图算法在网络关联数据挖掘中的应用探析

李聪瑞

（商洛学院，陕西 商洛 726000）

摘 要 大数据时代的网络数据呈现规模化、复杂化、关联化特征，传统数据挖掘算法难以适配非线性、多维度的网络关联数据处理需求。图算法依托图结构的节点与边关联逻辑，可精准刻画网络数据的关联关系与拓扑特征，成为网络关联数据挖掘的核心技术。本文立足于网络关联数据挖掘的核心需求，阐述图算法的核心原理与适配优势，分析主流图算法的应用场景，梳理技术应用中的现存问题，并提出针对性优化路径，以期为网络关联数据挖掘技术的落地升级提供理论参考。

关键词 图算法；网络数据；关联数据挖掘；拓扑分析；数据挖掘应用

中图分类号：TP311.13；TP393 **文献标志码**：A **DOI**：10.3969/j.issn.2097-3365.2026.30.004

0 引言

互联网与信息技术的持续迭代，推动社交网络、物联网、电商平台等场景的数据体量持续激增，各类数据不再是孤立个体，而是形成相互交织的网络化关联体系。关联数据挖掘是从海量杂乱数据中提取隐性关联价值的核心手段，常规挖掘方法仅能处理结构化线性数据。图算法基于拓扑结构解析数据关联，能够突破传统技术局限，直观呈现数据节点的连接关系，当前已广泛应用于各类网络数据挖掘场景，系统探析图算法的实际应用价值具有重要的现实意义。

1 图算法在网络关联数据挖掘中的应用理论基础

1.1 图算法核心原理与分类特征

图算法是以图论为理论支撑，依托节点、边、权重等核心要素构建运算逻辑的数据分析算法，核心原理是将所有网络数据要素进行图结构建模，通过标准化遍历、计算、推演操作，解析图网络的拓扑结构、关联强度、节点权重等核心指标，以此量化数据之间的关联关系。根据应用功能与运算逻辑的差异，主流图算法可划分为四大核心类型，分别为遍历类算法、社区发现类算法、路径分析类算法与权重排序类算法。各类算法具备独立的技术优势与适配场景，能够针对性解决网络关联数据挖掘中的不同问题，同时各类算法可组合应用，大幅提升复杂网络数据的挖掘精度与分析效率。

1.2 网络关联数据挖掘核心内涵

网络关联数据是区别于传统结构化数据的新型数据形态，核心特征是以节点代表数据个体，以连接边代表数据之间的依存、互动、关联关系，整体呈现非线性、网络化、动态化的结构特点[1]。网络关联数据挖掘的核心目标是通过规范化技术手段清洗海量冗余网络数据，挖掘数据表层信息下隐藏的关联规则、拓扑规律与潜在价值，实现无序网络数据的结构化转化。该技术广泛服务于舆情分析、智能推荐、风险识别、网络拓扑优化等多个领域，是大数据价值落地的关键技术支撑。相较于传统数据挖掘聚焦数据属性特征的分析模式，网络关联数据挖掘更加侧重数据个体之间的联动关系分析，更加适配当下复杂网络环境的数据处理需求。

1.3 图算法适配网络关联数据挖掘的核心优势

传统数据挖掘算法多基于数据表、线性序列完成数据统计分析，仅能够识别显性数据关联，无法捕捉跨维度、跨层级的隐性网络关联，同时面对动态更新、结构复杂的大规模网络数据，存在运算效率低、分析误差大、适配性差等诸多短板[2]。图算法的核心优势在于与网络关联数据的结构高度契合，可完整复刻数据的网络化分布特征，直观呈现数据节点的层级关系与联动逻辑。同时，图算法支持动态数据更新分析，能够实时适配网络数据的增减与关联变化，且具备极强的兼容性与拓展性，可适配不同行业、不同场景的复杂网络关联数据挖掘工作，有效弥补传统挖掘技术的短板。

2 主流图算法在网络关联数据挖掘中的具体应用

2.1 遍历类图算法实现网络数据全域挖掘

遍历类图算法是网络关联数据挖掘的基础算法类型，核心代表为深度优先算法与广度优先算法，主要

作者简介：李聪瑞（2005-），男，本科，研究方向：计算机科学与技术。

功能是完成网络拓扑结构的全域检索与数据梳理。在实际应用过程中，深度优先算法以纵向检索为核心逻辑，从单一核心节点出发持续延伸遍历深度，直至完成整条关联链路的全部数据挖掘，适用于挖掘深度层级较深的纵向数据关联关系[3]。广度优先算法以横向扩散为检索逻辑，从核心节点逐层向外拓展遍历范围，全面覆盖相邻节点的关联数据，能够实现网络局部数据的全覆盖挖掘。两类遍历算法相互配合，可完成社交网络关系梳理、设备网络节点排查、舆情传播链路追踪等基础挖掘工作，为后续深度数据解析提供完整的数据基础。

