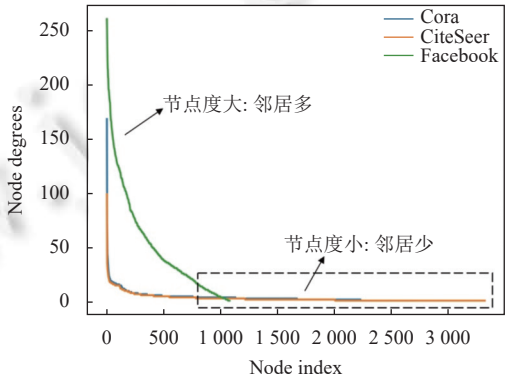


图 2 3 个数据集上的节点度分布示意图

综上所述, 本文的主要贡献如下.

(1) 研究一个可泛化图异常检测模型, 即模型可以检测来自不同应用领域的各种图数据集的异常, 仅需极少量正常样本进行知识获取, 而无需对目标数据进行任何训练或微调.

(2) 提出一种基于全局残差信息协作的可泛化图异常检测方法 (GRAD). 在该方法中, 巧妙地利用了线性 Transformer 高效地对全局信息进行建模, 并将其作为局部残差信息的补充进行上下文的重建学习和异常分数的计算.

(3) 在包含不同领域的多个基线数据集上测试所提出的 GRAD 方法. 实验结果表明, GRAD 方法有效地实现了多个目标数据集的异常检测. 与 SOTA 方法相比, GRAD 在 8 个目标数据集上的平均 AUROC 和平均 AUPRC 的性能分别提升了 1.9% 和 5.4%.

本文第 1 节回顾图异常检测的相关工作. 第 2 节介绍本文所研究的可泛化的图异常检测问题定义. 第 3 节介绍 GRAD 方法的架构设计与实现机制. 第 4 节通过实证分析验证方法的有效性. 第 5 节总结全文并展望未来研究方向.

1 相关工作

1.1 图神经网络

图神经网络 (GNN) 是近年来机器学习领域的重要研究方向, 旨在处理非欧几里得空间中的图结构数据[19,20],

在推荐系统[21]、分子预测[22]和交通流量预测[23]等任务上展现了优异的性能. 现有的 GNN 通常可被分类为两种类型: 基于空域的和基于谱域的. 前者基于图结构进行消息传播, 后者则受启发于信号处理中的频域操作. 近年来, 已有许多 GNN 模型及其变体被提出. 其中, Kipf 等人[6]提出的图卷积网络 (graph convolutional network, GCN) 通过局部一阶近似简化图卷积操作, 显著提升了模型效率, 成为里程碑式工作. 随后, Hamilton 等人[24]的 GraphSAGE 引入采样技术, 解决了大规模图训练的扩展性问题. 图注意力网络 (graph attention network, GAT)[25]通过注意力机制动态学习邻居权重, 进一步增强了模型表达能力. Wu 等人[26] 通过连续去除非线性和折叠连续层之间的权重矩阵来降低信息传递过程中的计算复杂性. Bo 等人[27]将数据解耦成高频和低频信号, 通过设计巧妙的自适应机制来动态汇集特征. 为提高图神经网络在分布漂移场景下的能力, 近期也涌现了一些相关工作. 例如, Yuan 等人[28]提出了一个用于节点级图 OOD (out-of-distribution) 泛化的结构感知不变学习框架 SING. 在其中, SING 设计了结构嵌入对齐损失来优化数据增广和由 GNN 建模得到的不变图表示. Xu 等人[29]提出一个通过渐进推理学习图因果不变性的模型 GPro. 具体而言, GPro 将图因果不变学习分解为多个由易到难的步骤, 并通过渐进推理过程不断强化 GPro 的感知能力, 从而提取对分布变化稳定的因果特征. 但这些方法聚焦在局部结构关系的建模上, 在图 OOD 的场景下进行全局关系的建模也是可改进的方向之一.

在图异常检测任务上应用 GNN 的目的是获得异常相关的节点表征, 与一般的节点/图分类数据集不同, 异常数据集通常具有较强的类不平衡性或异配性, 这使得 GNN 无法直接应用于异常数据集中. 在最近的研究中, 已经针对信息汇聚过程、滤波器设计等, 提出了多种 GNN 来处理问题[2], 以实现更有效的图异常检测. 比如, 在 GraphConsis 中, Liu 等人[30]揭示了节点连接中的不一致现象, 即异常节点很可能连接到正常节点以掩盖其异常. 然后, 它引入了一种基于一致性评分的方法, 该方法基于节点嵌入相似性和自关注机制, 为聚合中的不同连接分配权重. Tang 等人[16]基于图特征信号在谱域中的现象, 即异常的存在会导致光谱能量分布在低频的集中度降低, 在高频的集中度增加, 进而提出了 Beta 小波图神经网络, 其独特设计在于谱域和空域上的局部化带通滤波器.

1.2 图异常检测

早期的图异常检测方法主要依赖人工特征工程或专家开发的统计模型进行异常检测. 比如, Breunig 等人[31]利用局部离群因子 (local outlier factor, LOF), 通过密度估计识别异常. Perozzi 等人[32]基于邻域变化检测构建正态性度量, 该指标可以同时利用结构和属性来量化一致性和可分离性, 进而检测异常. 随着图神经网络在图数据相关任务, 如节点分类、链路预测等上的成功, 许多图异常检测研究利用 GNN 来建模节点表征. 在有监督信息的引导下, 该类方法在二分类的学习范式下展开[33]. 例如, Gao 等人[34]设计了一个标签 (预测) 感知的指标来计算聚合后的相似性得分, 并在此基础上修剪可能存在异质性的边, 从而进行异常检测.

在监督信息缺乏的情况下, 研究则集中在通过利用多种无监督学习技术来挖掘数据中的异常模式, 其中包括对比学习、重构学习等[35−37]. 在其中, Liu 等人[38]提出了一个基于子图对比学习的经典框架 CoLA, 其通过节点与其邻居间的对比实例对采样, 结合异常相关的目标来训练对比学习模型. TAM[39]发现了含注入异常与真实异常的 GAD 数据集中存在的“单类同质性”现象, 据此提出基于局部节点亲和力的新型无监督检测指标. 近期, Qiao 等人[40]基于先验信息, 设计多个正则化方法来生成伪异常数据来辅助模型学习, 展现了较好的性能. 吴江豪等人[41]通过利用结构熵和层次社区结构对异常节点进行检测. 然而, 这些方法关注在提升单个数据集上的性能, 当数据集不可训练或有新数据集到来之时, 所学得的模型的泛化性较差. 因此, 一个数据集一个模型的学习范式已不能满足信息快速涌现的现实场景, 亟需研究新型异常检测方法, 以期无需重新训练和微调, 模型即可在测试阶段对新数据进行检测.

2 问题定义与准备工作

2.1 问题定义

对于每个图, 可被表示为: $G=(V,E,\boldsymbol{X})$. 其中 V 是节点的集合: $V=\{v_1,\ldots,v_n\}$, E 是边的集合, n 为节点数量.

$\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n]^{\mathrm{T}}\in\mathbb{R}^{n\times d}$ 是图的特征矩阵, 图上节点之间的连接关系也可被表示为邻接矩阵 $\boldsymbol{A}\in\mathbb{R}^{N\times N}$, 其中若节点 v_i 和 v_j 之间存在边的连接, 则 $A_{ij}=1$.

本文关注的是节点级别的异常检测. 因此, 图上节点可以被划分为正常节点集 V_n 和异常节点集 V_a, 其中的标签信息分别被定义为 $y_i=0$ 和 $y_j=1$, $V_n\cup V_a=V$, $V_n\cap V_a=\varnothing$. 一般来说, 在每个图数据上, $|V_a|\ll|V_n|$. 图异常检测的目标是学习一个异常分数函数 (也可被称为异常检测模型) $f:V\to\mathbb{R}$, 可得到关系: $f(v)>f(v')$, 其中 $\forall v\in V_a$, $\forall v'\in V_n$.

