

生物医学领域基于知识图谱的知识发现方法综述^{*}

梅佳月 孙 佳 侯跃芳

(中国医科大学健康管理学院 沈阳 110122)

〔摘要〕 目的/意义 综述生物医学领域基于知识图谱的知识发现方法，为科研人员提供参考。方法/过程 通过系统检索及阅读分析相关文献，总结基于知识图谱的知识发现方法，对比各类知识发现方法的优势和不足；指出现阶段生物医学领域基于知识图谱的知识发现方法的局限性和挑战，并提出建议和展望。结果/结论 未来研究应提升知识发现结果的可解释性，构建有效的结果评估框架，建立规范化、多领域专家协同的知识发现流程，提高知识发现研究的质量和效率。

〔关键词〕 生物医学；知识图谱；知识发现

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2024.01.008

A Review of the Knowledge Discovery Methods Based on Knowledge Graph in the Biomedical Field

MEI Jiayue, SUN Jia, HOU Yuefang

School of Health Management, China Medical University, Shenyang 110122, China

〔Abstract〕 **Purpose/Significance** To review the knowledge discovery methods based on knowledge graph in the biomedical field, and to provide references for researchers. **Methods/Process** The paper summarizes the knowledge discovery methods based on knowledge graph by systematically searching and analyzing the relevant literatures, and compares the advantages and disadvantages of various knowledge discovery methods. It points out that the limitations and challenges of knowledge discovery methods based on knowledge graph in the biomedical field, and puts forward suggestions and prospects. **Result/Conclusion** In the future research, it is suggested to improve the interpretability of knowledge discovery results, build an effective result evaluation framework, and establish a standardized knowledge discovery process with multi-domain experts' collaboration, so as to improve the quality and efficiency of knowledge discovery research.

〔Keywords〕 biomedicine; knowledge graph; knowledge discovery

1 引言

随着科学技术的飞速发展，生物医学领域不断进步，大量科研文献随之产生。截至 2023 年 4 月，PubMed 包含的生物医学文献已超过 3 500 万篇^[1]，其中包含大量未被发现的有价值的知识信息。如何依据日益增加的现有知识数据进行隐含知识发现，

〔修回日期〕 2023-07-23

〔作者简介〕 梅佳月，硕士研究生；通信作者：侯跃芳，博士，教授，硕士生导师。

〔基金项目〕 辽宁省教育厅科学研究经费项目（人文社科类基础研究项目）（项目编号：JCRW2020005）。

是当前生物医学领域数据挖掘的重要研究内容。

知识图谱本质是语义网知识库,由实体和实体间的关系构成知识网络,形成直观且易于理解的知识结构框架^[2]。知识图谱能够对大量多源异构的医学信息进行有序化组织,有效关联和管理数据,从而提高海量医学数据的利用价值^[3]。在生物医学领域,知识图谱在临床决策支持、文献可视化分析、智能问答、智能检索等方面已得到广泛应用^[4]。

知识发现是从大量数据中提取出有用新知识的过程^[5]。利用知识图谱进行知识发现可以从海量数据中挖掘深层次的潜在知识关联,为生物医学领域科研人员提供建议和参考。本文综述了基于知识图谱的知识发现方法,总结分析各方法的优势和不足,并针对现阶段生物医学领域知识发现方法存在的局限性提出建议和展望。

2 基于知识图谱的知识发现方法

医学知识图谱聚焦生物医学领域,将大量复杂的生物医学数据以网络形式存储与组织,实现大规模医学数据的有效管理。知识实体之间不仅可以通过关系直接关联,还可以通过多个知识实体和关系间接关联起来,进而产生深层次的隐含知识,知识发现能够从知识图谱中挖掘出这种隐含的新知识。目前,生物医学领域基于知识图谱的知识发现方法主要包括知识推理、路径发现、链路预测和基于神经网络的发现方法。

2.1 知识推理

2.1.1 基于规则的知识推理 根据技术手段可将知识推理分为基于规则的知识推理和基于人工智能的知识推理。Swanson D R^[6]于 1986 年首次提出了一种 ABC 模型,即当两类不存在直接相关关系的知识实体 A 和 C 同时与另一个知识实体 B 相关时,通过推理可得 A 与 C 之间存在潜在关联。知识推理与 ABC 模型同理,是指领域专家根据实体类型和语义关系类型,结合专业知识,定义知识推理规则,构建规则库,实现知识推理。本文定义的推理规则之一:如果已知药物 X 抑制靶点 Y,靶点 Y 促进肿瘤

