

数据驱动的开源学术成果演化规律与合作模式分析

叶波甸, 高敏, 王伟, 陈阳

引用本文

叶波甸, 高敏, 王伟, 陈阳. 数据驱动的开源学术成果演化规律与合作模式分析[J]. 计算机科学, 2025, 52(8): 45-50.

YE Bodian, GAO Min, WANG Wei, CHEN Yang. Data-driven Analysis of Evolutionary Trends and Collaboration Patterns in Open Source Academic Achievements [J]. Computer Science, 2025, 52(8): 45-50.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[OpenRank动力学:面向开源生态的影响力评估与动态传播模型](#)

OpenRank Dynamics: Influence Evaluation and Dynamic Propagation Models for Open Source Ecosystems

计算机科学, 2025, 52(8): 62-70. <https://doi.org/10.11896/jsjcx.250300005>

[开源项目中的子社区发现与评价:以Apache IoTDB为例](#)

Sub-community Detection and Evaluation in Open Source Projects: An Example of Apache IoTDB

计算机科学, 2025, 52(7): 26-36. <https://doi.org/10.11896/jsjcx.250200108>

[端云人脸识别系统计算卸载策略设计](#)

Design of Computation Offloading Strategy for Device-Cloud Face Recognition System

计算机科学, 2025, 52(6A): 240600065-7. <https://doi.org/10.11896/jsjcx.240600065>

[开源软件组件漏洞检测与自动修复技术研究综述](#)

Survey of Open-source Software Component Vulnerability Detection and Automatic Repair Technology

计算机科学, 2025, 52(6): 1-20. <https://doi.org/10.11896/jsjcx.240400023>

[基于大语言模型自身的提示语公平性自动优化与评估](#)

Automatic Optimization and Evaluation of Prompt Fairness Based on Large Language Model Itself

计算机科学, 2025, 52(4): 240-248. <https://doi.org/10.11896/jsjcx.240900008>

数据驱动的开源学术成果演化规律与合作模式分析

叶波甸¹ 高敏¹ 王伟² 陈阳¹

1 复旦大学计算与智能创新学院 上海 200438

2 华东师范大学数据科学与工程学院 上海 200062

(bdye22@m.fudan.edu.cn)

摘要 开源已经成为当今软件开发领域中不可忽视的潮流,也是推动技术创新与进步的关键力量。深入探究开源发展的趋势及其合作模式,不仅有助于揭示学术界和工业界的发展态势,也能为相关研究人员或者政策制定者提供制定合理目标与规划的依据。基于 DBLP 数据库,收集 1998 至 2023 年间的 5990 篇开源主题论文,系统分析了开源领域的整体发展轨迹。通过分析论文发表的期刊/会议、标题、引用数等统计性信息,发现当前开源成果可以被分为开源软件设计开发与开源领域实证研究两种类型,且前者在数量上占据明显优势。为了更有效地揭示开源领域研究者间的合作关系以及对应国家间的合作模式,建模开源领域研究者合作的高阶关系,同时进一步挖掘研究者背后所反映的国家合作网络。研究表明,开源领域大多数研究者来自高校,并且他们的研究兴趣主要集中在软件工程或者开源软件方面。此外,在国家合作网络中占据重要地位的国家是以美国为代表的发达国家,而以中国为代表的发展中国家对开源领域的重视程度也在提高。通过对比各国的合作模式,发现开源领域中跨国合作的模式尚未形成主流。

关键词: 开源;高阶关系;合作网络;国家合作;演化规律

中图分类号 TP311.5

Data-driven Analysis of Evolutionary Trends and Collaboration Patterns in Open Source Academic Achievements

YE Bodian¹, GAO Min¹, WANG Wei² and CHEN Yang¹

1 College of Computer Science and Artificial Intelligence, Fudan University, Shanghai 200438, China

2 School of Data Science and Engineering, East China Normal University, Shanghai 200062, China

Abstract Open source has become a significant trend in software development, driving technological innovation and progress. Insights into current trends and collaboration models can help researchers and policymakers set reasonable goals. This paper analyzes 5990 papers related to open source from the DBLP database, published between 1998 and 2023, to explore the evolution of open-source related studies. The analysis of publication venues, titles, and citation counts reveals two main categories of research: those focused on open-source software and those on empirical studies, with the former being more prevalent. Additionally, the collaborative relationships among researchers and countries are modeled and the findings indicate that most researchers are affiliated with universities, primarily focusing on software engineering and open-source. Furthermore, collaborations tend to be concentrated within single countries, predominantly involving developed nations.