在常规的设置中, 通常需要对每个数据集训练出一个模型进行异常检测. 在这里, 我们面向更具有现实意义的设置, 开发一个通用的图异常检测模型, 可检测来自不同领域或不同图数据集上的异常节点, 仅需极少量正常样本, 无需对目标数据集进行任何新的训练. 遵循 Liu 等人[7]近期的研究工作, 该设置被称为可泛化的图异常检测问题. 形式上的, 令 $T_{\text{train}}=\left\{D^1_{\text{train}},\ldots,D^N_{\text{train}}\right\}$ 为训练数据集的集合, 其中每个来自任意域的标记数据集可被定义为 $D^i_{\text{train}}=(G^i_{\text{train}},y^i_{\text{train}})$. 在这里, 我们的目标是在 T_{train} 的基础上学习一个可泛化的图异常检测模型 $f(\cdot)$, 其中 $f(\cdot)$ 可用来识别任何测试数据集 $T_{\text{test}}=\{D^1_{\text{test}},\ldots,D^{N'}_{\text{test}}\}$ 上的异常节点. 在测试阶段, 遵循正常少样本推断的范式, 即对于每个目标数据集, 提供 n_k ($n_k\ll n$) 个正常样本作为上下文以获得相关知识.

2.2 特征统一化

本文的目的是设计一个通用的网络来进行跨数据集的图异常检测, 所以网络的输入维度需要统一. 另外, 鉴于图中节点特征的维度存在显著差异, 如 Cora 为 1 433, CiteSeer 为 3 327, 为消除图间特征异构性, 需将节点特征映射至共享特征空间. 参考已有研究成果, 本工作采用奇异值分解 (SVD) 这一有效的降维方法[42]. SVD 可将高维特征投影至统一的低维子空间, 并且能够较好地保持原始数据的距离关系. 对于任意一个数据集的特征矩阵: $\boldsymbol{X}\in\mathbb{R}^{n\times d}$, 操作过程可定义如下:

$$\boldsymbol{X}\in\mathbb{R}^{n\times d}\xrightarrow{\text{SVD}}\tilde{\boldsymbol{X}}\in\mathbb{R}^{n\times d'} \tag{1}$$

其中, d' 是所有数据集共享的预定义转换后的特征维度. 这里的降维方法不限于 SVD, 也可基于主成分分析 (PCA)[43]、典型相关性分析 (CCA)[44]等进行转换.

同时, 虽然特征投影能够统一维度, 但不同数据集中投影特征的语义含义仍存在差异, 即不同位置的特征对应着不同的含义. 因此, 借鉴 ARC[10]中提出的方法, 我们基于特征对异常检测任务的实际贡献进行对齐. 形式上, 给定特征矩阵 $\tilde{\boldsymbol{X}}$, 特征贡献度可定义如下:

$$c_k\left(\tilde{\boldsymbol{X}}\right)=-\frac{1}{|E|}\sum\nolimits_{(v_i,v_j)\in\mathcal{E}}\left(\tilde{\boldsymbol{X}}_{ik}-\tilde{\boldsymbol{X}}_{jk}\right)^2 \tag{2}$$

其中, 较低的 c_k 表示连接节点之间的第 k 个特征发生了显著变化, 对应于高频图信号. 基于此, 我们可以将投影后的特征按照贡献度进行重排序为 $\boldsymbol{X}'$, 从而进一步对齐每个数据集的特征空间.

3 基于全局残差信息协作的可泛化图异常检测方法 GRAD

本节针对可泛化图异常检测问题中的两个关键难点: 1) 学习有效的数据集无关的通用表示, 2) 局部与全局信息的高效融合, 提出了一种全局和局部残差信息联合感知的框架 GRAD. 具体来说, 如图 3 所示, 为了缓解在获取局部残差特征时结构信息不足或不准确造成的影响, 我们首先基于节点特征, 利用线性 Transformer 进行高效的结构学习, 并结合隐藏层表征得到全局残差信息, 该操作也可视作对局部和全局信息进行解耦学习. 在这里, 线性 Transformer 可建模节点之间的全局相关性, 以补充仅利用原始结构不能捕获的远程信息. 同时, 也可摆脱传统注意力机制计算过程中的高计算复杂度的困境. 接着, 我们将局部与全局残差信息按列拼接作为每个节点的通用性表示, 最后, 在少量正常节点的引导下, 进行基于注意力机制的重建学习. 在测试过程, 重建距离可作为异常分数, 即重建距离大的节点有更大的可能是异常节点. 接下来, 本文将分别介绍 GRAD 方法中的 3 个关键技术.

3.1 基于局部残差信息的表示学习

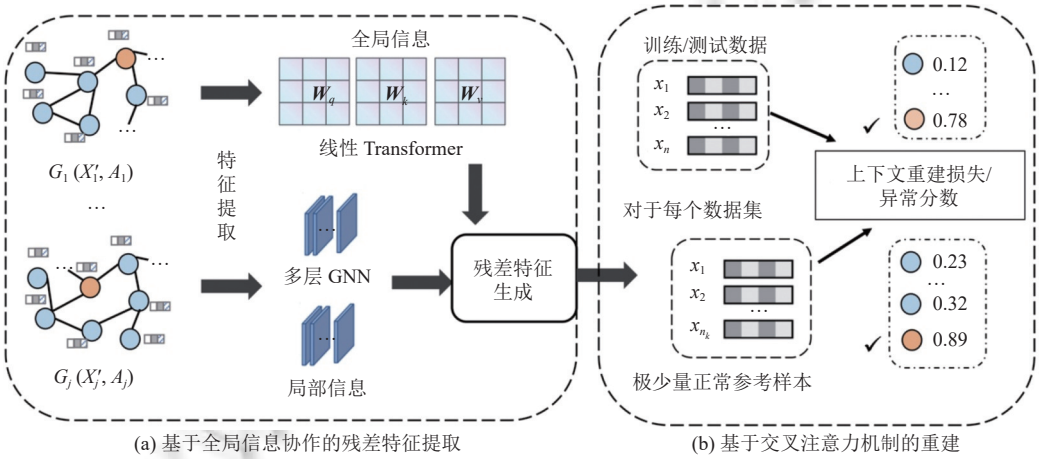
在常规的图异常检测问题下, 我们可采用基于 GNN 的图编码器来学习融合语义与结构信息的节点嵌入, 该嵌入将作为下游异常评分模块的输入. 但对于多个数据集, 正常与异常的模式之间是有差异的, 直接利用节点嵌入

可能导致模型过度拟合数据集特定的语义信息, 却无法识别跨数据集的共享异常特征, 从而导致泛化性差. 因此, 在本文中, 我们利用节点表示的残差进行学习, 即节点基于邻居传播前后的表示差值. 通过差分运算的操作, 可以有效抑制原始表示中与数据集相关的成分, 从而增强异常相关信号的相对强度. 在这种表示空间中, 残差幅值直接反映节点的异常程度, 较大的残差特征对应更高的异常概率. 在操作上, 首先, 我们选取先传播后转换的架构对节点进行表示学习. 对于一个图, 假设对齐后的初始节点特征矩阵为: $\boldsymbol{X}^0=\boldsymbol{X}'$, 则第 l 次传播时的特征可被形式化为:

$$\boldsymbol{X}^l=\tilde{\boldsymbol{A}}\boldsymbol{X}^{l-1} \tag{3}$$

其中, $\tilde{\boldsymbol{A}}$ 是归一化的邻接矩阵. 经过 l 次传播后, 可得到 l 个特征矩阵: $\{\boldsymbol{X}^0,\ldots,\boldsymbol{X}^l\}$. 接着, 我们借助多层感知器 $g(\cdot)$ 将特征映射到高维空间:

$$\boldsymbol{Z}^l=g\left(\boldsymbol{X}^l\right)\in\mathbb{R}^{n\times d_r} \tag{4}$$



