软件学报 ISSN 1000-9825, CODEN RUXUEW
2026,37(2):860−874 [doi: 10.13328/j.cnki.jos.007423] [CSTR: 32375.14.jos.007423]

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

基于链路聚合的图欺诈检测*

邱 天[1,2], 贾凌翔[1,2], 高 杨[1,2], 冯尊磊[1,2], 高 艺[1,2], 宋明黎[1,2]

[1](区块链与数据安全全国重点实验室 (浙江大学), 浙江 杭州 310027)

[2](杭州高新区 (滨江) 区块链与数据安全研究院, 浙江 杭州 310051)

通信作者: 冯尊磊, E-mail: zunleifeng@zju.edu.cn

摘 要: 随着信息技术发展, 信息网络、人类社会与物理空间交互加深, 信息空间风险外溢现象严峻. 欺诈事件激增, 欺诈检测成为重要研究领域. 欺诈行为给社会带来了诸多负面影响, 且逐渐呈现出智能化、产业化及高度隐蔽性等新兴特征, 传统的专家规则与深度图神经网络算法在应对上显得愈发局限. 当前反欺诈算法多从节点自身与邻居节点的局部信息出发, 或聚焦于用户个体, 或分析节点与网络拓扑关系, 或利用图嵌入技术学习节点表示, 这些视角虽然能具备一定的欺诈检测能力, 但是忽略了实体长程关联模式的关键作用, 缺乏对于海量欺诈链路之间共性模式的挖掘, 限制了全面的欺诈检测能力. 针对以上欺诈检测算法的局限性, 提出一种基于链路聚合的图欺诈检测模型 PA-GNN (path aggregation graph neural network), 包含不定长链路采样, 位置关联的统一链路编码, 链路信息交互聚合, 以及聚合关联的欺诈检测. 从节点出发的若干链路之间通过全局模式交互与相似度比对, 挖掘欺诈链路之间的共性规律, 从而更全面地揭示欺诈行为之间的关联模式, 并通过链路聚合继而实现欺诈检测. 在金融交易、社交网络和评论网络这 3 类欺诈场景下的多个数据集上的实验结果表明, 所提方法的曲线下面积 (AUC) 和平均精度 (AP) 指标相较于最优基准模型均有显著提升. 此外, 该方法为欺诈检测任务挖掘了潜在的共性欺诈链路模式, 驱动节点学习这些重要的模式并获得更具表现力的表示, 具备一定的可解释性.

关键词: 图神经网络; 欺诈检测; 链路聚合; 注意力机制; 特征表示

中图法分类号: TP309

中文引用格式: 邱天, 贾凌翔, 高杨, 冯尊磊, 高艺, 宋明黎. 基于链路聚合的图欺诈检测. 软件学报, 2026, 37(2): 860–874. http://www.jos.org.cn/1000-9825/7423.htm

英文引用格式: Qiu T, Jia LX, Gao Y, Feng ZL, Gao Y, Song ML. Path-aggregation-based Graph Fraud Detection. Ruan Jian Xue Bao/Journal of Software, 2026, 37(2): 860–874 (in Chinese). http://www.jos.org.cn/1000-9825/7423.htm

Path-aggregation-based Graph Fraud Detection

QIU Tian[1,2], JIA Ling-Xiang[1,2], GAO Yang[1,2], FENG Zun-Lei[1,2], GAO Yi[1,2], SONG Ming-Li[1,2]

[1](State Key Laboratory of Blockchain and Data Security (Zhejiang University), Hangzhou 310027, China)

[2](Hangzhou High-tech Zone (Binjiang) Institute of Blockchain and Data Security, Hangzhou 310051, China)

Abstract: With the development of information technology, the interaction between information networks, human society, and physical space deepens, and the phenomenon of information space risk overflow becomes more severe. Fraudulent incidents have sharply increased, making fraud detection an important research field. Fraudulent behavior has brought numerous negative impacts to society, gradually presenting emerging characteristics such as intelligence, industrialization, and high concealment. Traditional expert rules and deep graph neural network algorithms are becoming increasingly limited in addressing fraudulent activities. Current fraud detection methods often rely on local information from the nodes themselves and neighboring nodes, either focusing on individual users, analyzing the relationship between nodes and graph topology, or utilizing graph embedding technology to learn node representations. Although these approaches offer

* 基金项目: 国家重点研发计划 (2022YFB2703100)
收稿时间: 2024-08-16; 修改时间: 2025-02-05; 采用时间: 2025-03-03; jos 在线出版时间: 2025-07-30
CNKI 网络首发时间: 2025-07-31

certain fraud detection capabilities, they overlook the crucial role of long-range association patterns of entities and fail to explore common patterns among massive fraudulent paths, limiting comprehensive fraud detection capabilities. In response to the limitations of existing fraud detection methods, this study proposes a graph fraud detection model called path aggregation graph neural network (PA-GNN), based on path aggregation. The model includes variable-length path sampling, position-related unified path encoding, path interaction and aggregation, and aggregation-related fraud detection. Several paths originating from a node interact globally and compare their similarities, extracting common patterns among fraudulent paths, thus more comprehensively revealing the association patterns between fraudulent behaviors, and achieving fraud detection through path aggregation. Experimental results across multiple datasets in fraud scenarios, including financial transactions, social networks, and review networks, show that the area under the curve (AUC) and average precision (AP) metrics of the proposed method have significantly improved compared to the optimal benchmark models. In addition, the proposed method uncovers potential common fraudulent path patterns for fraud detection tasks, driving nodes to learn these important patterns and obtain more expressive representations, which offers a certain level of interpretability.

Key words: graph neural network (GNN); fraud detection; path aggregation; attention mechanism; feature representation

随着信息技术的不断演进, 信息网络空间、人类社会空间与实体物理空间之间的三元交互融合日益加深, 导致信息空间的风险显著外溢至人类社会和物理空间, 现象愈发严峻. 由于欺诈事件急剧增加, 欺诈检测[1−5]已成为一个日益突出的研究领域. 欺诈行为涵盖了金融交易、社交网络以及评论网络中的各种形式. 金融欺诈包括虚假宣传、内幕交易和操纵市场等手段, 旨在非法获取财富或操纵利益分配. 社交欺诈涉及虚构身份、伪造关系和欺骗他人的信任, 以获取个人利益或满足欲望. 评论欺诈指的是在网络平台或社交媒体上, 通过发布虚假、夸大或误导性的评论、评价或反馈, 以操纵公众舆论、影响消费者决策或提升特定产品、服务、个人或品牌的声誉. 这些欺诈行为给社会带来了严重的负面影响, 导致经济损失、破坏信任关系、破坏公正竞争环境以及损害消费者权益等问题. 因此, 打击欺诈行为和保护公众利益非常重要.

传统的基于专家规则的方法通过对数据进行分析和建模, 能够识别出一些常见的欺诈模式. 然而, 专家规则缺乏对不同欺诈场景的通用性. 因此, 引入人工智能技术成为了解决欺诈检测问题的新思路. 深度学习[6,7]作为人工智能技术的重要分支, 具有强大的数据处理和模式识别能力, 对欺诈检测展现出巨大潜力. 考虑到欺诈样本间存在错综复杂的拓扑结构和依赖关系, 研究人员提出将图神经网络 (graph neural network, GNN)[8−12]作为图深度学习解决方案.

随着欺诈行为逐渐呈现出智能化、产业化、高度隐蔽性等新兴特征, 目前的深度图神经网络算法已愈发难以满足欺诈检测的准确性要求. 对于欺诈检测的研究, 该问题的主要难点有: (1) 欺诈样本间实体通常具有高阶交互. 欺诈行为不是由单一的实体或简单的关系所构成的, 而是由多个实体通过复杂的交互作用共同形成的. 这种高阶交互关系使得欺诈行为的识别和检测变得更加困难, 因为需要同时考虑多个实体和它们之间的关系, 以及这些关系如何共同作用于欺诈行为. (2) 欺诈样本间依赖关系错综复杂. 这些依赖关系可能是隐式的, 也可能是显式的, 它们共同构成了欺诈行为的复杂网络. 在这种复杂的依赖关系中, 一个实体的变化可能会影响其他多个实体, 进而改变整个欺诈行为的模式. (3) 欺诈样本的高度隐蔽性. 欺诈者可能会采用各种手段来掩盖其欺诈行为, 如使用虚假身份、伪造数据、隐藏关键信息等. 这些手段使得欺诈行为在表面上看起来与正常行为无异, 如果仅局限于分析欺诈实体周围的局部信息, 将难以有效识别和检测这些经过精心伪装的欺诈活动. 这些特征共同构成了当前欺诈检测研究的主要挑战.

对于传统的图神经网络, 无论是 GCN[13]、GraphSAGE[14], 还是 GAT[15], 均属于局部网络范畴, 它们的共同局限性在于无法有效处理具有长程依赖特性的图结构. 这是因为它们主要聚焦于一阶邻居节点的信息, 对全局信息的利用相对有限, 不能很好地处理全图规模的任务. 此外, 传统图神经网络在多次迭代后会导致“过度平滑 (over-smoothing)”的问题, 即节点表示趋于相似, 使模型容易受到无关噪声的影响并且缺乏可解释性. 近年来, 图欺诈检测技术取得了显著进展, 主要可划分为谱方法和空间方法两大类. 谱方法侧重于利用图信号的不同频率成分, 通过组合多频信号[16]、分析光谱能量分布[17]以及运用标签感知的边缘指标[18]等策略, 优化或增强对图上欺诈信息的捕捉与处理能力. 相较于谱方法, 空间方法则更加关注节点间的空间关系, 采用上下文与关系融合[19]、类别平衡与采样优化[20]、强化关系感知的邻居选择[21]、同质与异质连接管理[22]以及异常特征识别与限制[23]等手段, 提升

图神经网络对欺诈行为和关系的检测与识别能力. 上述方法在图欺诈检测的先前研究中均展现出了优异的性能. 然而, 从本质上来说, 上述方法也都仅考虑了低阶邻居信息, 其“感受野”相对有限. 对于更为隐蔽复杂、涉及长程链路关联的欺诈行为, 它们往往难以有效应对. 因此, 为了提升检测精度, 有必要将分析范围扩展到更广泛的上下文信息.