Z,那么可以认为药物 X 可能治疗肿瘤 Z。

早期经典的 ABC 模型在数据量较小、实体关联较为简单清晰时取得了较好效果。WEEBER M 等^[7]构建了“沙利度胺(药物) - 白细胞介素 - 12 (免疫途径) - 疾病”规则来推断沙利度胺的治疗应用。陈延雪等^[8]人工构建推理规则,挖掘突发事件知识图谱中的隐含关系。利用 ABC 模型得到的知识发现结果准确性较高,可理解性较强,适用于小型数据集,但是推理发现的范围不够广,结果存在一定局限性。人工构建推理规则也存在成本较高、不适用于动态变化的知识图谱、迁移性较差等缺陷^[9]。

2.1.2 基于人工智能的知识推理 除了领域专家人工构建推理规则外,还可以通过机器学习等方法自动从知识图谱中挖掘归纳规则和特征,完成推理。该方法克服了人工定义规则带来的成本高等缺点,在小规模知识图谱中表现出较好推理效果^[10]。Schoenmackers S 等^[11]提出一阶规则学习算法,将知识图谱中的全部关系分别形成集合,以此作为特征来预测实体之间是否存在关系。Galárraga L 等^[12]提出一种关联规则挖掘算法 (association rule mining under incomplete evidence, AMIE),依次学习预测每种关系的规则,并利用这些规则进行预测,实现知识发现。付子轩等^[13]构建了急腹症知识图谱,采用“AMIE + 随机游走”算法实现更加准确有效的知识推理进而补全知识图谱。Noor A 等^[14]构建了一个基于规则的框架算法,利用该框架探索潜在的药物相互作用机制。

2.2 路径发现

2.2.1 基于规则的路径发现 为了解决 ABC 发现模型的局限性,Wilkowski B 等^[15]提出了 AB_nC 框架,其中 $B_n = (B_1, B_2, \dots, B_m)$ 表示 A 和 C 中的多个中间实体。与 AB_nC 模型同理,路径发现从复杂的知识网络中发现有价值的键路,但不同的是,路径发现可以挖掘较为复杂的实体间关系。Xian Y 等^[16]定义了 k-hop 多跳路径,认为 $Path_k = \{e_0 \xleftarrow{r_1} e_1 \xleftarrow{r_2} e_2 \cdots \xleftarrow{r_k} e_k\}$ 表示从实体 e_0 到实体 e_k 是由 k 个关系连接的 k + 1 个实体,通过路径能

够寻找可能存在的实体关联。胡正银等^[17]利用 Neo4j 的多深度关系节点查询,通过自定义路径长度,挖掘出若干条知识路径,发现“疫苗 (vaccines)”和“胎盘生长因子 (placental growth factor)”除了能直接通过“恶性肿瘤 (malignant neoplasms)”进行关联外,还可以通过“白血病, T 细胞 (leukemia, T-cell)”“癌症治疗 (cancer treatment)”等知识实体,通过“治疗 (TREATS)”“共现 (cooccurrence)”“使用 (USES)”“相关 (ASSOCIATED_ WITH)”等关系词进行关联,并发现两条路径具有一定的启发意义,见图 1。

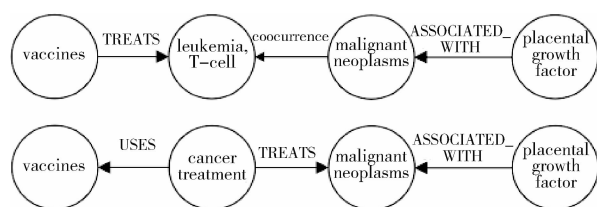


图 1 关联路径示例

除此之外,还有学者运用路径发现方法进行知识发现。Schutte D 等^[18]提出了“药物补充剂-基因-药物”“药物补充剂-基因 1-功能-基因 2-药物”等路径,用于发现药物-补充剂相互作用。Han B S 等^[19]将“神经乳糖酰胺”作为路径的头节点,将“动脉硬化”作为路径的尾节点,寻找基因、药物或细胞等中间实体,并通过排序挑选出最有可能具有生物学意义的相关实体。路径发现方法可以挖掘深层次知识,结果较为全面,同时也存在路径较长会导致可解释性较差的缺陷,因此选择合适的路径长度是路径发现方法的关键。