(a) 基于全局信息协作的残差特征提取 (b) 基于交叉注意力机制的重建

图 3 GRAD 的整体流程图

在此基础上, 我们可以计算残差特征, 当 $l>0$ 时, 对于第 l 层的表示被定义为:

$$\boldsymbol{R}^l=\boldsymbol{Z}^l-\boldsymbol{Z}^0 \tag{5}$$

接着, 将 l 个残差特征表示按列拼接, 得到局部视角的表示:

$$\boldsymbol{H}_{\text{local}}=\left[\boldsymbol{R}^1\|\ldots\|\boldsymbol{R}^l\right] \tag{6}$$

3.2 基于全局残差信息的表示学习

在获得局部视角下的残差特征信息后, 若直接应用于异常检测, 会存在以下两个关键问题: (1) 邻域依赖性问题: 残差计算对邻居表示具有强依赖性, 但实际网络的幂律分布特性 (图 2) 导致部分节点的邻域信息极度稀疏, 加之复杂系统中难以避免的噪声边的干扰, 直接影响残差特征的有效性; (2) 表示学习局限性: 基于 GNN 的框架受制于局部消息传递机制, 缺乏对全局异常模式的感知能力. 虽然深层架构可以扩展信息传播范围, 但会引发表征同质化 (over-smoothing) 的负面效应. 因此, 本文拟基于特征进行全局性的结构学习和残差表示学习. 在以往的研究中验证了 Transformer 中的全局注意力机制能够捕捉节点间隐含的依赖关系, 然而, 平凡的 Transformer 模型的计算复杂度通常随节点数量呈二次方增长, 这限制了其应用的高效性. 因此, 我们借鉴已有的研究基础[17], 通过调整矩阵运算规则等操作实现了线性复杂度. 对于全局相关性的建模, 也可从以下两方面入手: (1) 利用多层图神经网络来建模多跳信息, 从而扩充节点的全局视野. (2) 数据增广. 对于原始图数据, 基于邻接矩阵, 利用图扩散[14]技术 (如 PageRank 等) 进行增广, 得到一个相对稠密的图, 在此图上进行表示学习可建模全局相关性. 在本文中, 我们选用了高效便捷的线性 Transformer 进行建模. 具体的操作过程如下.

对于任意图数据, 给定初始特征矩阵 $\boldsymbol{X}'$, 首先利用单层的神经网络将其映射到隐空间: $\boldsymbol{Z}'=h(\boldsymbol{X}')$. 接着, 我们

将此表示映射至查询空间、键空间和值空间, 继而进行注意力系数计算. 形式化定义如下:

$$\boldsymbol{Q} = f_Q(\boldsymbol{Z}'),\ \tilde{\boldsymbol{Q}} = \frac{\boldsymbol{Q}}{\|\boldsymbol{Q}\|} \tag{7}$$

$$\boldsymbol{K} = f_K(\boldsymbol{Z}'),\ \tilde{\boldsymbol{K}} = \frac{\boldsymbol{K}}{\|\boldsymbol{K}\|},\ \boldsymbol{V} = f_V(\boldsymbol{Z}') \tag{8}$$

其中, f_Q、f_K、f_V 是单层神经网络. 基于以上表示, all-pair 注意力分数的计算表示如下:

$$\boldsymbol{D} = \mathrm{diag}\left(\boldsymbol{1} + \frac{1}{n}\tilde{\boldsymbol{Q}}\left(\tilde{\boldsymbol{K}}^{\mathrm{T}}\boldsymbol{1}\right)\right) \tag{9}$$

经过全局信息传播的表示可被定义如下:

$$\boldsymbol{Z}_{\mathrm{global}} = \boldsymbol{D}^{-1}\left[\boldsymbol{V} + \frac{1}{n}\tilde{\boldsymbol{Q}}\left(\tilde{\boldsymbol{K}}^{\mathrm{T}}\boldsymbol{V}\right)\right] \tag{10}$$

其中, $\boldsymbol{1}$ 表示 n 维全 1 的列向量, $\mathrm{diag}(\cdot)$ 是将 n 维向量转换为 $n\times n$ 的对角矩阵的过程. 类似的, 我们将经过全局相关性建模的表示与初始表示相减, 得到全局视角下的残差特征:

$$\boldsymbol{H}_{\mathrm{global}} = \boldsymbol{Z}_{\mathrm{global}} - \boldsymbol{Z}' \tag{11}$$

3.3 基于注意力机制的重建学习和异常分数计算

在获得局部与全局视角下的残差特征表示后, 我们进而将二者按列拼接, 作为节点的最终表示为:

$$\boldsymbol{H} = [\boldsymbol{H}_{\mathrm{local}} \| \lambda \times \boldsymbol{H}_{\mathrm{global}}] \tag{12}$$

其中, λ 是可调节的参数, 用来平衡局部信息和全局信息的使用程度.

此外, 由于对于每个数据集, 我们有可利用的少量正常 (上下文) 节点. 为了进一步利用它们, 同时为方便在测试阶段进行更好的适配, 将模型的训练设计为基于这些节点的重建学习. 即对于每个节点, 利用已知的 n_k 个正常节点特征的组合来重建其特征信息. 在这里, 同样利用了注意力机制. 操作过程如下.

基于融合后的特征 $\boldsymbol{H} \in \mathbb{R}^{n\times d_r}$, 少量正常节点和待训练节点的特征可以拆分出来, 进而分别被定义为: $\boldsymbol{H}_k \in \mathbb{R}^{n_k\times d_r}$ 和 $\boldsymbol{H}_q \in \mathbb{R}^{n_q\times d_r}$. 重建过程则可被形式化表达如下:

$$\boldsymbol{Q}_r = \boldsymbol{H}_q\boldsymbol{W}_q,\ \boldsymbol{K}_r = \boldsymbol{H}_k\boldsymbol{W}_k \tag{13}$$

$$\check{\boldsymbol{H}}_q = Softmax\left(\frac{\boldsymbol{Q}_r\boldsymbol{K}_r^{\mathrm{T}}}{\sqrt{\boldsymbol{d}_r}}\right)\boldsymbol{H}_k \tag{14}$$

其中, $\boldsymbol{W}_q$ 和 $\boldsymbol{W}_k$ 为可学习的参数矩阵. 接着, 对于图上的每一个节点 $\boldsymbol{v}_i$, 损失函数可被定义如下:

$$\mathcal{L}\left(\boldsymbol{H}_{q_i}, \check{\boldsymbol{H}}_{q_i}, y_i\right) = \begin{cases} 1-\cos(\theta), & \text{if } y_i = 0 \\ \max(0, \cos(\theta) - margin), & \text{if } y_i = 1 \end{cases} \tag{15}$$

其中, *margin* 是超参数, 默认为 0, 用于控制不相似样本的惩罚强度, $\max(\cdot,\cdot)$ 是最大化的操作; $\cos(\cdot)$ 计算如公式 (16):

$$\cos(\theta) = \frac{\boldsymbol{H}_{q_i}\cdot\check{\boldsymbol{H}}_{q_i}}{\|\boldsymbol{H}_{q_i}\|\cdot\|\check{\boldsymbol{H}}_{q_i}\|} \tag{16}$$

对应地, 在测试过程中, 节点 v_i 的重建分数可被定义为:

$$s(v_i) = \sqrt{\sum_{j=1}^{d_r}\left(\boldsymbol{H}_{q_i\_j} - \check{\boldsymbol{H}}_{q_i\_j}\right)^2} \tag{17}$$

其中, $\boldsymbol{H}_{q_i\_j}$ 和 $\check{\boldsymbol{H}}_{q_i\_j}$ 分别为节点重建前后特征向量上第 j 维的值.