为了有效检测欺诈行为, 需要采用更全局的检测技术, 来挖掘和识别隐藏在大量数据中的欺诈模式. 如何更有效地捕捉和利用图数据中的长程依赖关系仍是一个重要的研究方向. 针对上述图欺诈检测存在的问题, 本文提出了一种基于链路聚合的图欺诈检测模型 PA-GNN (path aggregation graph neural network), 如图 1 所示. PA-GNN 由 4 个部分构成: 不定长链路采样、位置关联的统一链路编码、链路信息交互聚合和聚合关联的欺诈检测. 首先从某待测节点出发, 通过随机游走采样若干条长度不等的链路, 这些链路可能包含共性欺诈模式. 进一步地, 通过对链路根据链路重要程度进行进一步筛选, 选出一批高质量的链路. 然后, 对所有链路进行编码, 得到统一长度、位置关联的嵌入表示. 接着, 进行全局链路信息交互与相似度比对, 挖掘欺诈链路间的共性规律, 以更全面地揭示欺诈行为间的关联模式, 再通过相似度加权之后的各条链路进行聚合作为中心待测节点的特征. 最后, 基于聚合关联的特征, 引入由多层感知机 (MLP) 与判别函数组成的欺诈检测器, 实现节点层面的欺诈检测. 相比其他“局部感知模型”, 本文方法的优势在于拥有更大的“感受野”, 能够通过全局信息交互来自动挖掘隐蔽的欺诈模式, 且有效规避了“过度平滑”的问题. 在金融交易、社交网络和评论网络这 3 类欺诈场景下的多个数据集上的实验结果表明, 相较于最优基准模型, 本文方法在曲线下面积 (AUC) 和平均精度 (AP) 指标上均有显著提升. 此外, 本文方法具备一定的可解释性, 可以为欺诈检测任务挖掘潜在的共性欺诈链路模式, 从而驱动节点学习这些重要的模式并获得更具表现力的表示.

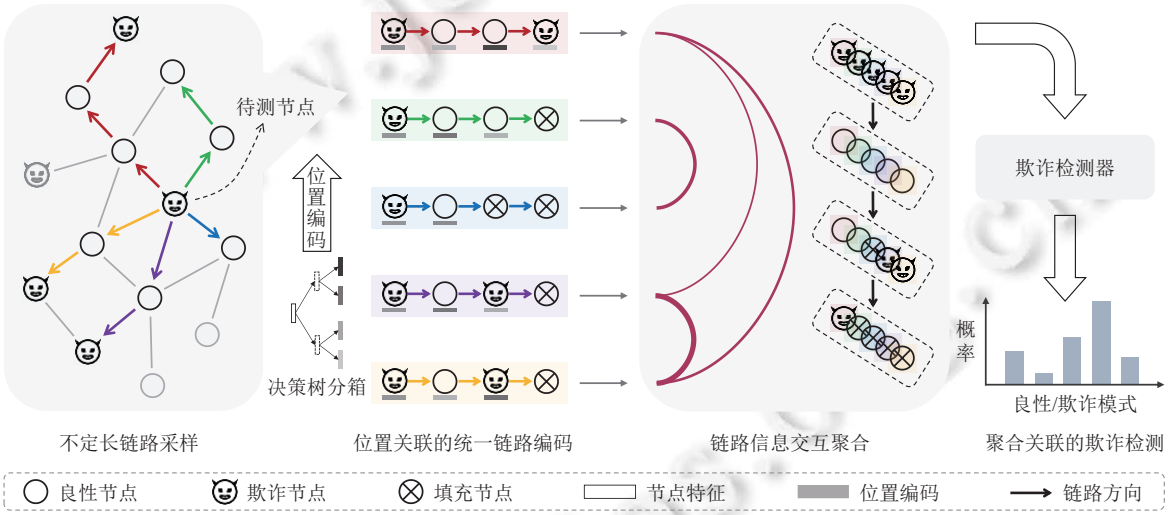


图 1 基于链路聚合的图欺诈检测模型 PA-GNN 框架图

综上所述, 本文所提出的基于链路聚合的图欺诈检测模型 PA-GNN 的主要贡献如下.

(1) 采用长距离链路采样编码、链路信息交互与聚合的方式, 针对欺诈场景下欺诈实体高阶交互、依赖关系错综复杂及高度隐蔽性的特点, 弥补了主流 GNN 算法无法处理长程依赖图结构的不足.

(2) 方法可广泛应用于金融交易、社交网络和评论网络等欺诈场景, 且相较于其他主流欺诈检测方法, 可以有效地实现性能的大幅提升.

(3) 方法具备一定的可解释性, 通过全局模式交互与相似度比对, 挖掘欺诈链路间的共性规律, 可以更全面地揭示欺诈行为间的关联模式.

本文第 1 节介绍图神经网络以及图欺诈检测的相关技术和研究现状. 第 2 节介绍本文构建的基于链路聚合的

图欺诈检测模型 PA-GNN. 第 3 节通过一系列对比实验验证了所提出的 PA-GNN 在不同欺诈场景下的有效性、泛化性与可解释性. 最后在第 4 节总结全文.

1 相关工作

1.1 图神经网络

考虑到样本间存在错综复杂的拓扑结构和依赖关系, 研究人员提出图神经网络 (GNN)[8−12], 通过学习图数据之间的关系, 实现了对非结构化数据的分析和处理, 并将其广泛应用于各类下游任务, 如图分类[24]、链路预测[25]、节点分类[26,27]、内容推荐[28,29]等. GNN 建立在消息传递的基础上, 通过使用神经模块递归聚合邻域信息来更新目标节点的表示. GNN 具有挖掘实体之间深刻的依赖性和相关性的能力, 所以基于 GNN 的欺诈检测成为新的主流方法. 图卷积网络 (GCN)[13]是最早提出的经典 GNN 算法之一. 它通过利用图卷积操作来聚合节点的邻居信息, 从而更新节点的表示. 图采样与聚合 (GraphSAGE)[14]是一种改进的 GCN 算法. 它通过采样节点的邻居子图, 然后聚合各个节点的特征进行更新, 相比于 GCN 具有更好的扩展性. 图注意力网络 (GAT)[15]是另一种经典的 GNN 算法. 它引入了注意力机制来学习中心节点对邻居节点的交互权重, 从而更好地适应图中节点之间的关联, 相比于 GCN 和 GraphSAGE 具有更好的表达能力. 但是无论是 GCN、GraphSAGE, 还是 GAT, 它们都是局部网络, 其缺点是无法处理具有长程依赖的图结构, 因为它们只考虑了一阶邻居节点的信息, 对于全局信息的利用较为有限. 欺诈样本间实体通常具有高阶交互, 且具有错综复杂的依赖关系以及高度隐蔽性, 仅靠局部信息无法准确识别欺诈行为. 因此, 针对这一缺点, 仍然需要进一步研究和改进 GNN 算法.

1.2 图欺诈检测算法

基于图的反欺诈算法[1−5]可以抽象为给定关系网络数据, 从中查找异常的节点. 本节首先回顾基于结构特征和图表示学习的传统图欺诈检测算法, 随后介绍基于谱方法和空间方法的最新检测技术.

传统的基于图的反欺诈算法主要依赖于结构特征. 节点的重要性通常与欺诈风险紧密相关, 因此, 识别网络中的关键节点对于有效的欺诈检测至关重要. 常用的节点重要性评价指标有中心性度量和 PageRank[30]等. 中心性度量又分为度中心性、加权度中心性、介数中心性[31]、接近中心性和特征向量中心性[32], 这些常用的节点重要性评价指标可以帮助了解网络中节点的重要程度. Drezewski 等人[33]聚焦于银行金融交易, 利用上述节点重要性特征表示网络结构, 识别用户在交易网络中的角色, 揭示可疑的洗钱参与者. Wang 等人[34]提出了一种通过检查交易网络中用户账户的 EgoNet 特征来检测异常交易行为的方法, 然后使用 Mahalanobis 距离来衡量用户账户与正常行为的偏离程度. Wang 等人[35]提出一种基于网络图结构的无监督异常检测算法 GBKD-Forest, 提取出入度、边连接以及 EgoNet 这 3 类结构特征, 在 Bagging 方法内随机抽取特征建立 KD-Tree 森林, 隔离出异常样本. 这类方法只能识别与已知欺诈模式相似的欺诈行为, 在识别未知欺诈类型方面存在着局限性, 且通常需要手动设置的中心度/相似度阈值, 无法很好地通用于各种欺诈场景.