2.2.2 基于人工智能的路径发现 对于长路径的知识发现,有专家利用人工智能方法进行研究。路径排序算法 (path ranking algorithm, PRA)^[20]是一种典型的自动挖掘路径规则算法。该算法将两个实体之间的多条序列路径作为特征,输入 logistic 回归模型实现推理^[21]。Zhao D 等^[22]提出了一种基于关系路径特征嵌入的卷积神经网络,运用 PRA 从知识图谱中提取疾病-药物关系路径特征,利用经过训练的模型挖掘新的治疗药物。Syafiandini A F 等^[23]利用 PRA 挖掘苍耳子化合物与糖尿病之间的有效路

径,得到治疗糖尿病的潜在药物。为提高 PRA 路径搜索的效率和知识发现效果,还有学者提出耦合路径排序算法 (coupled path ranking algorithm, CPRA)^[24]、增强链接预测 (enhance link prediction, ELP)^[25]等改进算法。

2.3 链路预测

链路预测通过各种算法计算两个知识实体之间的紧密程度和相似性,进而预测其产生关联的可能性^[26]。链路预测是一种基于知识图谱进行知识发现的重要方法。常见的用于知识发现的链路预测指标可以分为以下 3 种^[27]。

2.3.1 基于局部信息的相似性指标 基于局部信息的相似性指标一般通过节点局部信息计算得出。这类指标算法比较简单,适合大规模网络应用,但是计算涉及信息量有限、预测精度较低。常见的基于局部信息的相似性指标包括共同邻居指标 (common neighbors, CN)^[27]、含权的共同邻居指标 (weighted CN, WCN)^[28]、Adamic Adar 指标 (AA)^[29]、含权的 AA 指标 (weighted AA, WAA)^[28]等。其中,CN、AA 等指标只考虑节点相似性,而 WCN、WAA 等含权指标在此基础上还考虑了边的权重,在预测性能方面更有优势。Andrej K 等^[26]通过使用 CN、AA 等指标对医学主题词表 (Medical Subject Headings, MeSH) 网络进行链路预测,发现 AA 指标性能最佳。马代川等^[30]分析药物-蛋白相互作用网络结构特征,运用链路预测实现药物重定位。

2.3.2 基于路径的相似性指标 基于路径的相似性指标主要有 3 个,分别是局部路径指标 (local path, LP)、Katz 指标和 LHN-II^[31]。基于局部路径的相似性指标是在 CN 的基础上提出来的,CN 指标考虑两个节点之间的二阶路径,而 LP 考虑三阶路径^[32]。Katz 指标^[33]相当于考虑全部路径的 LP 指标。LHN-II^[34]的原理是如果 x 和 y 两个节点中任一邻居节点 z 与另一节点相似,那么节点 x 和节点 y 相似。Vural H 等^[35]开发了基于 Katz 指标的预测模型,预测长链非编码核糖核酸 (ribonucleic acid, RNA) 与环境因素之间的潜在关系,预测结果较好。袁华兵等^[36]提出了一种基于多关系路径的链路

预测方法,并将其与 Jaccard、CN 等 8 种链路预测算法进行对比分析,获得更为精确的结果。

2.3.3 基于随机游走的相似性指标 随机游走实际上是一个马尔可夫过程,认为随机游走粒子每一步都可以移动到任一邻居节点,且移动到每一个节点的概率相等^[37]。基于随机游走的相似性指标包括基于全局随机游走和局部随机游走,全局随机游走指标包括平均通勤时间 (average commute time, ACT)^[38]、重启随机游走 (random walk with restart, RWR)^[39]、SimRank^[40]等,局部随机游走指标包括局部随机游走 (local random work, LRW)^[41]、有叠加效应的局部随机游走 (superposed random work, SRW)^[41]等。Li X 等^[42]基于 SimRank 指标提出一种预测模型,以结直肠癌为例预测微小 RNAs (microRNAs, miRNAs) 与疾病之间的关联,结果表明该模型效果较好。Che Y 等^[43]基于 RWR 指标提出了一种 CN-RWR 指标,预测冠心病临床用药的联合使用,结果证明该指标具有较强预测能力。