Keywords Open source, Higher-order relationship, Collaboration network, International collaboration, Evolution

1 引言

“开源”(Open Source)作为一种代表性的软件开发方法^[1],旨在强调开放性、透明性和社区参与,具有深远的社会和经济影响。开源软件是开源的具体表现之一,初期主要由一些小型项目和非营利组织推动,例如 GNU 项目和 Linux 操作系统。随着这些开源项目的成功,越来越多的开发者和企业开始采纳开源模式,进一步推动了开源文化的普及与发展^[2]。

进入 21 世纪,开源的概念不断扩展到其他领域,推动了各个领域的创新和协作。以 GitHub 为代表的开源平台的出现,为开发者提供了共享与贡献代码的全球性的开源社区,使得开源项目的协作变得更加便捷。截至目前,GitHub¹⁾上的项目(Repository)已经超过 5.18 亿。不仅如此,根据开源贡献度评价算法 OpenRank 计算出^[3-4]的排名,许多全球知名度较高的企业都积极支持开源并参与开源项目。在此背景下,中国的开源产业也得到蓬勃发展,并且引起了政府的高度

¹⁾ <https://github.blog/newsinsights/octoverse/octoverse-2024/>

到稿日期:2025-02-05 返修日期:2025-06-04

基金项目:国家自然科学基金(62072115)

This work was supported by the National Natural Science Foundation of China(62072115).

通信作者:陈阳(chenyang@fudan.edu.cn)

重视,国家“十四五”规划纲要和国家软件发展战略对其作出重要部署。在2024开放原子开源生态大会上,国家工业和信息化部也表明中国对开源体系建设的重视^[5]。综上所述,开源已经被视为现代软件开发的重要组成部分,拥有巨大的发展潜能,影响着技术标准、行业规范、商业模式和政府决策。

随着开源项目在各行各业中的应用日益广泛,学术界的研究者也开始探索开源领域的重要议题,囊括了技术实现到开源社区的协作模式,再到社会影响等多方面。尽管已有大量学术文献聚焦于探究开源软件面临的各类问题和挑战,但现有工作对于如何研究开源领域的学术合作模式和趋势尚未形成系统的研究框架。一方面,目前的研究主要集中在开源软件的技术实现和应用案例,缺乏对开源成果类型的系统性梳理与分析。现有研究往往缺乏对不同类型开源成果发展进程的量化,这使得广大研究者在理解开源成果的宏观视图时面临困难。另一方面,虽然部分研究者开始关注开源项目的社区动态与合作关系,但关于哪些学者在开源领域拥有高影响力仍缺乏讨论。具体而言,现有研究往往缺乏对开源成果中的合作者角色的刻画,也未能全面分析合作网络对研究产生的影响。因此,深入探讨开源背景下的成果多样性,量化其发展进程,以及刻画推动开源研究的关键学者及其合作模式,对于填补当前研究空白并推动学术界对开源成果的理解与应用具有重要意义。

本文主要的研究贡献如下:

1) 开源主题出版物数据集的构建与分析。首次构建了一个涵盖开源主题文献的数据集,该数据集包含DBLP数据库中收录的包含“open-source”与“open source”关键词的主要学术文献。通过对这一数据集的深入分析,揭示了当前开源研究的整体态势,包括潜在的主要研究方向、关键会议与期刊,以及代表性论文。

2) 开源合作网络的构建与分析。构建了研究者间的高阶合作网络,分析了网络的内部结构,同时识别在合作网络中具有关键位置的研究者,探讨了他们在推动开源学术发展过程中的重要作用。