为了克服传统方法的局限性, 基于图表示学习的反欺诈方法逐渐受到关注. 这类方法利用图表示学习技术将节点或实体在图中的结构特征映射到低维嵌入空间中, 以便更好地揭示欺诈行为的潜在模式和规律. 经典的图嵌入模型包含 DeepWalk[36]和 Node2Vec[37]等. 使用这些模型可以获得节点嵌入, 然后再利用传统的分类方法对所得到的低维特征数据进行分析, 以实现欺诈检测的目标. DeepWalk[36]是一种流行的图嵌入算法, 旨在学习图中节点的连续表示. 它利用随机游走和语言建模的概念, 将图的结构信息编码为低维向量表示. 首先从图中的每个节点开始生成多条随机游走路径, 将这些路径视为句子. 然后, 它应用自然语言处理中的 skip-gram[38]模型, 根据随机游走中节点的共现关系来学习节点的向量表示. Node2Vec[37]通过引入一种灵活的偏好随机游走策略来扩展 DeepWalk. Node2Vec 不仅执行纯随机游走, 还探索了同质性和结构等效性邻域. 它通过引入两个超参数来控制深度优先和广度优先对图的探索权衡, 捕捉图中编码的结构属性.

近年来, 先进的图欺诈检测技术主要可以分为谱方法和空间方法两类. 在谱方法方面, Chai 等人[16]提出 AMNet, 一种新型的自适应多频图神经网络, 旨在自适应地组合每个节点的多个频率信号, 以解决现有 GNN 仅考虑单一

频率图信号而异常和正常节点则倾向于不同的频段的问题. Tang 等人[17]发现图异常导致光谱能量分布的“右移”现象, 并据此提出了 BWGNN, 利用 Beta 图小波在光谱和空间域中生成带通滤波器, 以更好地捕获图上的异常信息. Gao 等人[18]则关注于 GNN 在盲目平滑邻近节点表示时可能破坏异常判别信息的问题, 设计了一个标签感知边缘指标来计算聚合后相似度得分, 并在此基础上修剪可能的异质性边缘. 在空间方法方面, Liu 等人[19]提出了 CARE-GNN, 通过结合上下文嵌入、邻域信息度量和关系注意力来分别解决欺诈者造成的上下文不一致、特征不一致和关系不一致问题. Liu 等人[20]提出了 PC-GNN, 通过标签平衡采样器和邻域采样器, 解决了传统 GNN 中由于类别不平衡导致的特征稀释问题, 提高了对少数类 (欺诈者) 的检测性能. Peng 等人[21]提出的 RioGNN 则是一种新型图神经网络架构, 结合强化的关系感知邻居选择和标签感知神经相似性度量来改进表示学习并维护关系相关表示. Shi 等人[22]则提出了 H^2-FDetector, 通过引入同质性和异质性连接, 并设计了一种新的信息聚合策略, 分别在原型先验的指导下传播相似和不同的信息, 一定程度上解决了欺诈者隐藏在欺诈图中的问题. 最后, Gao 等人[23]观察到异常和正常节点之间的结构分布偏移程度有所不同, 提出了 GDN 方法, 该方法能够识别并限制关键异常特征以减轻异质邻居的影响, 运用原型向量推断和更新异常特征分布, 同时限制正常节点的其余特征以保持连通性并增强同质邻居的影响.

2 基于链路聚合的图欺诈检测模型

鉴于欺诈场景下欺诈实体间高阶交互的复杂性、依赖关系的错综复杂性及高度隐蔽性的特点, 以及主流图神经网络算法在处理长程依赖图结构方面的局限性, 本文提出了一种基于链路聚合的图欺诈检测模型 PA-GNN. 该模型由 4 个核心部分构成: 不定长链路采样、位置关联的统一链路编码、链路信息交互聚合以及聚合关联的欺诈检测. 对于一张含有欺诈节点的图 (graph), 有以下情况.

在不定长链路采样中, 对于图中给定的一个待测节点, 以其为起点, 随机采样多条链路. 这些链路的长度和经过的节点均具随机性, 以确保采样的多样性和广泛性. 在复杂关系网络图中, 节点的连接模式往往可以由少数关键链路来有效表征. 因此, 根据链路上节点的重要性指标, 筛选出最具价值的 Top-K 条链路. 这些链路在表征节点连接模式方面具有更高的代表性和信息量, 为后续共性欺诈模式的挖掘提供有力支持.

在位置关联的统一链路编码中, 把从不定长链路采样中筛选出的长度各异的链路, 通过统一且高效的方式编码成相同长度的嵌入向量, 以确保后续处理的一致性和有效性. 为增强节点特征在所有节点中的相对位置信息, 引入决策树分箱编码方法, 对节点属性值范围进行分组, 作为位置编码的依据, 并将位置编码融入链路的嵌入向量, 以达到节点位置标记与特征去噪的效果.

在链路信息交互聚合中, 首先利用链路间的自注意力机制[39]实现信息交互, 使链路嵌入之间能够相互传递和接收信息, 进而自动挖掘出链路中存在的共性欺诈模式. 期望通过这种精细设计的信息交互过程, 使得欺诈链路在嵌入空间中呈现出更加紧密的聚类效果, 即欺诈链路之间的相似度更高. 同时, 也希望欺诈链路与良性链路之间的相似度尽可能低, 即它们在嵌入空间中保持较远的距离. 反之, 良性链路之间应表现出较高的相似度, 而与欺诈链路保持明显的区分. 这样的设计有助于在后续的链路分析中更准确地识别和区分欺诈行为与良性行为. 接着执行链路聚合, 将链路层面的信息整合提炼为节点层面的表征. 这一过程不仅保留了链路间的交互信息, 还为中心待测节点赋予了一个富含上下文信息的表示. 这个表示中蕴含了节点在关系网络图中的连接模式、交互行为以及潜在的欺诈或良性特征, 为后续的欺诈检测提供了丰富的信息基础.

最后在聚合关联的欺诈检测中, 引入欺诈检测器, 运用一层 MLP 与判别函数来识别聚合节点表示中隐藏的欺诈模式. 这一步骤为欺诈检测提供了准确的判断和有力的依据, 从而实现对欺诈行为的有效识别和防范.

2.1 问题定义

令 $G=(V,E)$ 为一张关系网络图, $V=\{v_1,v_2,\ldots,v_N\}$, $E=\{e_1,e_2,\ldots,e_M\}$, $X=\{\boldsymbol{x}_{v_1},\boldsymbol{x}_{v_2},\ldots,\boldsymbol{x}_{v_N}\}$ 分别代表节点集、边集和节点特征集, $v_n\in\mathbb{N}$ 代表节点的编号, $\boldsymbol{x}_{v_n}\in\mathbb{R}^D$, D 为节点特征数. $Y=\{y_{v_1},y_{v_2},\ldots,y_{v_N}\}$ 是图 G 的节点真实标签集. 本文聚焦于欺诈场景下的关系网络图, 图中节点分为良性与欺诈两类. 基于图的欺诈检测本质上是节点二分类

问题, $y_{v_n} \in \{0, 1\}$, $y_{v_n} = 1$ 表示节点 v_n 为欺诈节点, 否则为良性节点.

2.2 不定长链路采样

对于给定的一个待测节点, 从其自身出发, 通过随机游走采样 I 条链路组成 $\mathbf{R} \in \mathbb{N}^{I \times J}$, 每条链路 $\mathbf{R}_i \in \mathbb{N}^J$ 以该节点为起点, 共包含 J 个按顺序前后相连的节点. 为了进一步丰富链路长度的多样性, 对链路进行随机覆盖的操作. 定义一个二值矩阵 $\mathbf{M} \in \mathbf{1}^{J \times J}$, $\mathbf{M} = \mathbf{U} - \mathbf{I}$, 其中 $\mathbf{U} \in \mathbf{1}^{J \times J}$ 是全 1 的上三角矩阵, $\mathbf{I} \in \mathbf{1}^{J \times J}$ 是单位矩阵. 从 $\mathbf{M}$ 中随机可重复地采样 I 行, 得到 $\tilde{\mathbf{M}} \in \mathbf{1}^{I \times J}$, 接着, 通过以下操作获得不同长度的链路 $\tilde{\mathbf{R}} \in \mathbb{R}^{I \times J}$:

$$\tilde{\mathbf{R}}_{i,j} = \begin{cases} \mathbf{R}_{i,j}, & \text{if } \tilde{\mathbf{M}}_{i,j} = 0 \\ -1, & \text{if } \tilde{\mathbf{M}}_{i,j} = 1 \end{cases} \tag{1}$$

为了提升采样链路的质量, 需对链路进行进一步筛选. 选用链路所经过节点的度中心性之和作为评估链路重要程度的依据, 从随机采样的 I 条链路中筛选出 Top-K 条最有价值的链路 $\tilde{\mathbf{R}} \in \mathbb{R}^{K \times J}$:

$$\tilde{\mathbf{R}} \in \mathbb{R}^{K \times J} = \text{Top-}K^{\downarrow \text{degree}}(\tilde{\mathbf{R}} \in \mathbb{R}^{I \times J}) \tag{2}$$

2.3 位置关联的统一链路编码

为了方便后续的链路信息交互, 对采样得到的不同长度的链路进行统一编码和降维, 使其长度相同. 同时, 根据链路节点自身特征进行决策树分箱, 并以此为依据设计位置编码, 将其与链路编码融合, 进一步增强特征.