2.4 基于神经网络的发现方法

神经网络是通过大量数据训练来求解某个问题的算法。知识图谱中基于神经网络的知识发现能够对复杂非线性关系进行建模,深入学习语义三元组的结构特征和语义信息,实现实体和关系的有效预测。目前在基于知识图谱的知识发现中,基于循环神经网络 (recurrent neural network, RNN)、卷积神经网络 (convolutional neural network, CNN) 和图神经网络 (graph neural network, GNN) 的方法应用较为广泛。

2.4.1 基于 RNN 的发现方法 RNN 是一种基于循环结构提取知识序列特征的模型,通过输入知识路径结构或实体文本描述进行预测。陆洪钰^[44]基于 RNN 模型提出一种神经网络推理模型 (time embedded neural network reasoning model, TM-NNR),能够适应知识图谱中不断更新的动态知识推理以及缓解模型可解释性差等问题。Li D 等^[45]提出一种基于知识的循环神经网络 (knowledge-based RNN, KBRNN) 模型算法,从脑瘫知识图谱中提取诊断知识,具有一定通用性,同时在性能方面更具有竞

争力。

2.4.2 基于 CNN 的发现方法 通过卷积操作获取知识局部特征来预测图谱中缺失的关系。Yan J 等^[46]提出一种结合知识图谱和深度学习的加权 CNN-LSTM 模型,该模型大大提高了水质预测效果。刘禹琪^[47]利用基于深度学习的卷积神经网络模型 ConvE (convolutional 2D knowledge graph embeddings) 对中医名方知识图谱中药物与疾病之间的隐含关系进行预测,实现知识图谱补全。

2.4.3 基于 GNN 的发现方法 将图谱中节点和关系的图形结构和节点内容信息作为输入进行训练,从知识图谱中提取拓扑结构特征进行预测,实现知识发现。该方法能够同时考虑知识图谱的语义信息和结构信息,逐渐成为热门研究方法^[48]。黄紫阳^[49]利用节点类型感知的多关系图神经网络学习知识图谱中的药物特征,挖掘图谱中复杂的语义关联模式,实现了药物-药物相互作用预测。Dai G 等^[50]在 GNN 基础上设计了一种多关系图注意力网络模型来优化知识图谱补全工作,该模型考虑了不同节点的权重影响,优化了网络结构,显著提高预测性能。

3 基于知识图谱的知识发现方法对比

上述各种知识发现方法中,知识推理规则的路径较短,发现结果的准确性和可解释性均存在一定优势。路径发现方法挖掘较长路径,能够发现更深层次的隐含知识,同时路径越长其准确性和可解释性就越差。链路预测简单、高效,但这类方法往往容易忽略节点之间的语义关系。基于神经网络的发现方法具有强大的特征捕捉能力,效果较好,但由于整个过程由算法训练得出,仍存在解释性不足等问题。随着科学技术的发展和进步,越来越多学者运用深度学习等方法对知识图谱进行发现和挖掘,其中不乏深度学习与传统推理相结合的研究。总体来看,利用人工智能的知识发现方法往往具有更高效率,在处理大规模知识图谱时更有优势,但仍然存在可解释性较差等不足,见表 1。

表 1 基于知识图谱的知识发现方法对比

知识发现方法	具体方法	优势	不足	代表文献
知识推理	基于规则的知识推理	结果准确性较高，可理解性较强	依赖领域专家的专业知识，推理规则难以归纳	Syafiandini A F 等 ^[22]
	基于人工智能的知识推理	自动挖掘推理规则，节省人工成本	规则学习搜索空间大，效率低；发现深度有限，适用于小型数据集	Noor A 等 ^[14]
路径发现	基于规则的路径发现	能够挖掘深层次的隐含知识	路径越长可解释性越差	胡正银等 ^[17]
	基于人工智能的路径发现	自动挖掘推理路径，为知识发现节省时间和人力成本	难以处理关系稀疏的数据	Bakal G 等 ^[51]
链路预测	基于局部信息的链路预测	计算简单，具有较好的普适性	预测精确度有限；忽略网络结构，效果相对一般	Andrej K 等 ^[26]
	基于路径的链路预测	利用更多的网络信息，效果相对较好	忽略网络结构	袁华兵等 ^[36]
	基于随机游走的链路预测	发现结果的精确度高	无法准确预测边的语义信息	Dasgupta S 等 ^[52]
基于神经网络的发现方法	RNN、CNN、GNN	具有强大的特征捕捉能力；减少人工参与所带来的知识偏见和限制	神经网络结构相对复杂，模型难以训练；可解释性较差	黄紫阳 ^[49]