3.4 GRAD 方法流程

根据前文介绍, 我们给出所提方法 GRAD 的详细流程, 如算法 1 所示.

算法 1. GRAD 方法.

输入: 训练数据集 T_{train}, 建模残差特征以及重建学习中的多个初始化神经网络;
输出: 模型参数, 包括: 训练轮数 E; 特征传播次数 L.

初始化模型参数
#训练过程:
1. **for** $D^i_{\text{train}} \in T_{\text{train}}$ **do**
2. 基于公式 (1)、(2) 对齐特征;
3. **end**
4. **for** epoch in E **do**:
5. **for** $D^i_{\text{train}} \in T_{\text{train}}$ **do**
6. 获得图数据的节点特征矩阵 $\boldsymbol{X}'$, 邻接矩阵 $\boldsymbol{A}$, 标签信息 y;
7. **for** l in L **do**
8. $\boldsymbol{Z}^l \leftarrow$ 通过公式 (3)、(4) 计算节点在局部视角下的表示;
9. $\boldsymbol{R}^l \leftarrow$ 通过公式 (5)、 (6) 计算节点在局部视角下的残差特征;
10. **end**
11. $\boldsymbol{H}_{\text{local}} \leftarrow$ 将 L 层的残差特征进行按列拼接;
12. $\boldsymbol{H}_{\text{global}} \leftarrow$ 通过公式 (7)–(11) 计算节点在全局视角下的残差特征;
13. $\boldsymbol{H} \leftarrow$ 将局部和全局视角下的残差特征进行按列拼接;
14. $\boldsymbol{H}_k$ 和 $\boldsymbol{H}_q \leftarrow$ 划分出查询节点和上下文节点的特征;
15. $\mathcal{L} \leftarrow$ 通过公式 (15) 计算损失;
16. 基于梯度下降算法更新模型参数;
17. **end**
18. **end**
#测试过程:
1. **for** $D^i_{\text{test}} \in T_{\text{test}}$ **do**
2. 固定模型参数, 执行训练过程中步骤 6–14;
3. $s(\cdot) \leftarrow$ 通过公式 (17) 计算异常得分;
4. **end**

4 实验分析

在本节, 我们给出了基准数据集、度量标准以及与对比方法的介绍, 并设计实验来验证方法的有效性.

4.1 实验数据

为综合评估, 我们选取了来自多个领域的图数据集, 共 12 个, 其中包括社交网络、引文网络和电子商务评价网络. 遵循方法 ARC[10], 在每个领域上的大型数据集训练模型, 在剩余的数据集上进行测试. 具体地, 训练数据集 T_{train} 包括: PubMed、Flickr、Questions 和 YelpChi; 测试数据集 T_{test} 包括: Cora、CiteSeer、ACM、BlogCatalog、Facebook、Weibo、Reddit 和 Amazon. 这些数据集选自不同的领域, 以确保提出的模型可学习广泛的异常模式. 同时, 上述数据的多样性可以较好地评估模型适应新的和看不见的图数据的能力. 另外, 在每个数据集中, 正常 (上下文) 节点数量 n_k 设置为 10. 数据集的统计信息如表 1 所示, 详细信息如下.

- 引文网络: Cora、CiteSeer、PubMed、ACM, 节点为论文, 边为引用关系, 节点属性为词袋向量.
- 社交网络: BlogCatalog、Flickr: 节点为用户, 边为关注关系, 节点属性包含用户在社交网络中生成的个性化文本内容, 如博客文章或带有标签描述的共享照片. Facebook: 用户好友关系网络. Weibo: 用户-话题图, 短时密集发帖用户标记为可疑. Reddit: 论坛帖子网络, 被封禁用户被标记为异常节点, 属性为帖子文本向量. Question: 来自问答平台 Yandex Q, 节点为用户, 边表示一年内是否存在问答交互. 节点特征由用户描述文本的 FastText 词向量

均值生成.

● 电子商务评价网络: Amazon: 构建用户-产品关系图, 用于识别受雇撰写虚假产品评论的用户. YelpChi: 识别恶意评论, 基于同一用户的评论关系建图.

表 1 实验数据集的统计信息

类型	数据集	训练	测试	节点数量	边数量	特征维度	平均度	异常数量	异常比例 (%)
引用网络	Cora	—	√	2 708	5 429	1 433	3.90	150	5.53
	CiteSeer	—	√	3 327	4 732	3 703	2.77	150	4.50
	ACM	—	√	16 484	71 980	8 337	8.37	597	3.62
	PubMed	√	—	19 717	44 338	500	4.50	600	3.04
社交网络	BlogCatalog	—	√	5 196	171 743	8 189	66.11	300	5.77
	Flickr	√	—	7 575	239 738	12 047	63.30	450	5.94
	Facebook	—	√	1 081	55 104	576	50.97	25	2.31
	Weibo	—	√	8 405	407 963	400	48.53	868	10.30
	Reddit	—	√	10 984	168 016	64	15.30	366	3.33
	Questions	√	—	48 921	153 540	301	3.13	1 460	2.98
电子商务评价网络	Amazon	—	√	10 244	175 608	25	17.18	693	6.76
	YelpChi	√	—	23 831	49 315	32	2.07	1 217	5.10

4.2 评价指标及对比方法

在本文, 我们采用常用评价指标 AUROC 和 AUPRC 来评估检测模型的性能[45,46]. 对于 AUROC (area under the ROC curve, ROC 曲线下面积), 用于衡量模型在不同阈值下区分正负样本的能力, 取值范围为 [0, 1], 值越大表示模型性能越好. ROC 曲线由 *TPR* (true positive rate) 和 *FPR* (false positive rate) 构成, 计算方式如下:

$$\begin{cases} TPR = \dfrac{TP}{TP+FN} \\ FPR = \dfrac{FP}{FP+TN} \end{cases},$$

其中, *TP* (true positive) 表示正确预测的正样本, *FP* (false positive) 表示错误预测的正样本, *TN* (true negative) 表示正确预测的负样本, *FN* (false negative) 表示错误预测的负样本. 对于 AUPRC (area under the PR curve, PR 曲线下面积), 用于评估模型在正样本 (如异常) 上的精确率-召回率表现, 取值范围为 [0, 1], 值越接近 1 说明模型对正类的识别越精准. PR 曲线由 *Precision* (精确率) 和 *Recall* (召回率) 构成, 计算方式如下:

$$\begin{cases} Precision = \dfrac{TP}{TP+FP} \\ Recall = \dfrac{TP}{TP+FN} \end{cases}.$$

为验证所提出方法 GRAD 的有效性, 我们选取以下相关的方法进行比较.

● GCN[6]: 图神经网络领域的基础模型, 通过邻居聚合机制高效处理图结构数据.

● GAT[25]: 创新性地引入注意力机制, 动态调整节点贡献权重, 能根据不同下游任务优化注意力分布以获得更优质的节点表征.

● BGNN[47]: 将梯度提升决策树 (GBDT) 与 GNN 相结合, 专门处理含表格型节点特征的图数据. GBDT 负责特征处理, GNN 则建模图结构信息, 已在多种表格特征图数据上表现优异.

● BWGNN[16]: 创新之处在于配置频谱/空间局部带通滤波器, 可有效解决异常检测中的“右移现象”, 即谱能量集中分布于高频而非低频区域的特征问题.

● GHRN[48]: 基于图谱理论的监督式异常检测方法, 指出异质性与图的频率呈正相关. 通过增强高频成分来削减类间连接边, 显著提升异常检测性能.