2.2 社区发现算法实现网络数据集群分类

社区发现算法是挖掘网络数据集群关联特征的核心算法，典型代表为 Louvain 算法、GN 算法，核心应用价值是从复杂全域网络中拆分出关联性更强的子网络集群。复杂网络数据整体杂乱无序，不同数据节点的关联强度存在显著差异，社区发现算法可通过计算节点之间的关联密度与耦合程度，将关联紧密的节点划分为同一社区集群，将关联薄弱的节点进行剥离区分。在实际场景中，该算法可应用于社交圈层划分、用户群体聚类、产业关联集群分析、网络舆情阵营分类等场景。通过精准的集群分类，能够帮助工作人员快速区分数据层级，精准定位核心数据群体，提升关联数据挖掘的针对性与有效性。

2.3 路径分析算法挖掘数据关联传导规律

路径分析算法以最短路径算法、关键路径算法为核心，聚焦网络数据节点之间的关联传导路径与联动逻辑挖掘。网络关联数据的价值不仅体现在静态连接关系中，更蕴含在动态的关联传导过程中，数据节点的交互传播、影响传导、联动变化均依托固定路径完成。路径分析算法可精准计算任意两个数据节点之间的最优关联路径，识别网络中的核心传导链路与中转节点，同时能够排查网络数据关联中的断裂节点与异常路径。该算法广泛应用于电商物流路径优化、网络风险传导溯源、信息传播路径分析、交通网络数据联动挖掘等场景，能够精准捕捉数据动态关联的内在规律，实现从静态数据梳理到动态规律挖掘的升级[4]。

2.4 权重排序算法定位网络核心数据节点

权重排序算法以 PageRank 算法、度中心性算法、介数中心性算法为核心，主要用于量化网络数据节点的重要程度，实现核心关联数据的精准筛选。大规模网络关联数据中，不同节点的影响力与价值存在巨大差异，普通挖掘方式无法区分数据优先级。权重排序算法通过计算节点的连接数量、中转频次、辐射范围等核心指标，对所有数据节点进行权重排序，精准定位网络中的核心枢纽节点、高价值关联数据与关键影响因子。在实际应用中，该算法可用于舆情核心传播账号筛选、金融风险核心节点识别、电商核心流量用户挖掘、网络设备核心枢纽定位等工作，有效提升海量网络数据中高价值信息的提取效率。

3 图算法在网络关联数据挖掘应用中的现存问题

3.1 大规模复杂网络数据运算效率不足

当前各行业网络数据呈现爆发式增长趋势，网络节点数量持续增多，数据关联链路更加复杂，形成超大规模的复杂网络体系。传统基础图算法的运算逻辑较为固化，面对海量节点与高密度关联网络时，会出现运算冗余、迭代速度缓慢、算力消耗过大等问题。部分高精度图算法虽然能够保证挖掘精度，但运算流程繁杂，无法适配实时性、动态化的网络数据挖掘需求。尤其在物联网、全域社交网络等超大数据场景中，算法运算滞后性问题突出，难以实现网络关联数据的实时挖掘与动态分析，制约技术的实际应用效果。

3.2 动态异构网络数据挖掘精度有限

现阶段，多数常规图算法适用于结构统一、属性单一的同构网络数据挖掘，而实际应用场景中的网络关联数据多为动态异构数据。不同数据节点的属性特征、关联权重、更新频率存在显著差异，同时网络数据时刻处于新增、删除、关联变更的动态变化状态。常规图算法无法精准适配异构数据的差异化特征，难以精准识别动态变化的隐性关联关系，容易出现数据关联漏判、误判等问题。此外，部分算法对噪声数据的过滤能力较弱，海量冗余噪声数据会干扰节点权重与关联强度的计算结果，进一步降低网络关联数据挖掘的精准度。

3.3 算法组合应用适配性与通用性较差

单一图算法仅能完成网络关联数据的单项维度挖掘，无法满足复杂场景下多维度、全方位的数据挖掘需求，实际工作中需要依托多种算法组合完成综合分析。当前行业内尚未形成标准化的算法组合应用体系，工作人员多依托经验进行算法搭配，存在算法组合不合理、功能重叠、适配性不足等问题。不同图算法的运算逻辑、数据口径、评价标准存在差异，组合应用过程中容易出现数据衔接断层、运算结果冲突等问题[5]。同时，各类图算法的场景通用性较差，针对金融、医疗、工业等垂直领域的专属网络数据缺乏定制化算法模型，技术落地适配性不足。