4 讨论

基于知识图谱的知识发现能够全面、深入地对知识信息进行挖掘，发现有价值的潜在信息，为科研人员提供线索和参考。本文聚焦生物医学领域，阐述基于知识图谱的知识发现方法研究进展，对比分析此类方法的优势和不足。目前此类方法仍在不断革新，知识发现效果也趋于更加准确与高效。针对现有方法的局限性提出以下建议。

4.1 提高知识发现结果的可解释性

基于 ABC 模型构建的知识推理规则清晰且容易解读，而后来的 AB_nC 模型路径较长，构建推理规则存在一定困难，导致知识发现结果的可解释性较差。链路预测往往只能判断两个节点的相似性，而不能给出具体语义关系。深度学习的整个过程由算法训练得出，因此仍然存在解释性差的缺陷。未来应继续思考如何提升知识发现过程及结果的可解释性。

4.2 探寻更加有效的结果评估方法

目前知识发现结果的常见评估方法有基于 ABC

模型的闭合式验证、基于证据的评估方法^[53]、时间片划分法^[54]和专家评估法^[55]等。本课题组基于 ABC 模型，通过探寻两个知识实体之间的共同相关节点进行闭合式验证^[56]，这种评估方法与基于证据的评估方法均比较容易实现，但需要耗费大量人力且通用性较差；时间片划分法只能评估已经发现的知识，无法对新发现的知识进行验证；专家评估法的评估效果仍然存在人工影响。以上评估方法均存在一定不足，构建一个科学可靠的金标准数据集是解决知识发现评估问题的关键。

4.3 降低专家参与的执行难度

生物医学领域专业性较强，因此在知识发现过程中，除了数据专家和情报专家的分析外，还需要生物医学领域专家提供专业知识。然而不同领域专家对领域知识发现的了解程度不同。未来需要确立一套多领域专家协同的统一知识发现流程，为各领域专家建立规范可量化的执行标准，提高知识发现研究的准确性。