● DOMINANT[49]: 关注图异常检测模型设计中网络稀疏性和数据非线性的问题, 结合 GCN 与深度自编码器,

通过联合重构邻接矩阵和节点特征, 依据重构误差识别结构/属性异常.

● CoLA[38]: 采用对比自监督学习, 通过采样对比实例对并利用图局部信息, 无监督地建模节点与其邻域子结构的关系, 将节点与子图的相似性作为自监督信号.

● HCM-A[50]: 将跳数预测作为自监督任务, 同时建模局部与全局上下文信息, 设计两种异常评分指标, 并引入贝叶斯学习优化异常捕获能力.

● TAM[39]: 基于单类同质性和局部亲和性原理设计, 通过在截断邻接矩阵上端到端优化自定义的亲和度指标来实现异常检测.

● ARC[10]: 一种基于上下文学习的可泛化的图异常检测方法, 能够仅使用少数正常节点实时识别异常, 该模型的关键设计在于采用先验启发的残差表示来解决局部信息的不一致问题.

4.3 实验结果与分析

表 2 和表 3 中展现了在基准数据集上我们所提出的方法 GRAD 与对比方法的性能结果. 可以发现, 对比不同范式下的方法, GRAD 有着较好的表现, 在不用微调的情况下, 在多个数据集上达到了具有竞争力的性能. 具体地, 相比于近期的 SOTA 方法 ARC, 在 8 个数据集上的平均 AUROC 和平均 AUPRC 分别提高了 1.9% 和 5.4%. 特别地, 在 Amazon 数据集上有着显著的提升, 潜在的原因是在该数据集中全局残差信息对异常检测的重要性较高. 对于在监督信息的引导下, 基于数据集进行事先训练的方法, 如 GCN 和 BGNN 等, 它们的竞争力是最弱的, 其性能受限于对新数据集的泛化能力, 这体现了对可泛化图异常检测模型的需求. 对于无监督范式下的方法, 如 CoLA 和 DOMINANT 等, 展现出了较好的竞争力, 这归功于它们设计针对图异常检测的预置任务, 来进行的通用表示学习, 这具有一定的泛化能力. 比如, 在 CoLA 中, 作者采样对比实例对并利用图局部信息, 无监督地建模节点与其邻域子结构的异常关系. 另一个发现是, 在测试中针对数据集进行微调时, 并没有展现出一致的性能提升, 如在 Facebook 和 Weibo 上出现了负迁移的情况, 这体现了在此问题中需要改进微调策略.

表 2 AUROC 方面的异常检测性能 (平均值±标准差) (%)

Type	Method	Cora	CiteSeer	ACM	BlogCatalog	Facebook	Weibo	Reddit	Amazon	Average
Supervised-pre-train only	GCN	59.64±8.3	60.27±8.1	60.49±9.6	56.19±6.3	29.51±4.8	76.64±17.6	50.43±4.4	46.63±3.4	54.98±6.6
	GAT	50.06±2.6	51.59±3.4	48.79±2.7	50.40±2.8	51.88±2.1	53.06±7.4	51.78±4.0	50.52±17.2	51.01±4.0
	BGNN	42.45±11.5	42.32±11.8	44.00±13.6	47.67±8.5	54.74±25.2	32.75±35.3	50.27±3.8	52.26±3.3	45.81±4.1
	BWGNN	54.06±3.2	52.61±2.8	67.59±0.7	56.34±1.2	45.84±4.9	53.38±1.6	48.97±5.7	55.26±16.9	54.26±3.4
	GHRN	59.89±6.5	60.27±8.1	60.49±9.6	56.19±6.3	29.51±4.8	76.64±17.6	50.43±4.4	46.63±3.4	54.98±6.6
Unsupervised-pre-train only	DOMINANT	66.53±1.1	69.47±2.0	70.08±2.3	74.25±0.6	51.01±0.7	**92.88±0.3**	50.05±4.9	48.94±2.6	65.40±1.8
	CoLA	63.29±8.8	62.84±9.5	66.85±4.4	50.04±3.2	12.99±11.6	16.27±5.6	52.81±6.6	47.40±7.9	46.56±6.0
	HCM-A	54.28±4.7	48.12±6.8	53.70±4.6	55.31±0.5	35.44±13.9	65.52±12.5	48.79±2.7	43.99±0.7	50.64±3.3
	TAM	62.02±2.3	72.27±0.8	74.43±1.5	49.86±0.7	65.88±6.6	71.54±0.1	55.43±0.3	56.06±2.1	63.44±1.8
Unsupervised-pre-train & fine-tune	DOMINANT	72.23±0.3	74.69±0.3	74.34±0.1	74.61±0.1	49.92±0.5	92.21±0.1	52.14±5.0	59.06±2.8	68.65±1.1
	CoLA	67.62±4.2	70.75±3.4	69.11±0.6	62.49±3.3	64.70±18.8	31.55±6.0	58.12±0.6	52.51±6.6	59.61±4.2
	HCM-A	56.45±4.9	55.54±4.0	57.69±3.5	55.10±0.2	36.57±10.7	71.89±2.7	49.15±2.7	42.20±0.5	53.07±2.4
	TAM	62.56±2.1	76.54±1.3	**86.29±1.5**	57.69±0.8	**76.26±3.7**	71.73±0.1	56.62±0.4	57.13±1.5	68.10±1.4
Few-shot without fine-tune	ARC	85.33±0.4	90.64±0.3	79.27±0.1	**74.81±0.2**	69.57±1.4	89.24±0.3	58.95±1.2	69.77±3.8	77.20±0.9
	Ours	**86.89±0.8**	**91.07±0.5**	79.95±0.1	74.26±0.1	65.46±1.8	90.07±0.2	**60.01±0.4**	**84.66±2.9**	**79.05±0.8**

注: 加粗为目前最好的实验结果, 下划线为次好的实验结果

表 3 AUPRC 方面的异常检测性能 (平均值±标准差) (%)

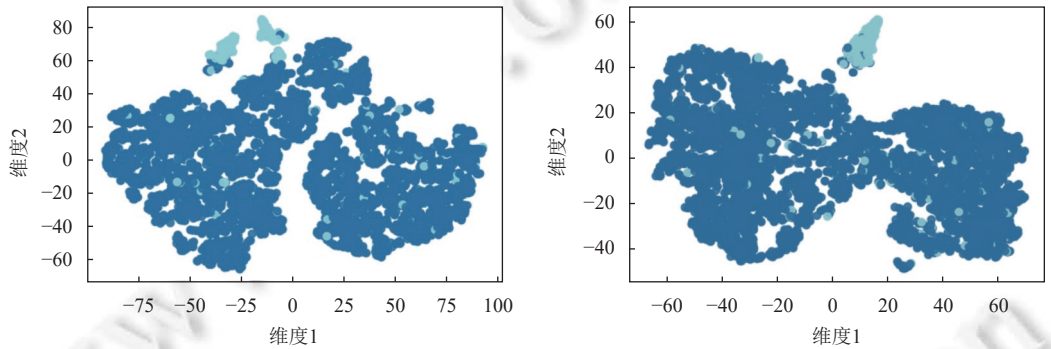
Type	Method	Cora	CiteSeer	ACM	BlogCatalog	Facebook	Weibo	Reddit	Amazon	Average
Supervised-pre-train only	GCN	7.41±1.5	6.40±1.4	5.27±1.1	7.44±1.0	1.59±0.1	67.21±15.2	3.39±0.3	6.96±2.0	13.21±2.8
	GAT	6.49±0.8	5.58±0.6	4.70±0.7	12.81±2.0	3.14±0.3	33.34±9.8	3.73±0.5	15.74±17.8	10.69±4.1
	BGNN	4.90±1.2	3.91±1.0	3.48±1.3	5.73±1.4	3.81±2.1	30.26±29.9	3.52±0.5	7.51±0.5	7.89±4.7
	BWGNN	7.25±0.8	6.35±0.7	7.14±0.2	8.99±1.1	2.54±0.6	12.13±0.7	3.69±0.8	13.12±11.8	7.65±2.1
	GHRN	9.56±2.4	7.79±2.0	5.61±0.7	10.94±2.5	2.41±0.6	28.53±7.3	3.24±0.3	7.54±2.0	9.45±2.2