对于第 k 条链路 $\tilde{\mathbf{R}}_k$, 其经过的所有节点的特征组成的集合为 $\{\boldsymbol{x}_{\tilde{\mathbf{R}}_{k,j}}\}_{j=1}^{J}$. 在不定长链路采样中, 不同长度的链路在尾部用标号为−1 的节点进行填充, 标号为−1 的节点对应的特征 $\boldsymbol{x}$ 则用全零填充. 首先将链路节点特征依次进行拼接, 得到链路特征 $\check{\boldsymbol{h}}_k \in \mathbb{R}^{DJ}$:

$$\check{\boldsymbol{h}}_k = \bigcup \left\{ \boldsymbol{x}_{\tilde{\mathbf{R}}_{k,j}} \right\}_{j=1}^{J} \tag{3}$$

其中, $\bigcup$ 是拼接操作, 拼接后的链路特征长度为 $D \cdot J$. 接着, 对链路特征进行线性映射, 得到嵌入表示 $\boldsymbol{h}_k \in \mathbb{R}^C$:

$$\boldsymbol{h}_k = \check{\boldsymbol{h}}_k \cdot \mathbf{W} + \boldsymbol{b} \tag{4}$$

其中, $\mathbf{W} \in \mathbb{R}^{DJ \times C}$ 和 $\boldsymbol{b} \in \mathbb{R}^C$ 分别为可学习的权重和偏置.

为了增强每个节点在所有节点当中的位置信息, 本文受到 Transformer[39]中位置编码的启发, 针对欺诈检测任务的特有属性和需求, 设计了一种独特的位置编码方法. 具体地, 引入决策树分箱编码, 它对节点各属性值的范围进行分组. 对于每个属性, 确定组的数量和每组的范围对于最佳欺诈检测至关重要. 以节点标签作为监督, 采用决策树将节点的每个属性分类到不同的叶节点中. 叶子节点的数量对应组的数量, 分割条件值决定每个组的范围. 利用训练好的每个属性的决策树, 将所有分割条件值按升序排序, 然后用于将整个范围划分为若干个互斥的箱 (bin). 每个属性值会有对应的箱号, 于是得到各个节点的位置编码 $\left\{ \boldsymbol{x}^{\text{pos}}_{\tilde{\mathbf{R}}_{k,j}} \right\}_{j=1}^{J}$, 其中 $\boldsymbol{x}^{\text{pos}}_{\tilde{\mathbf{R}}_{k,j}} \in \mathbb{R}^D$, 再对位置编码进行标准化操作. 在链路中, 标号为−1 的填充节点对应的位置编码 $\boldsymbol{x}^{\text{pos}}$ 同样用全零填充. 随后, 采用与公式 (3) 和 (4) 相同的拼接和线性映射操作, 得到链路位置编码的嵌入表示 $\boldsymbol{h}^{\text{pos}}_k \in \mathbb{R}^{DJ}$:

$$\check{\boldsymbol{h}}^{\text{pos}}_k = \bigcup \left\{ \boldsymbol{x}^{\text{pos}}_{\tilde{\mathbf{R}}_{k,j}} \right\}_{j=1}^{J} \tag{5}$$

$$\boldsymbol{h}^{\text{pos}}_k = \check{\boldsymbol{h}}^{\text{pos}}_k \cdot \mathbf{W}^{\text{pos}} + \boldsymbol{b}^{\text{pos}} \tag{6}$$

其中, $\mathbf{W}^{\text{pos}} \in \mathbb{R}^{DJ \times C}$ 和 $\boldsymbol{b}^{\text{pos}} \in \mathbb{R}^C$ 分别为可学习的权重和偏置.

最后, 将链路特征嵌入与位置编码相加, 得到最终的链路嵌入表示 $\boldsymbol{z}_k \in \mathbb{R}^C$:

$$\boldsymbol{z}_k = \boldsymbol{h}_k + \boldsymbol{h}^{\text{pos}}_k \tag{7}$$

2.4 链路信息交互聚合

在该阶段, 待测节点自身的链路之间会进行信息交互, 以挖掘共性的欺诈模式. 对于待测节点, 首先堆叠该节点出发的所有链路的嵌入, 得到 $\mathbf{Z} \in \mathbb{R}^{K \times C}$:

$$\mathbf{Z} = [\boldsymbol{z}_1; \boldsymbol{z}_2; \ldots; \boldsymbol{z}_K] \tag{8}$$

接着, 实现链路嵌入之间的自注意力机制[39], 得到更新后的嵌入表示 $\hat{\mathbf{Z}} \in \mathbb{R}^{K\times C}$:

$$\mathbf{Q}=\mathbf{Z}\cdot\mathbf{W}^{\text{query}},\ \mathbf{K}=\mathbf{Z}\cdot\mathbf{W}^{\text{key}},\ \mathbf{V}=\mathbf{Z}\cdot\mathbf{W}^{\text{value}} \tag{9}$$

$$\hat{\mathbf{Z}}=Softmax\left(\mathbf{Q}\cdot\mathbf{K}^{\mathrm{T}}\right)\cdot\mathbf{V} \tag{10}$$

其中, $\mathbf{W}^{\{\text{query, key, value}\}} \in \mathbb{R}^{C\times C}$ 为可学习的转换权重矩阵. $Softmax(\mathbf{Q}\cdot\mathbf{K}^{\mathrm{T}})$ 操作可以得到每条链路对自身及其他所有链路的关注度. 随着训练数据量的增加, 模型对于欺诈链路与良性链路模式的认知会更加清晰, 即, 在注意力机制中, 欺诈链路与欺诈链路、良性链路与良性链路之间的关注度/相似度会更高, 欺诈链路与良性链路模式之间的关注度/相似度会更低. 此外, 采样链路通常较长, 涉及高阶邻居, 链路信息交互可视为全局特征融合. 相较于 GCN 等局部网络, 链路信息交互的"感受野"更大, 更适合检测采用各种手段掩盖的欺诈行为.

接下来进行信息聚合操作. 类似于传统 GNN 聚合邻居节点的特征, 本方法聚合的是链路特征. 链路信息交互部分输出从待测节点出发的所有链路的嵌入 $\hat{\mathbf{Z}} \in \mathbb{R}^{K\times C}$, 聚合得到 $\bar{\boldsymbol{z}} \in \mathbb{R}^{C}$:

$$\bar{\boldsymbol{z}}=\bigcap\left\{\hat{\mathbf{Z}}_k\right\}_{k=1}^{K} \tag{11}$$

其中, $\bigcap$ 代表聚合操作. 本文采用平均聚合, 聚合后的特征 $\bar{\boldsymbol{z}}$ 作为更新后的待测节点的嵌入.

2.5 聚合关联的欺诈检测

通过前面提到的链路聚合操作, 将更新后的待测节点的嵌入 $\bar{\boldsymbol{z}}$ 馈送到欺诈检测器中. 欺诈检测器运用一层 MLP 和 *Sigmoid* 判别函数产生一个预测值 $p \in [0, 1]$, 表示待测节点被预测为欺诈的概率:

$$p=Sigmoid(MLP(\bar{\boldsymbol{z}})) \tag{12}$$

图 G 中往往包含数以万计的节点和边, 一次性对所有节点采样链路并输入到图神经网络中需要相当大的显存与计算开销, 这在实际应用中往往是不现实的. 为了在有限的硬件资源下高效地训练大规模的图数据集, 本文采用 minibatch 的方式, 通过将训练数据划分为若干小批量来实现模型的渐进式训练, 即在每次迭代时, 仅从图 G 中随机采样一小批节点, 再从每个节点出发进行不定长链路采样、位置关联的统一链路编码、链路信息交互聚合、聚合关联的欺诈检测. 链路信息交互不需要任何的监督信号, 模型最终输出每个节点为欺诈的概率 p. 节点标签为 y, 构建二元交叉熵损失 L:

$$L=y\cdot\log p+(1-y)\cdot\log(1-p) \tag{13}$$

随后采用反向传播梯度下降优化模型参数.

3 实验分析

3.1 实验数据

本文实验选取金融交易、社交网络和评论网络这 3 类典型的欺诈场景中共 5 个常用公开数据集 (Elliptic[40]、T-Finance[17]、T-Social[17]、YelpChi[41]和 Amazon[42]) 来验证本文提出的反欺诈模型. 表 1 给出了数据集所对应的详细项目信息.

表 1 本文使用的欺诈检测数据集

数据集	欺诈场景	总节点数	欺诈节点数	节点特征数	总边数	平均度数
Elliptic	金融交易	46 564	4 545	93	73 248	1.57
T-Finance	金融交易	39 357	1 804	10	42 445 086	1 078.46
T-Social	社交网络	5 781 065	174 280	10	146 211 016	25.29
YelpChi	评论网络	45 954	6 677	32	7 693 958	167.43
Amazon	评论网络	11 944	821	25	8 796 784	736.50

● Elliptic 数据集[40]来源于真实的比特币交易网络, 是一个由表示比特币交易的节点和表示交易流的边组成的图结构数据集. 节点分为合法、非法和未知这 3 类, 分别代表不同类型的交易实体, 如交易所、矿工、交易活动等. 每个交易实体与 93 个特征相关联. Elliptic 数据集在区块链安全、金融欺诈检测等领域具有极高的应用价值.

通过分析交易网络中的节点和边, 可以识别出如洗钱等潜在的非法活动, 为金融机构和监管部门提供有力的数据支持. 训练集、验证集、测试集的划分比例分别为 45.86%、18.34%、35.80%, 遵循事务实体时间戳, 符合数据集划分的官方建议.

● T-Finance 数据集最早由 Tang 等人[17]发布, 旨在查找金融交易网络中的异常账户. 这些节点是独特的匿名账户, 具有与注册天数、日志活动和交互频率相关的 10 维特征. 图中的边连接有交易记录的两个账户. 如果节点属于欺诈、洗钱和在线赌博等类别, 人类专家会将节点注释为异常. 训练集、验证集、测试集的划分比例为 40%、20%、40%.