5 结语

随着生物医学知识数量的不断增长，大规模知

识图谱将成为研究趋势。如何优化知识发现方法,提高模型效率,增强知识发现结果的准确性和有效性是研究的重点和难点。未来研究中,建议提升知识发现结果的可解释性,构建有效的结果评估框架,规范化、多领域专家协同的知识发现流程,从而提高知识发现研究整体效果。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 PubMed. National Library of Medicine [EB/OL]. [2023 - 04 - 13]. <https://pubmed.ncbi.nlm.nih.gov/about/>.
- 2 谭玲, 鄂海红, 匡泽民, 等. 医学知识图谱构建关键技术及研究进展 [J]. 大数据, 2021, 7 (4): 80 - 104.
- 3 范媛媛, 李忠民. 中文医学知识图谱研究及应用进展 [J]. 计算机科学与探索, 2022, 16 (10): 2219 - 2233.
- 4 胡红娟, 周阳, 匡泽民, 等. 医学知识图谱应用研究进展 [J]. 医学信息学杂志, 2022, 43 (5): 30 - 33.
- 5 唐晓萍. 数据挖掘与知识发现综述 [J]. 电脑开发与应用, 2002 (4): 31 - 32.
- 6 SWANSON D R. Undiscovered public knowledge [J]. The library quarterly, 1986, 56 (2): 103 - 118.
- 7 WEEBER M, KLEIN H, DE JONG VAN DEN BERG L T, et al. Using concepts in literature - based discovery: simulating swanson's raynaud - fish oil and migraine - magnesium discoveries [J]. Journal of the American society for information science and technology, 2001, 52 (7): 548 - 557.
- 8 陈延雪, 杨长春, 葛天一, 等. 基于知识推理的医疗应急响应机制研究 [J]. 小型微型计算机系统, 2022, 43 (3): 638 - 643.
- 9 蒋川宇, 韩翔宇, 杨文蕊, 等. 医学知识图谱研究与应用综述 [J]. 计算机科学, 2023, 50 (3): 83 - 93.
- 10 田玲, 张谨川, 张晋豪, 等. 知识图谱综述——表示、构建、推理与知识超图理论 [J]. 计算机应用, 2021, 41 (8): 2161 - 2186.
- 11 SCHOENMACKERS S, DAVIS J, ETZIONI O, et al. Learning first - order horn clauses from web text [C]. Massachusetts: The 2010 Conference on Empirical Methods in Natural Language Processing, 2010.
- 12 GALÁRRAGA L, TEFLIOUDI C, HOSE K, et al. Fast rule mining in ontological knowledge bases with AMIE ++ [J]. The VLDB journal, 2015, 24 (6): 707 - 730.
- 13 付子轩, 周鹏, 任海燕, 等. 基于知识图谱的中西医结合急腹症诊疗推理分析 [J]. 中国实验方剂学杂志, 2023, 29 (11): 190 - 199.
- 14 NOOR A, ASSIRI A. A rule - based inference framework to explore and explain the biological related mechanisms of potential drug - drug interactions [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1155/2022/9093262>.
- 15 WILKOWSKI B, FISZMAN M, MILLER C M, et al. Graph - based methods for discovery browsing with semantic predications [J]. AMIA annual symposium proceedings, 2011: 1514 - 1523.
- 16 XIAN Y, FU Z, MUTHUKRISHNAN S, et al. Reinforcement knowledge graph reasoning for explainable recommendation [C]. Paris: The 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2019.
- 17 胡正银, 刘蕾蕾, 代冰, 等. 基于领域知识图谱的生命医学学科知识发现探析 [J]. 数据分析与知识发现, 2020, 4 (11): 1 - 14.
- 18 SCHUTTE D, VASILAKES J, BOMPELLI A, et al. Discovering novel drug - supplement interactions using SuppKG generated from the biomedical literature [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1016/j.jbi.2022.104120>.
- 19 HAN B S, DAHEE L, MINJOO K, et al. Enriching plausible new hypothesis generation in PubMed [J]. Plos one, 2017, 12 (7): e180539.
- 20 LAO N, COHEN W W. Relational retrieval using a combination of path - constrained random walks [J]. Machine learning, 2010, 81 (1): 53 - 67.
- 21 董文波, 孙仕亮, 殷敏智. 医学知识推理研究现状与发展 [J]. 计算机科学与探索, 2022, 16 (6): 1193 - 1213.
- 22 ZHAO D, WANG J, SANG S, et al. Relation path feature embedding based convolutional neural network method for drug discovery [J]. BMC medical informatics and decision making, 2019, 19 (S2): 59.
- 23 SYAFIANDINI A F, SONG G, AHN Y, et al. An automatic hypothesis generation for plausible linkage between xanthium and diabetes [J]. Scientific reports, 2022, 12 (1): 17547.
- 24 WANG Q, LIU J, LUO Y, et al. Knowledge base completion via coupled path ranking [C]. Berlin: The 54th Annual Meeting of the Association for Computational Linguistics, 2016.
- 25 陈钦况, 陈珂, 伍赛, 等. 关于主动学习下的知识图谱补充研究 [J]. 计算机科学与探索, 2020, 14 (5): 769 - 782.
- 26 ANDREJ K, THOMAS R, DIMITAR H. Link prediction on