表 3 AUPRC 方面的异常检测性能 (平均值±标准差) (%)(续)

Type	Method	Cora	CiteSeer	ACM	BlogCatalog	Facebook	Weibo	Reddit	Amazon	Average
Unsupervised-pre-train only	DOMINANT	12.75±0.7	13.85±2.3	15.59±2.6	35.22±0.8	2.95±0.1	**81.47±0.2**	3.49±0.4	6.11±0.2	21.42±0.9
	CoLA	11.41±3.5	8.33±3.7	7.31±1.4	6.04±0.5	1.90±0.6	7.59±3.2	3.71±0.6	11.06±4.4	7.16±2.2
	HCM-A	5.78±0.7	4.18±0.7	4.01±0.6	6.89±0.3	2.08±0.6	21.91±11.8	3.18±0.2	5.87±0.0	6.73±1.8
	TAM	11.18±0.7	11.55±0.4	23.20±2.3	10.57±1.1	8.40±0.9	16.46±0.1	3.94±0.1	10.75±3.1	12.01±1.1
Unsupervised-pre-train & fine-tune	DOMINANT	21.35±0.7	23.02±1.5	22.74±0.9	35.79±0.6	3.56±0.1	77.69±1.4	3.84±0.7	7.48±0.4	24.43±0.8
	CoLA	13.91±5.5	19.51±3.7	8.48±0.5	10.43±1.2	**15.19±11.7**	8.03±1.1	4.07±0.1	7.27±1.1	10.86±3.0
	HCM-A	6.41±1.3	4.76±0.5	4.41±0.6	6.62±0.1	2.23±0.7	27.20±5.5	3.10±0.1	5.64±0.1	7.54±1.1
	TAM	13.62±0.5	18.66±1.4	**58.04±8.1**	13.90±0.5	11.11±3.2	16.47±0.0	3.93±0.1	11.56±1.8	18.41±1.9
Few-shot without fine-tune	ARC	47.26±0.9	**48.38±0.7**	40.14±0.1	34.80±0.1	8.96±1.7	65.46±0.9	4.14±0.2	15.45±4.3	33.07±1.1
	Ours	**49.27±1.1**	46.58±1.5	40.47±0.1	**36.31±0.1**	5.36±1.1	65.21±0.6	**4.55±0.1**	**60.24±4.8**	**38.49±1.0**

注: 加粗为目前最好的实验结果, 下划线为次好的实验结果

同时, 我们在 Amazon 数据集上对残差特征进行了可视化, 如图 4 所示, 其中深色为正常类样本, 浅色为异常类样本. 可以发现, 全局残差特征可以辅助局部残差特征使得正常和异常类中样本之间的距离更近, 即类内更紧, 类间更开, 其中在异常类的可视化中较为明显, 这也反映了全局残差特征可在一定程度上提升分类的准确性.



(a) 局部残差特征 (b) 全局残差特征

图 4 可视化展示

4.4 消融实验

本节探究不同组件对所提模型 GRAD 的影响. 具体地, 我们设计了 3 种变体, 分别是: 1) 将残差表示舍弃, 只使用原本隐表示的变体: 无残差; 2) 将原始图结构舍弃, 只保留全局信息的变体: 无局部. 此变体同时也模拟了图结构缺失的场景. 在此场景下, 只利用线性 Transformer 编码器对隐关系进行建模; 3) 将全局信息的建模舍弃, 只使用基于 GNN 的局部信息建模: 无全局. 实验结果展示在表 4 中.

表 4 在 8 个数据集上的消融实验 (%)

变体	Cora	CiteSeer	ACM	BlogCatalog	Facebook	Weibo	Reddit	Amazon
原模型	**86.89±0.8**	**91.07±0.5**	**79.95±0.1**	74.26±0.1	65.46±1.8	90.07±0.2	**60.01±0.4**	**84.66±2.9**
无残差	61.49±1.3	62.28±2.7	73.00±0.6	73.89±0.3	47.00±2.7	**94.36±0.1**	46.95±3.7	79.62±5.0
无局部	73.72±0.8	75.85±0.4	74.95±0.2	**75.77±0.6**	54.17±1.2	90.70±0.5	56.06±3.2	84.59±2.1
无全局	85.60±0.8	89.85±0.6	79.41±0.3	74.10±0.2	**66.59±2.0**	89.47±0.9	59.49±0.6	72.56±5.1

注: 加粗为目前最好的实验结果, 下划线为次好的实验结果

可以发现, 3 个组件都对模型性能具有一定的贡献. 特别地, 残差信息的影响最大, 这体现了在可泛化的图异常检测场景下提取数据集无关的特征的必要性. 一个特例是在 Weibo 数据集上的表现是相反的, 潜在的原因是该数据集的异常模式倾向存在于原始隐表示中, 该现象启发了我们对残差表示自适应学习的探索, 将作为未来的研

究方向之一. 另外, 可以发现, 在结构缺失的情况下, 受益于提出的全局信息模型, 变体无局部的性能在大部分数据集上是较好的, 其中, 在 BlogCatalog 上展现出了最佳性能. 需要说明的是, 该处的无全局变体可视作为方法 ARC, 性能有区别的原因在于无全局变体的实现是在 GRAD 的基础上进行, 在超参数的设置上与 ARC 有所不同. 可以发现, 在 8 个数据集的平均 AUROC 性能上, ARC 为 77.20%, 无全局变体为 77.13%, GRAD 为 79.05%, 因此可验证所提出方法的有效性.

4.5 参数影响分析

本节探究模型中 3 个参数的敏感性, 即局部与全局残差信息之间的权衡参数 λ、隐藏层大小 d_r 和正常样本数据 n_k. 对于 λ, 我们的参数选择范围设置在{1, 0.1, 0.01, 0.001, 0.000 1}之中. 从图 5 中可以发现, 如 Cora、CiteSeer、Facebook 这类节点数量较少的数据集, 性能最优在 0.001 处取得, 如 Reddit 和 Amazon 这类节点数量较多的数据集, 性能最优在 0.01 处取得. 背后的原因在于在节点数量多的图中, 全局信息更为丰富, 在此任务中的贡献会较大. 为更符合现实场景, 我们将此参数统一设置为 0.01. 对于 n_k, 我们的参数选择范围设置在{2, 5, 10, 15, 20}之中, 对应的结果展示在图 6. 可以发现, 整体的趋势是, 模型的性能随着上下文正常样本数量的增加而提高. 同时, 在上下文节点数量较小时, 模型的性能在大多数数据集上的表现仍然是良好的. 对于 d_r, 我们的参数选择范围设置在: {128, 256, 512, 1 024, 2 048}之中. 从图 7 中的结果可以发现的是, 大部分数据集在 1 024 处取得最佳性能, 在大于 1 024 后的性能增益较少. 同时, 该参数的选择是相对不敏感的, 较小的取值也能达到较好的性能表现. 在本文中, 该值被统一设置为 1 024.