● T-Social 数据集同样由 Tang 等人[17]发布, 旨在查找社交网络中的异常账户. 它具有与 T-Finance 相同的节点注释和功能, 并且两个节点如果保持好友关系超过 3 个月则相互连接. T-Social 的规模非常庞大, 拥有 578 万个节点和 1.5 亿条边, 是本文使用的其他数据集的至少 100 倍, 具有很高的挑战性. 训练集、验证集、测试集的划分比例分别为 40%、20%、40%.

● YelpChi 数据集[41]是基于 Yelp (美国最大点评网站, 具有较大的顾客群体和社会影响力) 的一个行为图数据集, 数据集中个体特征与图结构化特征以稀疏矩阵的形式进行存储, 涵盖业务、评论、用户信息等. 该数据集被广泛应用于金融风控、欺诈检测、反洗钱等研究任务中. 训练集、验证集、测试集的划分比例分别为 40%、20%、40%.

● Amazon 数据集[42]是由亚马逊公司提供的大规模电子商务数据集. 与 YelpChi 数据集类似, 这个数据集包含数百万个产品的信息, 涵盖了多个类别和品牌. 它的特点是提供了丰富的商品特征和用户评价, 包含 24 个产品类别下的超过 1.4 亿条评论和产品元数据. 数据集使用乐器下的评论, 并将用户作为图的节点. 该数据集也被广泛应用于欺诈检测任务的验证. 训练集、验证集、测试集的划分比例分别为 40%、20%、40%.

3.2 基准模型

本文将所提方法与非 GNN 模型、传统 GNN 模型、谱方法以及空间方法进行比较.

非 GNN 模型仅依据节点自身特征进行欺诈检测, 未考虑节点间复杂的连接关系. 本文选用经典的多层感知机 (MLP)[43]和随机森林 (RF)[44]模型作为没有图信息情况下的基础模型.

传统 GNN 模型不仅考虑节点自身特征, 还会纳入邻居节点的信息, 通常通过聚合一阶邻居的信息来感知局部连接模式以获取更多有用的信息. 本文选用 GCN[13]和 GAT[15]两个经典模型, 它们的区别在于对邻居节点权重的计算方式, GCN 综合考虑入度和出度, 而 GAT 使用注意力机制.

另外, 本文方法还与目前先进的谱方法和空间方法进行比较. 谱方法包括 AMNet[16]、BWGNN[17]和 GHRN[18], 空间方法包括 CARE-GNN[19]、PC-GNN[20]、RioGNN[21]、H^2-FDetector[22]和 GDN[23]. 谱方法涉及利用图信号的不同频率成分, 可通过自适应组合多频信号、分析光谱能量分布和利用标签感知边缘指标等手段, 设计特定机制以优化或增强对图上欺诈信息的捕获与处理能力. 空间方法则主要关注节点间的空间关系, 通过利用标签感知相似性度量、强化学习选择邻居、跨关系聚合、引入异构网络和新型信息聚合策略等手段, 提升图神经网络对欺诈行为和关系的检测与识别能力. 这些方法在之前的图欺诈检测工作中表现出了很好的性能.

3.3 评价指标

考虑到欺诈数据集中正负样本不平衡, 评估指标应同时兼顾精确率 (precision) 和召回率 (recall), 本文采用 AUC 和 AP 两个评价指标来评估反欺诈模型的性能.

AUC (area under the curve) 通常指的是 AUC-ROC (area under the receiver operating characteristic curve), 即 ROC 曲线下面积. ROC 曲线是通过改变分类阈值, 绘制出真阳性率 (true positive rate, TPR) 与假阳性率 (false positive rate, FPR) 之间的关系曲线.

AP (average precision), 即平均精度, 是基于 P-R (precision-recall) 曲线计算得出. P-R 曲线通过改变分类阈值, 绘制出精确率 (precision) 与召回率 (recall) 之间的关系. AP 是 P-R 曲线下面的面积, 反映了模型在不同阈值下的平均精度. 欺诈检测数据集的正负样本通常是极度不平衡的, 由于欺诈行为 (正样本) 往往远少于良性行为 (负样本), 一般模型会对良性行为产生高召回率, 此时 AUC 就会产生过于乐观的分数. P-R 曲线更能准确地反映模型对

欺诈行为的识别能力, 获得误报会对 AP 值产生显著影响. 因此, 以 AP 作为评价指标, 在正负样本不平衡的情况下具有显著优势, 它能够更直接地衡量模型在识别少数但重要的欺诈行为上的性能.

AUC 和 AP 值的范围均是 0–1, 值越接近 1 表示模型性能越好. 本文的所有数值结果以百分数 (%) 形式呈现, 范围为 0–100%.

3.4 实验设置

本文的模型参数设置见表 2, n_path (*I*) 代表初始采样链路个数, k_path (*K*) 代表筛选链路个数, l_path (*J*) 代表链路最大长度, d_path (*C*) 代表链路编码长度, n_bin (*B*) 代表位置编码中的决策树分箱个数, n_attn 和 n_head 分别代表链路自注意力层数和注意力头个数.

表 2 模型参数设置

数据集	n_path (*I*)	k_path (*K*)	l_path (*J*)	d_path (*C*)	n_bin (*B*)	n_attn	n_head
Elliptic	100	100	5	256	4	2	1
T-Finance	200	150	5	128	4	2	1
T-Social	200	100	30	64	32	2	8
YelpChi	200	200	20	256	32	2	8
Amazon	60	30	20	64	4	2	1

在训练参数设置方面, 所有实验均使用 Adam 优化器, 初始学习率为 1E–3, 权重衰减为 5E–4, 梯度裁剪范数为 1.0. 采用 ReduceLROnPlateau 自适应的学习率调整策略, 在模型性能指标 (验证集上的损失或准确率) 停止改善时自动降低学习率为原来的 0.5 倍. 控制每组对比实验采用完全相同的 batch_size, 从头开始训练 50 轮次, 并采用自动混合精度 (AMP) 策略来加快训练速度. 所有数值结果是 3 个不同随机种子下测试集上的平均值. 本文的实验环境为 NVIDIA GeForce RTX 3090 GPU 和 Intel(R) Xeon(R) Silver 4210R CPU @ 2.40 GHz.

3.5 实验结果与分析

本文提出的反欺诈模型 PA-GNN 在金融交易欺诈数据集 Elliptic、T-Finance, 社交网络欺诈数据集 T-Social, 以及评论网络欺诈数据集 YelpChi、Amazon 上, 与非 GNN 模型 MLP、RF, 传统 GNN 模型 GCN、GAT, 谱方法 AMNet、BWGNN、GHRN, 以及空间方法 CARE-GNN、PC-GNN、RioGNN、H^2-FDetector、GDN 进行了全面的性能对比, 以验证本文方法的有效性与先进性. 如表 3 所示, 由实验结果可得以下结论.

表 3 本文方法 (PA-GNN) 与基准模型性能比较 (%)

类别	方法	金融交易				社交网络		评论网络			
		Elliptic		T-Finance		T-Social		YelpChi		Amazon	
		AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP
非GNN	MLP	80.08	61.71	91.31	65.95	67.78	5.28	68.27	29.18	92.62	66.23
	RF	82.69	65.20	94.42	81.16	80.38	42.50	89.47	<u>69.77</u>	94.47	83.02
传统GNN	GCN	85.28	38.27	92.34	78.47	91.72	56.35	60.99	24.89	80.89	29.68
	GAT	76.54	27.15	92.89	78.94	93.40	52.26	73.98	34.86	94.50	85.16
谱方法	AMNet	88.89	<u>69.49</u>	95.72	83.41	—	—	83.57	52.85	97.90	87.87
	BWGNN	89.63	48.39	96.29	85.50	97.77	79.53	<u>90.64</u>	68.05	<u>98.08</u>	88.52
	GHRN	<u>89.95</u>	55.20	<u>96.76</u>	83.78	<u>98.26</u>	<u>87.66</u>	90.59	66.68	97.15	<u>89.84</u>
空间方法	CARE-GNN	87.84	37.19	89.95	61.84	78.32	41.15	83.99	52.96	96.96	85.64
	PC-GNN	88.52	34.70	89.90	65.11	89.29	51.40	81.75	49.06	95.46	84.85
	RioGNN	86.35	29.06	91.30	62.60	81.74	17.62	85.72	56.45	96.91	87.62
	H^2-FDetector	63.18	10.53	—	—	—	—	89.88	57.48	96.05	84.94
	GDN	88.67	65.21	95.82	<u>85.69</u>	88.43	52.34	90.42	67.38	97.26	86.78
链路聚合	PA-GNN	**91.08 +1.13**	**75.52 +6.03**	**97.24 +0.48**	**89.58 +3.89**	**99.60 +1.34**	**96.01 +8.35**	**91.86 +1.22**	**73.58 +3.81**	**98.12 +0.04**	**92.43 +2.59**

注: “加粗”和“下划线”分别表示“最好”和“次好”; “—”表示“显存溢出”

(1) 总体上, 本文所提出的方法在所有数据的两种评估指标 (AUC、AP) 上都显示出良好的结果, 其中 AP 指标在不同数据集上实现了约 2%–9% 的增幅. 在规模最大、挑战性最高的 T-Social 数据集上, 本文方法在 AP 指标上能比目前最先进的方法高出将近 9%. 其他已有的方法的表现非常不稳定, 如 GCN 在 T-Finance 和 T-Social 数据集上表现较为良好, 但在其余 3 个数据集上的表现就不尽人意, RioGNN 方法在 T-Finance、YelpChi 和 Amazon 数据集上表现出色, 但在 Elliptic 和 T-Social 数据集上表现糟糕. 相比之下, 本文方法在各个场景中的表现更加稳定, 充分体现了其强大的泛化性能.