3) 国家合作网络的构建与分析。为深入挖掘合作团队的特征,探讨了不同国家之间的合作模式及其潜在影响,为研究者理解开源领域的国际合作模式提供了新的视角,也有助于推动跨区域的协作与知识共享。

2 相关工作

在软件技术不断发展的背景下,开源软件自20世纪90年代起受到广泛关注。在开源社区领域的研究中,Liu等^[6]探索了2007—2008年国外相关研究聚焦的开源软件的多个关键领域,研究结论表明开源社区的动力机制来源于参与开源项目的满足感和个人能力的提高,因此开发者愿意在开源社区中积极参与知识共享。与之相反,Miller等^[7]研究了开源社区中的网络恶意行为的特征,发现开源社区中的恶意行为的突出表现为用户提出的强硬要求和傲慢言论,以及技术分歧引发的侮辱性评论。这些研究结果为未来开源社区的恶意行为检测提供了重要启示。

在开源软件协同演化的研究中,软件网络和开发者网络

的对比分析逐渐成为重要的研究方向。Li等^[8]的研究表明,开源软件的开发主要依赖于开发者的自我管理和志愿贡献,其中核心团队在项目的开发和实施中起到了主导作用。通过观测软件项目生命周期中开发者网络的动态变化,该研究以GNU/Linux公共数据集为例,探讨了开发者网络与软件项目的协同演化机制,包含开发者数量、志愿参与度、开发者经验与项目演化的关系,揭示了开源软件开发过程中团队结构与项目演化的协同机制,为进一步理解开源软件开发中的团队演变和项目发展提供了有力的理论依据。He等^[9]通过收集开源软件项目Vuze多个版本的源代码和开发者交互数据,从网络规模、平均度、最短路径、聚类系数及社区结构等角度对比了两类网络的演化特征,为理解开源软件项目的协同演化机制提供了重要的理论依据。Xia等^[10]首先分析了OpenStack项目的Git提交数据,然后基于Git提交数据中的父子哈希码关系构建了代码协作修订网络,以此探讨其结构与演化特征。该研究发现,OpenStack的开发过程可被视为一个稳定的网络型系统,其中核心开发人员发挥了主导作用,但一般开发人员在总工作量中占据了主体地位。此外,OpenStack的子项目与社区演化之间存在显著的关联性,开发人员的工作领域表现出高度的集中性和持续稳定性。

3 研究方法

3.1 数据收集

本文的数据主要来自DBLP数据库。DBLP是一个被广泛使用的计算机科学文献数据库,它提供文献检索和引用服务,是研究人员获取计算机科学领域研究成果的一个重要资源。截至2024年9月9日,DBLP已经收录超过七百万篇出版物,覆盖超过三百万位学者。不仅如此,DBLP的数据也被广泛应用于文献计量学和科研影响力分析等研究领域,是全球研究者的重要参考工具^[11-13]。本文的目标是识别一组与开源相关的论文,因此,本文检索DBLP XML数据集(截至2024-09-01)中每篇论文的标题,并且仅保留1998年(包括1998年)后的论文。通过使用关键词“open-source”和“open source”获得相应论文子集,同名论文仅被统计一次,每一篇论文均被独立处理。DBLP为同名学者创建不同的后缀名,即同名学者的名字后面会用序号进行标识。由于出版商的多样性,且为了保证数据质量,本文最后爬取收录于计算机科学领域4个最具权威的出版商平台的论文,即ACM,IEEE,Elsevier和Springer。如图1所示,ACM,IEEE,Elsevier和Springer平台论文占据初筛后整体数量的80%以上。后文中的数据处理与分析以这四大出版商平台中的文献为基础。