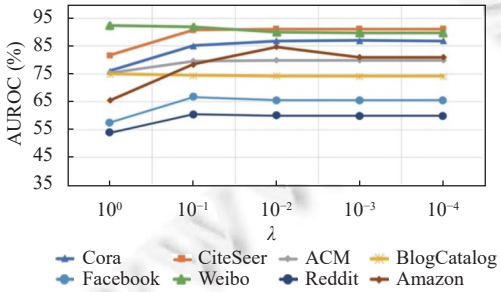


图 5 局部与全局信息的权衡参数 λ

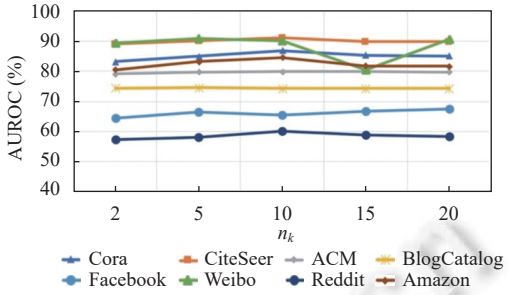


图 6 正常样本数量参数 n_k

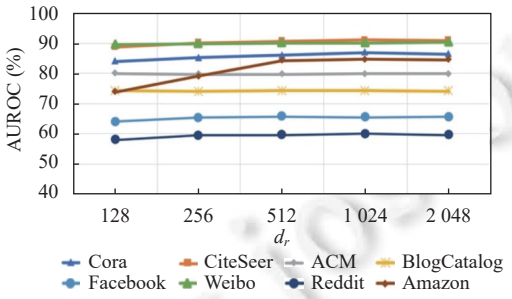


图 7 隐藏层大小参数 d_r

5 总 结

图异常检测作为识别网络中显著偏离常规节点的重要任务, 在医疗、金融等隐私敏感领域具有广泛应用. 但现有方法面临两大核心问题: 一是依赖数据集特定训练范式导致的高成本与弱泛化性, 难以适应跨数据应用等场景; 二是基于邻居信息的残差特征策略存在固有缺陷, 包括 GNN 建模的局限性、幂律分布下的邻居稀疏性以及噪声干扰. 针对这些问题, 本文提出全局与局部残差信息联合感知的泛化图异常检测方法 GRAD. 该方法通过结合

线性 Transformer 高效地建模全局节点表征, 并在全局和局部视角下将节点表示转换为自身及邻居之间的残差形式以生成数据集无关的通用特征, 最后结合少样本重建学习实现无需目标数据微调的高效检测. 实验表明, GRAD 在跨领域基准数据集上优于现有方法. 在本文, 可以发现, 所计算出的残差特征可视为对一阶信息的利用, 所以对于多阶残差信息的有效融合是未来值得探索的内容. 同时, 对于线性 Transformer, 它仍是基于多层神经网络进行建模, 这扩大了模型规模, 在异常检测这种数据量少的场景下, 增加了模型过拟合的风险. 因此, 利用数据增广技术来扩充训练数据也是未来的研究方向之一.

References

[1] Wu ZH, Pan SR, Chen FW, Long GD, Zhang CQ, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans. on Neural Networks and Learning Systems, 2021, 32(1): 4–24. [doi: 10.1109/TNNLS.2020.2978386]

[2] Qiao HZ, Tong HH, An B, King I, Aggarwal C, Pang GS. Deep graph anomaly detection: A survey and new perspectives. IEEE Trans. on Knowledge and Data Engineering, 2025, 37(9): 5106–5126. [doi: 10.1109/TKDE.2025.3581578]

[3] Pan JJ, Liu YX, Zheng X, Zheng YZ, Liew AWC, Li FY, Pan SR. A label-free heterophily-guided approach for unsupervised graph fraud detection. In: Proc. of the 39th AAAI Conf. on Artificial Intelligence. Philadelphia: AAAI Press, 2025. 12443–12451. [doi: 10.1609/aaai.v39i12.33356]

[4] Li JD, Dani H, Hu X, Liu H. Radar: Residual analysis for anomaly detection in attributed networks. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. Melbourne: AAAI Press, 2017. 2152–2158.

[5] Erfani SM, Rajasegarar S, Karunasekera S, Leckie C. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recognition, 2016, 58: 121–134. [doi: 10.1016/j.patcog.2016.03.028]

[6] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv:1609.02907, 2017.

[7] Xiao SX, Wang SP, Dai YF, Guo WZ. Graph neural networks in node classification: Survey and evaluation. Machine Vision and Applications, 2022, 33(1): 4. [doi: 10.1007/s00138-021-01251-0]

[8] Zhang MH, Chen YX. Link prediction based on graph neural networks. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 5171–5181.

[9] Ma XX, Wu J, Xue S, Yang J, Zhou C, Sheng QZ, Xiong H, Akoglu L. A comprehensive survey on graph anomaly detection with deep learning. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(12): 12012–12038. [doi: 10.1109/TKDE.2021.3118815]

[10] Liu YX, Li SY, Zheng Y, Chen QF, Zhang CQ, Pan SR. ARC: A generalist graph anomaly detector with in-context learning. In: Proc. of the 38th Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2025. 50772–50804.

[11] Qiao HZ, Niu CX, Chen L, Pang GS. AnomalyGFM: Graph foundation model for zero/few-shot anomaly detection. In: Proc. of the 31st ACM SIGKDD Conf. on Knowledge Discovery and Data Mining V.2. Toronto: ACM, 2025. 2326–2337. [doi: 10.1145/3711896.3736843]

[12] Zhu JW, Pang GS. Toward generalist anomaly detection via in-context residual learning with few-shot sample prompts. In: Proc. of the 2024 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024. 17826–17836. [doi: 10.1109/CVPR52733.2024.01688]

[13] Zheng WQ, Huang EW, Rao N, Katariya S, Wang ZY, Subbian K. Cold brew: Distilling graph node representations with incomplete or missing neighborhoods. arXiv:2111.04840, 2022.

[14] Zhang JQ, Wang SZ, Chen SC. Reconstruction enhanced multi-view contrastive learning for anomaly detection on attributed networks. arXiv:2205.04816, 2022.

[15] Jin M, Liu YX, Zheng Y, Chi LH, Li YF, Pan SR. ANEMONE: Graph anomaly detection with multi-scale contrastive learning. In: Proc. of the 30th Int'l Conf. on Information & Knowledge Management. Queensland: ACM, 2021. 3122–3126. [doi: 10.1145/3459637.3482057]

[16] Tang JH, Li JJ, Gao ZQ, Li J. Rethinking graph neural networks for anomaly detection. In: Proc. of the 39th Int'l Conf. on Machine Learning. Maryland: PMLR, 2022. 21076–21089.

[17] Wu QT, Zhao WT, Yang CX, Zhang HR, Nie F, Jiang HT, Bian YT, Yan JC. SGFormer: Simplifying and empowering Transformers for large-graph representations. In: Proc. of the 37th Int'l Conf. on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2023. 64753–64773.

[18] Wu QT, Zhao WT, Li Z, Wipf DP, Yan JC. Nodeformer: A scalable graph structure learning Transformer for node classification. In: Proc. of the 36th Int'l Conf. on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2022. 27387–27401.

[19] Geisler S, Kosmala A, Herbst D, Günnemann S. Spatio-spectral graph neural networks. In: Proc. of the 38th Int'l Conf. on Neural

Information Processing Systems. Vancouver: Curran Associates Inc., 2025. 49022–49080.

[20] Han HY, Liu XR, Shi F, Torkamani MA, Aggarwal CC, Tang JL. Towards label position bias in graph neural networks. In: Proc. of the 37th Int'l Conf. on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2023. 22999–23018.

[21] Wu SW, Sun F, Zhang WT, Xie X, Cui B. Graph neural networks in recommender systems: A survey. ACM Computing Surveys, 2022, 55(5): 97. [doi: 10.1145/3535101]

[22] Liu Q, Allamanis M, Brockschmidt M, Gaunt AL. Constrained graph variational autoencoders for molecule design. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 7806–7815.