(2) 目前的图欺诈检测方法能够实现较高的 AUC 值, 但 AP 值普遍较低. 原因在于欺诈场景数据集中良性样本远多于欺诈样本. 当数据集中正负样本数量极不均衡时, 模型可能更容易被训练成偏向多数类 (良性样本), 从而导致在 P-R 曲线上表现不佳, 即 AP 值较低. 然而, ROC 曲线对类别分布不敏感, 因此 AUC 值可能仍然很高. 本文方法在提升 AP 值方面非常明显, 说明在保证高召回率的同时, 也实现了高精确率, 有效克服了欺诈场景下正负样本不均衡带来的训练挑战. 这可能得益于本文方法的“全局感知”能力, 可检测到更多隐蔽的欺诈行为.

(3) 在处理具有复杂关联特性的图数据时, 传统非 GNN 方法, 尤其是 MLP, 难以达到理想的异常检测效果, 其 AUC 和 AP 值普遍较低. 原因在于这些方法仅能利用节点自身的特征, 无法捕捉节点间的关联关系. 欺诈者往往伪装成正常者, 其特征可能与正常者无异或差别甚小, 因此仅凭节点属性区分欺诈与良性行为极具挑战性. 传统 GNN 方法, 如 GCN 和 GAT, 在欺诈检测数据集上亦存在局限. 这些方法通过聚合邻居节点信息来更新节点表示, 从而部分捕捉图的结构特征. 然而, 在处理涉及高阶交互的异常时, 这些方法往往力不从心. 欺诈行为常涉及复杂的交互模式和隐藏的关系链, 可能跨越多个节点和边, 形成高阶图结构. 传统 GNN 方法主要关注一阶邻居信息, 难以有效捕捉这类高阶、非局部的交互模式. 例如, 在社交网络和评论网络数据集中, 尽管传统 GNN 方法的 AUC 和 AP 值较高, 但仍低于一些专为异常检测设计的 GNN 变体. 谱方法和空间方法在本次实验中表现出优势, 主要因为它们针对欺诈场景的特性进行设计. 谱方法认为欺诈节点和良性节点处于不同频域, 于是在消息聚合时尽量减少两者间信号的相互干扰, 相比传统 GNN 方法, 可有效避免节点信息的“平滑”现象. 空间方法则考虑邻居节点中可能同时包含欺诈节点和良性节点, 因此应有区分性地进行邻居选择与聚合. 然而, 所有这些方法本质上仅考虑低阶邻居信息, “感受野”有限, 只能在一定范围内有效识别欺诈行为. 面对更加隐蔽复杂、涉及长程链路关联的欺诈行为, 这类“局部感知”的方法将无能为力. 相比之下, 本文方法的感知范围更广, 考虑了欺诈行为的长程关联模式, 能够发现更复杂隐蔽的欺诈行为, 因此在性能上能超越目前先进的图欺诈检测技术.

3.6 消融实验

本节验证本文方法中的各个部件对于最终性能的影响, 主要涉及的问题如下.

(1) 不定长链路采样中, 采样不定长链路是否能获得比采样等长链路更好的效果?

(2) 不定长链路采样中, 从随机采样的链路中进一步挑选 Top-K 条高质量链路是否带来了更好的性能?

(3) 位置关联的统一链路编码中, 链路节点的位置编码是否必要?

(4) 链路信息交互聚合中, 链路间的自注意力模式交互是否必要?

针对上述问题, 本文在 T-Social 数据集上进行了多种组合的消融实验, 实验结果如表 4 所示.

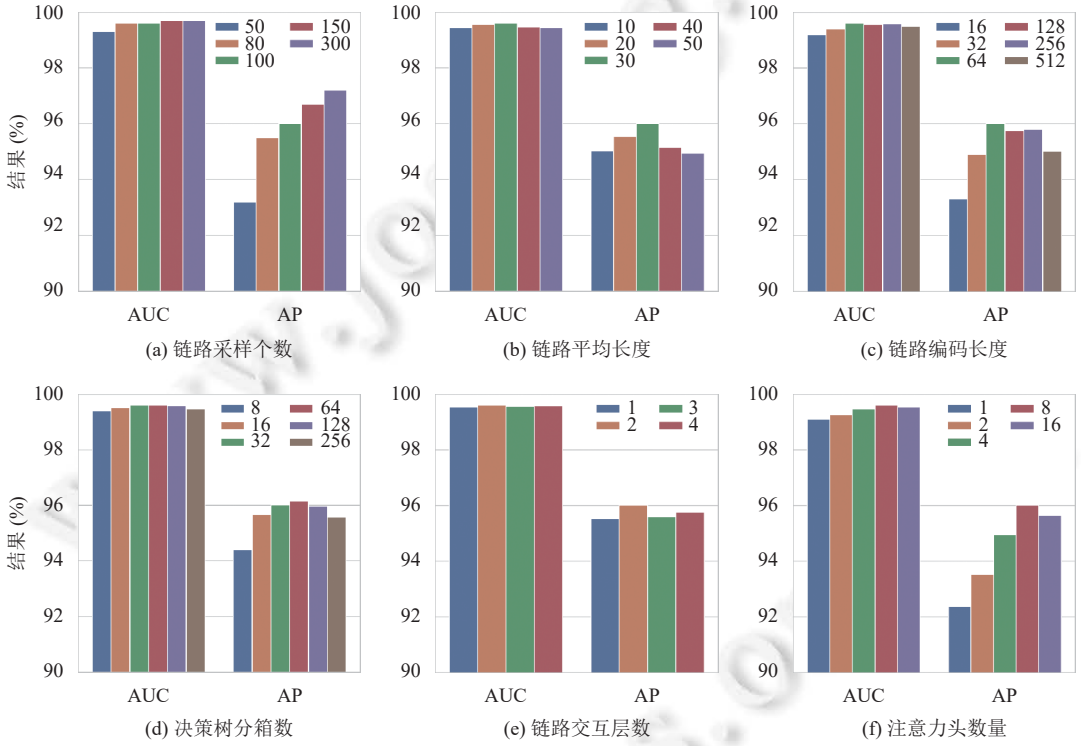
表 4 T-Social 数据集上的消融实验结果 (%)

不定长链路	Top-K链路筛选	位置编码	链路信息交互	AUC	AP
√	√	√	√	**99.60**	**96.01**
×	√	√	√	99.53	95.62
√	×	√	√	99.52	95.22
√	√	×	√	99.01	91.06
√	√	√	×	93.35	56.38
√	×	×	√	98.71	88.77
×	√	×	√	98.85	89.72
×	×	√	√	99.51	95.20
×	×	×	√	98.88	89.52
×	×	×	×	88.91	43.63

实验显示, 所有部件均能单独发挥作用. 采样不定长的链路可以提升性能, 因为欺诈链路在长度上也具有多样性. 对初步采样的链路, 根据链路经过节点的总体度中心性进行进一步筛选, 可以获得更多高质量的链路, 有助于后续的链路模式挖掘, 并提升最终性能. 位置编码起到较为重要的作用, 移除节点特征中融入的位置编码将造成约−5% 的 AP 下降. 这可能是因为位置编码标识了特征所处的范围, 起到了去噪的效果. 链路信息交互聚合是本方法的重要步骤, 如果链路间缺乏交互而直接进行聚合, 模型对欺诈行为的检测性能将显著下降, 因为模型未能充分利用所有链路提取共性的欺诈链路模式, 单纯的链路聚合本质也只是一种特殊的局部信息感知. 此外, 本节所涉及部件的任意组合均能实现性能的进一步提升, 并在使用所有部件时达到最优效果.

3.7 模型参数影响分析

本节充分探讨本文方法受模型参数的影响状况, 具体涉及不定长链路采样中的链路采样个数和链路平均长度, 位置关联的统一链路编码中的链路编码长度和决策树分箱数, 以及链路信息交互聚合中的链路交互层数和注意力头数量. 在 T-Social 数据集上, 模型参数对 AUC 和 AP 指标的影响分析结果如图 2 所示.



(a) 链路采样个数 (b) 链路平均长度 (c) 链路编码长度

(d) 决策树分箱数 (e) 链路交互层数 (f) 注意力头数量

图 2 T-Social 数据集上的模型参数影响分析结果

实验结果显示, 在不定长链路采样方面, 链路采样个数的增加意味着参与模式交互的链路数增加, 从而使模型能够一次接触更多的欺诈模式, 提高共性模式挖掘的效果, 但也会相应地增加计算开销. 随着链路平均长度的增加, 模型性能通常先上升后下降. 这可能是因为较短的链路只能实现低阶的模式交互, 无法处理欺诈模式的复杂关系和高阶交互, 而过长的链路则可能导致信息冗余, 且链路长度越长, 链路编码后的信息损失越严重. 在位置关联的统一链路编码方面, 链路编码长度存在一个最优值, 链路编码过短, 会造成信息损失严重, 链路编码过长则会导致模型过拟合. 同样, 决策树分箱数对于位置编码的表现也存在一个最优值. 箱数过少会导致位置编码在特征之间缺乏区分度, 而箱数过多则会使位置编码过于细粒度, 无法有效地去除特征噪声. 在链路信息交互聚合方面, 随着链路交互层数的增加, 模型性能会逐渐达到饱和, 无法进一步提升. 这与 GCN、GAT 等传统图神经网络不同, 本文模型的链路交互层数增加不会使节点特征“平滑”导致性能下降, 而是自然存在一个饱和点. 另外, 注意力头数量

增多会带来更好的模型性能, 但注意力头数量与链路编码长度的比值过大会增加计算负担, 同时也可能会导致性能下降.