- a network of co - occurring MeSH terms: towards literature - based discovery [J]. *Methods of information in medicine*, 2016, 55 (4): 340 - 346.
- 27 吕林媛, 周涛. 链路预测 [M]. 北京: 高等教育出版社, 2013.
- 28 MURATA T, MORIYASU S. Link prediction based on structural properties of online social networks [J]. *New generation computing*, 2008, 26 (3): 245 - 257.
- 29 ADAMIC L A, ADAR E. Friends and neighbors on the web [J]. *Social networks*, 2003, 25 (3): 211 - 230.
- 30 马代川, 罗代兵. 基于网络链路预测的药物分子重定位研究 [J]. *化学研究与应用*, 2020, 32 (5): 752 - 757.
- 31 张月霞, 冯译萱. 链路预测的方法与发展综述 [J]. *测控技术*, 2019, 38 (2): 8 - 12.
- 32 ZHOU T, LÜ L, ZHANG Y. Predicting missing links via local information [J]. *The European physical journal B*, 2009, 71 (4): 623 - 630.
- 33 KATZ L. A new status index derived from sociometric analysis [J]. *Psychometrika*, 1953, 18 (1): 39 - 43.
- 34 王林, 商超. 无标度网络中的链路预测问题研究 [J]. *计算机工程*, 2012, 38 (3): 67 - 70.
- 35 VURAL H, KAYA M. Prediction of new potential associations between LncRNAs and environmental factors based on KATZ measure [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1016/j.combiomed.2018.09.019>.
- 36 袁华兵, 刘敏, 杨延庆. 知识图谱中基于多关系路径的链路预测方法 [J]. *电子测量技术*, 2021, 44 (5): 135 - 138.
- 37 李栋. 基于网络结构的链路预测方法研究 [D]. 沈阳: 沈阳航空航天大学, 2018.
- 38 KLEIN D J, RANDIĆ M. Resistance distance [J]. *Journal of mathematical chemistry*, 1993, 12 (1): 81 - 95.
- 39 BRIN S, PAGE L. The anatomy of a large - scale hypertextual Web search engine [J]. *Computer networks and ISDN systems*, 1998, 30 (1): 107 - 117.
- 40 JE H G, WIDOM J. Simrank: a measure of structural - context similarity [C]. Edmonton: The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2002.
- 41 LIU W, LÜ L. Link prediction based on local random walk [J]. *EPL (Europhysics letters)*, 2010, 89 (5): 58007.
- 42 LI X, LIN Y, GU C, et al. SRMDAP: simRank and density - based clustering recommender model for miRNA - disease association prediction [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1155/2018/5747489>.
- 43 CHE Y, CHENG W, WANG Y, et al. A random walk with restart model based on common neighbors for predicting the clinical drug combinations on coronary heart disease [J]. *Journal of healthcare engineering*, 2021 (14): 4597391.
- 44 陆洪钰. 基于逻辑规则与深度学习的知识推理算法研究 [D]. 西安: 西安电子科技大学, 2021.
- 45 LI D, QU J, TIAN Z, et al. Knowledge - based recurrent neural network for TCM cerebral palsy diagnosis [J]. *Evidence - based complementary and alternative medicine: eCAM*, 2022 (34): 7708376.
- 46 YAN J, GAO Q, YU Y, et al. Combining knowledge graph with deep adversarial network for water quality prediction. [J]. *Environmental science and pollution research international*, 2022, 30 (4): 10360 - 10376.
- 47 刘禹琪. 中医名方知识图谱构建与链路预测模型的研究及应用 [D]. 长春: 东北师范大学, 2021.
- 48 孙水发, 李小龙, 李伟生, 等. 图神经网络应用于知识图谱推理的研究综述 [J]. *计算机科学与探索*, 2023, 17 (1): 27 - 52.
- 49 黄紫阳. 基于生物医学知识图谱和药物分子结构的药物 - 药物相互作用预测 [D]. 武汉: 华中农业大学, 2022.
- 50 DAI G, WANG X, ZOU X, et al. MRGAT: multi - relational graph attention network for knowledge graph completion [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1016/j.neunet.2022.07.014>.
- 51 BAKAL G, TALARI P, KAKANI E V, et al. Exploiting semantic patterns over biomedical knowledge graphs for predicting treatment and causative relations [EB/OL]. [2023 - 06 - 01]. <https://doi.org/10.1016/j.jbi.2018.05.003>.
- 52 DASGUPTA S, JAYAGOPAL A, HONG A L J, et al. Adverse drug event prediction using noisy literature - derived knowledge graphs: algorithm development and validation [J]. *JMIR medical informatics*, 2021, 9 (10): e32730.
- 53 郎宇翔. 基于图神经网络的文献知识发现研究 [D]. 北京: 中国科学院大学 (中国科学院文献情报中心), 2022.
- 54 NIAN Y, HU X, ZHANG R, et al. Mining on Alzheimer's diseases related knowledge graph to identify potential AD - related semantic triples for drug repurposing [J]. *BMC bioinformatics*, 2022, 23 (6): 1 - 15.
- 55 HWAN K Y, MIN S. A context - based ABC model for literature - based discovery [J]. *Plos one*, 2019, 14 (4): e215313.
- 56 孙佳. 基于网络链路预测与路径发现的药物 - 疾病知识发现研究 [D]. 沈阳: 中国医科大学, 2023.