[23] Wang SZ, Cao JN, Yu PS. Deep learning for spatio-temporal data mining: A survey. IEEE Trans. on Knowledge and Data Engineering, 2022, 34(8): 3681–3700. [doi: 10.1109/TKDE.2020.3025580]

[24] Hamilton WL, Ying R, Leskovec J. Inductive representation learning on large graphs. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 1025–1035.

[25] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. In: Proc. of the 6th Int'l Conf. on Learning Representations. OpenReview.net. 2018.

[26] Wu F, Zhang TY, de Souza AH Jr, Fifty C, Yu T, Weinberger KQ. Simplifying graph convolutional networks. In: Proc. of the 36th Int'l Conf. on Machine Learning. Long Beach: PMLR, 2019. 6861–6871.

[27] Bo DY, Wang X, Shi C, Shen HW. Beyond low-frequency information in graph convolutional networks. In: Proc. of the 35th AAAI Conf. on Artificial Intelligence. AAAI Press, 2021. 3950–3957. [doi: 10.1609/aaai.v35i5.16514]

[28] Yuan RW, Tang YQ, Zhang WS. A structure-aware invariant learning framework for node-level graph OOD generalization. In: Proc. of the 31st SIGKDD Conf. on Knowledge Discovery and Data Mining V.1. Toronto: ACM, 2025. 1879–1890. [doi: 10.1145/3690624.3709227]

[29] Xu YM, Shi B, Peng Z, Liu HX, Dong B, Chen C. Out-of-distribution generalization on graphs via progressive inference. In: Proc. of the 39th AAAI Conf. on Artificial Intelligence. Philadelphia: AAAI Press, 2025. 12963–12971. [doi: 10.1609/aaai.v39i12.33414]

[30] Liu ZW, Dou YT, Yu PS, Deng YT, Peng H. Alleviating the inconsistency problem of applying graph neural network to fraud detection. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2020. 1569–1572. [doi: 10.1145/3397271.3401253]

[31] Breunig MM, Kriegel HP, Ng RT, Sander J. LOF: Identifying density-based local outliers. In: Proc. of the 2000 ACM SIGMOD Int'l Conf. on Management of Data. Dallas: ACM, 2000. 93–104. [doi: 10.1145/342009.335388]

[32] Perozzi B, Akoglu L. Scalable anomaly ranking of attributed neighborhoods. In: Proc. of the 2016 SIAM Int'l Conf. on Data Mining. Miami: Society for Industrial and Applied Mathematics, 2016. 207–215.

[33] Guo JY, Li RH, Zhang Y, Wang GR. Graph neural network based anomaly detection in dynamic networks. Ruan Jian Xue Bao/Journal of Software, 2020, 31(3): 748–762 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5903.htm [doi: 10.13328/j.cnki.jos.005903]

[34] Gao Y, Wang X, He XN, Liu ZG, Feng HM, Zhang YD. Alleviating structural distribution shift in graph anomaly detection. In: Proc. of the 16th ACM Int'l Conf. on Web Search and Data Mining. Singapore: ACM, 2023. 357–365. [doi: 10.1145/3539597.3570377]

[35] Liu YX, Jin M, Pan SR, Zhou C, Zheng Y, Xia F, Yu PS. Graph self-supervised learning: A survey. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(6): 5879–5900. [doi: 10.1109/TKDE.2022.3172903]

[36] Duan JC, Wang SW, Zhang P, Zhu E, Hu JT, Jin H, Liu Y, Dong ZB. Graph anomaly detection via multi-scale contrastive learning networks with augmented view. In: Proc. of the 37th AAAI Conf. on Artificial Intelligence. Washington: AAAI Press, 2023. 7459–7467. [doi: 10.1609/aaai.v37i6.25907]

[37] Zheng Y, Jin M, Liu YX, Chi LH, Phan KT, Chen YPP. Generative and contrastive self-supervised learning for graph anomaly detection. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(12): 12220–12233. [doi: 10.1109/TKDE.2021.3119326]

[38] Liu YX, Li Z, Pan SR, Gong C, Zhou C, Karypis G. Anomaly detection on attributed networks via contrastive self-supervised learning. IEEE Trans. on Neural Networks and Learning Systems, 2022, 33(6): 2378–2392. [doi: 10.1109/TNNLS.2021.3068344]

[39] Qiao HZ, Pang GS. Truncated affinity maximization: One-class homophily modeling for graph anomaly detection. In: Proc. of the 37th Int'l Conf. on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2023. 49490–49512.

[40] Qiao HZ, Wen QS, Li XL, Lim EP, Pang GS. Generative semi-supervised graph anomaly detection. In: Proc. of the 38th Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2025. 4660–4688.

[41] Wu JH, Duan L, Yue K, Li AS, Yang PZ. Structural-entropy-based anomaly detection in attributed graph. Ruan Jian Xue Bao/Journal of Software, 2025, 36(11): 5031–5044 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/7374.htm [doi: 10.13328/j.cnki.jos.007374]

[42] Stewart GW. On the early history of the singular value decomposition. SIAM Review, 1993, 35(4): 551–566. [doi: 10.1137/1035134]

[43] Abdi H, Williams LJ. Principal component analysis. WIREs Computational Statistics, 2010, 2(4): 433–459. [doi: 10.1002/wics.101]

[44] Hardoon DR, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: An overview with application to learning methods. Neural Computation, 2004, 16(12): 2639–2664. [doi: 10.1162/0899766042321814]

[45] Tang JH, Hua FR, Gao ZQ, Zhao PL, Li J. GADBench: Revisiting and benchmarking supervised graph anomaly detection. In: Proc. of the 37th Int'l Conf. on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2023. 29628–29653.

[46] Pang GS, van den Hengel A, Shen CH, Cao LB. Toward deep supervised anomaly detection: Reinforcement learning from partially labeled anomaly data. In: Proc. of the 27th SIGKDD Conf. on Knowledge Discovery & Data Mining. Singapore: ACM, 2021. 1298–1308. [doi: 10.1145/3447548.3467417]

[47] Ivanov S, Prokhorenkova L. Boost then convolve: Gradient boosting meets graph neural networks. In: Proc. of the 2021 Int'l Conf. on Learning Representations. Vienna: OpenReview.net. 2021.

[48] Gao Y, Wang X, He XN, Liu ZG, Feng HM, Zhang YD. Addressing heterophily in graph anomaly detection: A perspective of graph spectrum. In: Proc. of the 2023 ACM Web Conf. Austin: ACM, 2023. 1528–1538. [doi: 10.1145/3543507.3583268]

[49] Ding KZ, Li JD, Bhanushali R, Liu H. Deep anomaly detection on attributed networks. In: Proc. of the 2019 SIAM Int'l Conf. on Data Mining. Calgary: Society for Industrial and Applied Mathematics, 2019. 594–602.

[50] Huang TJ, Pei YL, Menkovski V, Pechenizkiy M. Hop-count based self-supervised anomaly detection on attributed networks. In: Proc. of the 2023 European Conf. on Machine Learning and Knowledge Discovery in Databases. Grenoble: Springer, 2023. 225–241. [doi: 10.1007/978-3-031-26387-3_14]

附中文参考文献

[33] 郭嘉琰, 李荣华, 张岩, 王国仁. 基于图神经网络的动态网络异常检测算法. 软件学报, 2020, 31(3): 748–762. http://www.jos.org.cn/1000-9825/5903.htm [doi: 10.13328/j.cnki.jos.005903]

[41] 吴江豪, 段亮, 岳昆, 李昂生, 杨培忠. 基于结构熵的属性图异常检测. 软件学报, 2025, 36(11): 5031–5044. http://www.jos.org.cn/1000-9825/7374.htm [doi: 10.13328/j.cnki.jos.007374]

作者简介

张家强, 博士生, CCF 学生会员, 主要研究领域为图机器学习, 数据挖掘.

陈松灿, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为机器学习, 模式识别, 神经计算.