3.8 欺诈链路模式关联分析

由于隐私保护需求, 当前公开的图欺诈数据集仅包含脱敏后的节点特征, 原始详细特征因隐私限制无法获取. 因此, 本文对所提出欺诈检测方案的可解释性研究聚焦于欺诈链路模式的关联分析.

为了探究不同欺诈链路模式之间的潜在关联, 本文选用 T-Social 数据集, 并选取了 5 类欺诈链路模式, 分别为: 欺诈-良性、欺诈-欺诈、欺诈-欺诈-良性-欺诈、欺诈-良性-欺诈和欺诈-良性-欺诈-良性-欺诈. 每类链路模式下均采样 10 条链路, 总共组成 50 条链路. 经过链路编码后, 这些链路被送入链路信息交互聚合模块, 根据公式 (10) 中的 $Softmax(\mathbf{Q}\cdot\mathbf{K}^{\mathrm{T}})$ 计算链路之间的相关性得分. 结果可视化如图 3 所示. 热图的横纵坐标均代表参与自注意力交互的链路的索引, 其中纵坐标是 query 的索引, 横坐标是 key 的索引. 在自注意力机制中, key 是 query 的副本. 如颜色条所指示的, 链路间的相关性得分介于 0–1 之间, 得分越高的区域越亮. 观察发现, 同一种链路模式下的所有链路之间存在密切的关联. 同时, 相似的链路模式之间也可能存在共性关联.

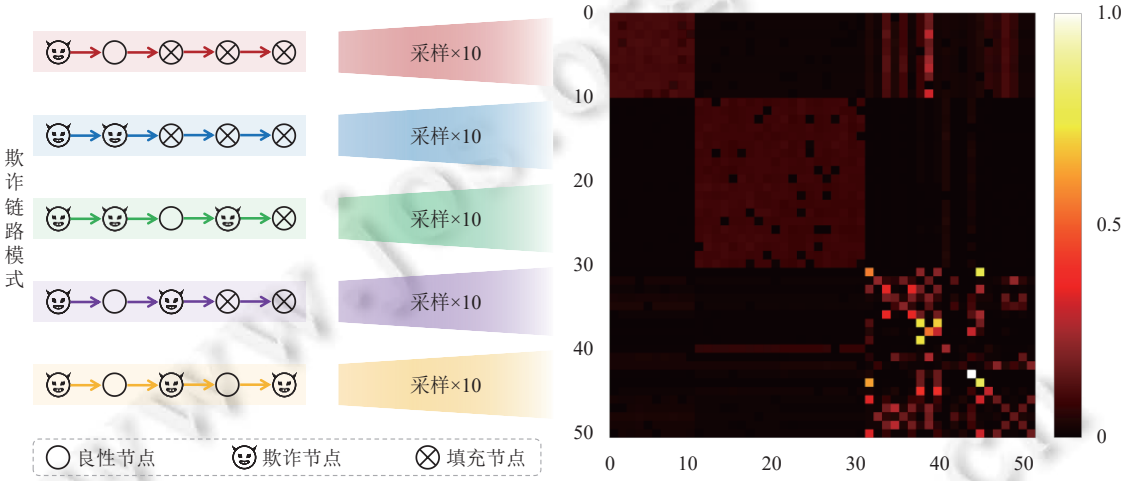


图 3 T-Social 数据集上的欺诈链路模式相关性得分

(1) “欺诈-欺诈”与“欺诈-欺诈-良性-欺诈”链路模式之间存在较高的关注度, 可能原因在于这两种模式都涉及连续的欺诈节点, 表明部分欺诈行为在网络中可能倾向于形成集群或连续的模式.

(2) “欺诈-良性-欺诈”与“欺诈-良性-欺诈-良性-欺诈”链路模式之间存在较高的关注度, 可能原因在于这两种模式都展示了欺诈行为在网络中通过良性节点进行传播或隐藏的潜在策略. “欺诈-良性-欺诈”模式表明欺诈节点可能利用良性节点作为跳板, 进行欺诈行为的扩散或转移, 从而增加欺诈活动的复杂性和隐蔽性. 而“欺诈-良性-欺诈-良性-欺诈”模式则进一步揭示了欺诈行为可能在网络中形成更为复杂的传播链路, 通过多个良性节点的桥接, 构建出更加难以察觉的欺诈网络结构.

(3) “欺诈-良性”链路模式对于“欺诈-良性-欺诈”和“欺诈-良性-欺诈-良性-欺诈”链路模式存在较高关注度, 可能原因在于“欺诈-良性”模式在某种程度上代表了欺诈行为的起始和存在, 是更复杂欺诈模式的基础. 因此, 当分析更复杂的欺诈模式时, 如“欺诈-良性-欺诈”和“欺诈-良性-欺诈-良性-欺诈”, 自然会关注到它们是否由基础的“欺诈-良性”模式发展而来.

(4) 与 (3) 相反, “欺诈-良性-欺诈”和“欺诈-良性-欺诈-良性-欺诈”链路模式对“欺诈-良性”链路模式几乎没有关注, 可能原因在于这两种模式已经包含了比“欺诈-良性”模式更复杂的信息, 即欺诈行为如何通过良性节点进行传播和隐藏. 这些信息已经超越了基础“欺诈-良性”模式的范畴, 因此, 在分析这两种模式时, 可能不会特别关注它们是否包含基础的“欺诈-良性”模式, 而是更侧重于它们自身的复杂性和欺诈行为的传播策略.

以上欺诈链路模式关联分析表明, 本文所提出的方法具备一定的可解释性, 可以通过链路信息交互发现潜在的共性欺诈模式, 相似的欺诈模式之间会获得更大的模式交互系数, 从而驱动节点学习这些重要的共性欺诈模式并获得更具表现力的表示.

4 总 结

本文关注了信息空间风险外溢至人类社会和物理空间所带来的严峻挑战, 尤其是欺诈行为的日益增多及其对社会造成的负面影响, 并聚焦于图欺诈检测领域, 指出了传统欺诈检测方法的局限性. 考虑到欺诈场景下欺诈实体间高阶交互的复杂性、依赖关系的错综复杂性及高度隐蔽性的特点, 本文提出了一种基于链路聚合的图欺诈检测模型 PA-GNN. 该模型通过长距离链路采样编码、链路信息交互与聚合的方式, 有效弥补了主流 GNN 算法在处理长程依赖图结构方面的不足. 实验结果表明, 该模型在金融交易、社交网络和评论网络这 3 类欺诈场景下均实现了显著的性能提升, 并具备一定的可解释性, 能够更好地揭示不同欺诈行为间的关联模式. 本文为图欺诈检测领域提供了新的思路和方法, 有望在实际应用中发挥重要作用, 为构建更加安全、可靠的信息空间环境做出积极贡献, 为欺诈行为的预防和打击提供有力支持.

References

[1] 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]

[2] Kim H, Lee BS, Shin WY, Lim S. Graph anomaly detection with graph neural networks: Current status and challenges. IEEE Access, 2022, 10: 111820–111829. [doi: 10.1109/ACCESS.2022.3211306]

[3] Liu HL, Liu YX, Xu JY, Chen SH, Qiao L. Research progress in application of graph anomaly detection in financial anti-fraud. Computer Engineering and Applications, 2022, 58(22): 41–53 (in Chinese with English abstract). [doi: 10.3778/j.issn.1002-8331.2203-0233]

[4] Ren J, Xia F, Lee I, Noori Hoshyar A, Aggarwal C. Graph learning for anomaly analytics: Algorithms, applications, and challenges. ACM Trans. on Intelligent Systems and Technology, 2023, 14(2): 28. [doi: 10.1145/3570906]

[5] Chen JL, Chen X, Jing YJ, Wang SY. Survey of application of graph neural network in anomaly detection. Computer Engineering and Applications, 2024, 60(13): 51–65 (in Chinese with English abstract). [doi: 10.3778/j.issn.1002-8331.2310-0234]

[6] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521(7553): 436–444. [doi: 10.1038/nature14539]

[7] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge: MIT Press, 2016.

[8] 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]

[9] Zhou J, Cui GQ, Hu SD, Zhang ZY, Yang C, Liu ZY, Wang LF, Li CC, Sun MS. Graph neural networks: A review of methods and applications. AI Open, 2020, 1: 57–81. [doi: 10.1016/j.aiopen.2021.01.001]

[10] Zhou Y, Zheng HX, Huang X, Hao SF, Li DA, Zhao JM. Graph neural networks: Taxonomy, advances, and trends. ACM Trans. on Intelligent Systems and Technology, 2022, 13(1): 15. [doi: 10.1145/3495161]

[11] Liu J, Shang XQ, Song LY, Tan YC. Progress of graph neural networks on complex graph mining. Ruan Jian Xue Bao/Journal of Software, 2022, 33(10): 3582–3618 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6626.htm [doi: 10.13328/j.cnki.jos.006626]

[12] Song LY, Ma ZY, Li ZH, Shang XQ. Review on temporal graph neural networks for financial risk prediction. Ruan Jian Xue Bao/Journal of Software, 2024, 35(8): 3897–3922 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/7087.htm [doi: 10.13328/j.cnki.jos.007087]

[13] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: Proc. of the 5th Int'l Conf. on Learning Representations. Toulon: OpenReview.net, 2017. 1–14.