3.4 图算法模型智能化融合程度较低

大数据与人工智能技术深度融合的行业背景下，数据挖掘工作逐步向智能化、自动化方向升级，但当前多数图算法仍停留在传统数理运算层面，与智能算法的融合深度不足。传统图算法仅能完成数据关联的量化分析与规律梳理，无法实现关联趋势预测、风险预判、智能分类等高阶功能。同时，现有图算法模型的自主学习能力较弱，无法根据持续更新的网络数据迭代优化运算逻辑，需要人工持续调整参数适配数据变化。智能化技术融合的缺失，导致图算法仅能实现基础的数据挖掘功能，难以挖掘网络关联数据的深层潜在价值，无法适配智能化数据应用的行业发展趋势。

4 图算法在网络关联数据挖掘中的优化应用路径

4.1 优化算法运算逻辑，提升大规模数据处理效率

针对大规模复杂网络数据运算效率不足的问题，需从算法底层运算逻辑入手完成迭代优化。技术人员可对传统图算法的冗余运算流程进行精简，简化非核心迭代步骤，引入分布式运算框架，将大规模网络数据拆解为多个子模块并行运算，大幅提升算法处理海量数据的速度。同时可搭建增量运算机制，针对网络中新增、变更的数据节点进行针对性运算，无需重复全域遍历数据，有效减少算力消耗。此外，还可根据网络数据规模分级适配算法模型，小型简单网络采用基础算法提升效率，大型复杂网络采用优化后的并行算法，实现运算效率与挖掘精度的双向平衡，适配大规模动态网络数据挖掘需求。

4.2 优化异构数据适配机制，提升数据挖掘精度

为解决动态异构网络数据挖掘精度不足的问题，需构建适配差异化数据特征的算法优化体系。首先，可在算法前置环节增设智能数据清洗模块，精准过滤网络中的冗余数据、噪声数据与无效节点，减少异常数据对挖掘结果的干扰。其次，对传统图算法的权重计算逻辑进行优化，针对不同属性、不同类型的数据节点设置差异化权重系数，适配异构网络的数据特征。最后，搭建动态更新适配机制，实时捕捉网络数据的关联变化与节点迭代状态，动态调整算法运算参数，精准识别动态网络中的隐性关联关系，有效降低数据漏判误判概率，全面提升复杂异构网络关联数据的挖掘精准度。

4.3 构建标准化算法组合应用体系，强化场景适配

行业需结合不同场景的挖掘需求，构建规范化、标准化的图算法组合应用体系，规避算法搭配的随意性问题。通过梳理各类图算法的功能优势、适用范围与运算特征，搭建算法匹配数据库，根据不同行业、不同场景的网络数据结构与挖掘目标，自动匹配最优算法组合模式。针对算法组合的数据衔接问题，统一各类算法的数据口径与运算标准，搭建数据互通接口，实现多算法运算数据的无缝衔接，避免结果冲突与数据断层。同时，针对金融、工业、舆情等垂直领域，研发定制化图算法模型，结合行业数据特征优化运算逻辑，全面提升算法的场景适配性与通用性。

4.4 推动图算法与智能技术深度融合，实现智能挖掘

依托人工智能、机器学习等前沿技术，推动图算法向智能化、高阶化方向升级，拓展数据挖掘的深度与广度。将图算法与机器学习算法深度融合，让算法模型具备自主学习能力，可依托海量历史网络数据持续迭代运算逻辑，自主优化参数设置，无需人工频繁干预调整。利用深度学习技术优化图神经网络模型，在传统拓扑结构分析的基础上，新增数据特征学习、关联趋势预测、异常关联识别等功能，实现网络关联数据从规律挖掘到趋势预判的升级。同时，搭建智能化数据挖掘平台，集成算法运算、数据分析、结果可视化、智能研判等功能，实现网络关联数据挖掘的自动化与智能化，充分释放网络数据的深层价值。

5 结束语

网络关联数据挖掘是大数据时代数据价值转化的核心手段，图算法凭借独特的拓扑分析优势，成为复杂网络关联数据挖掘的核心技术。本文系统梳理图算法的理论基础与主流应用场景，剖析当前技术应用中效率不足精度有限适配性较差等问题，提出针对性的算法优化与体系完善路径。未来需持续推动图算法的技术迭代与智能融合，完善标准化应用体系，进一步提升复杂网络关联数据的挖掘质量，让图算法更好地服务于各行业大数据分析与智能化发展。

参考文献：

[1] 鲍雨，肖明坤．算力网络技术在天翼视联的应用研究 [J]. 江苏通信，2025,41(03):19-22,45.
[2] 许郑．多元异构网络安全数据可视化融合研究 [J]. 信息与电脑（理论版），2024,36(22):77-79.
[3] 马健兵，徐池．基于论文语义与引用网络的技术发展趋势分析 [J]. 软件导刊，2024,23(11):100-106.
[4] 孙婷婷．巡查机器人路径规划算法与应用综述 [J]. 物联网技术，2024,14(09):127-130.
[5] 任诚，刘静，王宇，等．物联网中虚拟网络功能转发图算法设计 [J]. 计算机科学与探索，2025,19(06):1632-1639.