[14] Hamilton WL, Ying RT, Leskovec J. Inductive representation learning on large graphs. In: Proc. of the 31st Conf. on Neural Information Processing Systems. Long Beach: NIPS, 2017. 1–19.

[15] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. In: Proc. of the 7th Int'l Conf. on Learning Representations. Vancouver: OpenReview.net, 2018. 1–12.

[16] Chai ZW, You SQ, Yang Y, Pu SL, Xu JR, Cai HY, Jiang WH. Can abnormality be detected by graph neural networks? In: Proc. of the 31st Int'l Joint Conf. on Artificial Intelligence. Vienna: IJCAI, 2022. 1945–1951. [doi: 10.24963/ijcai.2022/270]

[17] 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. Baltimore: PMLR, 2022. 21076–21089.

[18] 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]

[19] 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]

[20] Liu Y, Ao X, Qin ZD, Chi JF, Feng JH, Yang H, He Q. Pick and choose: A GNN-based imbalanced learning approach for fraud detection. In: Proc. of the 2021 Web Conf. Ljubljana: ACM, 2021. 3168–3177. [doi: 10.1145/3442381.3449989]

[21] Peng H, Zhang RT, Dou YT, Yang RY, Zhang JY, Yu PS. Reinforced neighborhood selection guided multi-relational graph neural networks. ACM Trans. on Information Systems, 2022, 40(4): 69. [doi: 10.1145/3490181]

[22] Shi FZ, Cao YN, Shang YM, Zhou YC, Zhou C, Wu J. H^2-FDetector: A GNN-based fraud detector with homophilic and heterophilic connections. In: Proc. of the 2022 ACM Web Conf. ACM, 2022. 1486–1494. [doi: 10.1145/3485447.3512195]

[23] Gao Y, Wang X, He XN, Liu ZG, Feng HM, Zhang YD. Alleviating structural distribution shift in graph anomaly detection. In: Proc. of the 13th ACM Int'l Conf. on Web Search and Data Mining. Singapore: ACM, 2023. 357–365. [doi: 10.1145/3539597.3570377]

[24] Wang ZH, Shen HW, Cao Q, Cheng XQ. Survey on graph classification. Ruan Jian Xue Bao/Journal of Software, 2022, 33(1): 171–192 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6323.htm [doi: 10.13328/j.cnki.jos.006323]

[25] Kumar A, Singh SS, Singh K, Biswas B. Link prediction techniques, applications, and performance: A survey. Physica A: Statistical Mechanics and its Applications, 2020, 553: 124289. [doi: 10.1016/j.physa.2020.124289]

[26] Zhang LY, Sun HH, Sun YF, and Shi BB. Review of node classification methods based on graph convolutional neural networks. Computer Science, 2024, 51(4): 95–105 (in Chinese with English abstract). [doi: 10.11896/jsjkx.230600071]

[27] Ye DK, Yan ZQ, Xu JY, Chen F, Weng W. U-GCN based node classification algorithm. Microelectronics & Computer, 2025, 42(5): 9–17 (in Chinese with English abstract). [doi: 10.19304/J.ISSN1000-7180.2024.0272]

[28] Du XY, Chen Z, Xiang XG. Explainable tag-aware recommendation based on disentangled graph neural network. Ruan Jian Xue Bao/Journal of Software, 2023, 34(12): 5670–5685 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6754.htm [doi: 10.13328/j.cnki.jos.006754]

[29] Qian ZS, Zhao C, Yu QY, Li DM. Information fusion recommendation approach combining attention CNN and GNN. Ruan Jian Xue Bao/Journal of Software, 2023, 34(5): 2317–2336 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6405.htm [doi: 10.13328/j.cnki.jos.006405]

[30] Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: Bringing order to the Web. Stanford InfoLab, 1999. https://gwern.net/doc/technology/google/1998-page.pdf

[31] Freeman LC. A set of measures of centrality based on betweenness. Sociometry, 1977, 40(1): 35–41. [doi: 10.2307/3033543]

[32] Bonacich P, Lloyd P. Eigenvector-like measures of centrality for asymmetric relations. Social Networks, 2001, 23(3): 191–201. [doi: 10.1016/S0378-8733(01)00038-7]

[33] Drężewski R, Sepielak J, Filipkowski W. The application of social network analysis algorithms in a system supporting money laundering detection. Information Sciences, 2015, 295: 18–32. [doi: 10.1016/j.ins.2014.10.015]

[34] Wang Y, Wang LM, Yang J. Egonet based anomaly detection in E-bank transaction networks. In: Proc. of the 3rd Int'l Conf. on Material Engineering and Advanced Manufacturing Technology. Shanghai: IOP, 2020. 012038. [doi: 10.1088/1757-899X/715/1/012038]

[35] Wang K, Chen DW. Graph structure based anomaly behavior detection. In: Proc. of the 2nd Int'l Conf. on Computer Engineering, Information Science & Application Technology. Wuhan: ICCIA, 2016. 531–538. [doi: 10.2991/iccia-17.2017.90]

[36] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM, 2014. 701–710. [doi: 10.1145/2623330.2623732]

[37] Grover A, Leskovec J. Node2Vec: Scalable feature learning for networks. In: Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016. 855–864. [doi: 10.1145/2939672.2939754]

[38] Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. In: Proc. of the 1st Int'l Conf. on Learning Representations. Scottsdale, 2013.

[39] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.

[40] Weber M, Domeniconi G, Chen J, Weidele DKI, Bellei C, Robinson T, Leiserson CE. Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv:1908.02591, 2019.

[41] Rayana S, Akoglu L. Collective opinion spam detection: Bridging review networks and metadata. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Sydney: ACM, 2015. 985–994. [doi: 10.1145/2783258.2783370]

[42] McAuley J, Leskovec J. From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews. In: Proc. of the 2013 Int'l Conf. on World Wide Web. Rio de Janeiro: ACM, 2013. 897–908. [doi: 10.1145/2488388.2488466]

[43] Hinton GE. Connectionist learning procedures. In: Kodratoff Y, Michalski RS. Machine Learning: An Artificial Intelligence Approach, Volume III. San Mateo: Morgan Kaufmann, 1990. 555–610. [DOI: 10.1016/B978-0-08-051055-2.50029-8]

[44] Breiman L. Random forests. Machine Learning, 2001, 45(1): 5–32. [doi: 10.1023/A:1010933404324]

附中文参考文献

[3] 刘华玲, 刘雅欣, 许珺怡, 陈尚辉, 乔梁. 图异常检测在金融反欺诈中的应用研究进展. 计算机工程与应用, 2022, 58(22): 41–53. [doi: 10.3778/j.issn.1002-8331.2203-0233]

[5] 陈佳乐, 陈旭, 景永俊, 王叔洋. 图神经网络在异常检测中的应用综述. 计算机工程与应用, 2024, 60(13): 51–65. [doi: 10.3778/j.issn.1002-8331.2310-0234]

[11] 刘杰, 尚学群, 宋凌云, 谭亚聪. 图神经网络在复杂图挖掘上的研究进展. 软件学报, 2022, 33(10): 3582–3618. http://www.jos.org.cn/1000-9825/6626.htm [doi: 10.13328/j.cnki.jos.006626]

[12] 宋凌云, 马卓源, 李战怀, 尚学群. 面向金融风险预测的时序图神经网络综述. 软件学报, 2024, 35(8): 3897–3922. http://www.jos.org.cn/1000-9825/7087.htm [doi: 10.13328/j.cnki.jos.007087]

[24] 王兆慧, 沈华伟, 曹婧, 程学旗. 图分类研究综述. 软件学报, 2022, 33(1): 171–192. http://www.jos.org.cn/1000-9825/6323.htm [doi: 10.13328/j.cnki.jos.006323]

[26] 张丽英, 孙海航, 孙玉发, 石兵波. 基于图卷积神经网络的节点分类方法研究综述. 计算机科学, 2024, 51(4): 95–105. [doi: 10.11896/jsjkx.230600071]

[27] 叶岱昆, 颜钟棋, 徐嘉忆, 陈奋, 翁伟. 基于 U-GCN 的节点分类算法. 微电子学与计算机, 2025, 42(5): 9–17. [doi: 10.19304/J.ISSN1000-7180.2024.0272]

[28] 杜晓宇, 陈正, 项欣光. 基于解耦图神经网络的可解释标签感知推荐算法. 软件学报, 2023, 34(12): 5670–5685. http://www.jos.org.cn/1000-9825/6754.htm [doi: 10.13328/j.cnki.jos.006754]

[29] 钱忠胜, 赵畅, 俞情媛, 李端明. 结合注意力 CNN 与 GNN 的信息融合推荐方法. 软件学报, 2023, 34(5): 2317–2336. http://www.jos.org.cn/1000-9825/6405.htm [doi: 10.13328/j.cnki.jos.006405]

作者简介

邱天, 博士生, 主要研究领域为机器学习, 表征学习, 基于图的数据挖掘.

贾凌翔, 博士, 研究员, CCF 专业会员, 主要研究领域为图模式挖掘, 人工智能辅助药物研发.

高杨, 博士, 研究员, CCF 专业会员, 主要研究领域为图学习, 时间序列挖掘, 强化学习.

冯尊磊, 博士, 副教授, 博士生导师, CCF 专业会员, 主要研究领域为医学图像分析, 模型优化, 基于图的数据挖掘.

高艺, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为物联网软件, 智能边缘计算.

宋明黎, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为计算机视觉, 模式识别, 人工智能, 机器学习.