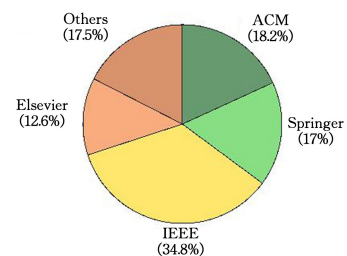


图1 初筛出版商统计图

Fig. 1 Statistics diagram of filtered publishers

3.2 数据处理

为了保证分析的针对性,需要获取作者所属机构的相关信息。由于 DBLP 提供的论文元数据中尚未包含作者的所属机构,因此本文利用 OpenAlex^[14-15] 来收集用户的机构信息。OpenAlex 由非盈利组织 OurResearch 维护,是一个完全开放的全球研究系统目录与科学知识图谱,继承自微软的学术知识图谱,目前依然保持月度更新的状态。OpenAlex 包含了 2.09 亿篇出版物、超过百万位作者及超过十万数量级的会议/期刊文献。通过输入论文的 DOI,可以利用 OpenAlex 提供的 PyAlex API 来提取出学者所属机构与对应国家。最终,本文筛选出研究文献 5990 篇,涉及到学者 17800 位。

3.3 数据建模

现有的研究大多基于普通图对学术合作网络进行建模,然而普通图在表达学术合作的复杂关系时存在一些局限性。这是因为普通图只能表示成对节点之间的二元关系,如果使用普通图建模,我们无法得知两个有合作关系的研究者之间是因为哪一篇论文建立联系,而超图可以保留这一高阶关联。为弥补这些不足,超图(Hypergraph)的概念^[15-16]应运而生。超图不仅能够将多个节点通过一条超边进行连接,从而有效表达多元合作关系,还能够更好地捕捉合作网络中的高阶交互和复杂结构。因此,采用超图对本文中涉及到的网络结构进行建模,能够更全面、细致地刻画合作的多方动态与交互模式,有助于更深入地理解学术合作的潜在规律与特点。

4 研究结论

4.1 开源研究成果概述

由于 DBLP 对 2024 年的论文收录尚不完整,因此本文仅讨论 1998—2023 年之间的开源主题文献。从图 2 所示开源领域出版数量的总体增长趋势可以推测出,在开源的概念被提出后,开源领域逐渐得到广大研究者的关注与探索,相应的研究成果得以发表。因此,从宏观来看,开源相关的研究工作整体呈现上升的趋势。除了分析文献数量的整体趋势外,本文继续探索各出版商平台与开源领域相关度的变化趋势。图 3 展示了不同出版商平台收录论文数量的增长情况,可以发现 IEEE 平台的增长趋势最为显著,并且与其他三大平台拉开明显差距。Springer 平台的收录数量相对最少,且波动较大,目前趋于稳定。ACM 与 Elsevier 在 2020 年前后的收录情况趋近。

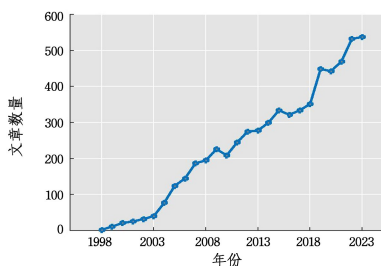


图 2 1998—2023 间包含开源关键词的论文数量

Fig. 2 The number of papers containing open source keywords between 1998 and 2023

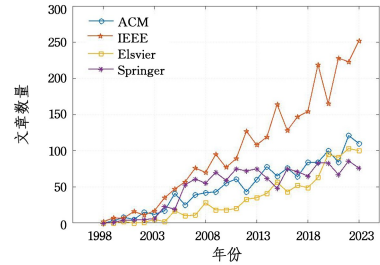


图 3 1998—2023 四大出版商的收录情况

Fig. 3 Inclusion of the four major publishers between 1998 and 2023

随后,统计了论文发表的期刊/会议的情况,并在表 1 中列出前 15 篇论文数量占比最高的期刊/会议。

表 1 论文发表地统计与排名

Table 1 Statistics and ranking of venues

序号	发表地	比例/%
1	Journal of Open Source Software	4.0
2	Computer Physics Communications.	1.8
3	SoftwareX	1.5
4	IEEE Access	1.4
5	Environmental Modelling and Software	1.3
6	Hawaii International Conference on System Sciences	1.2
7	IEEE/ACM International Conference on Mining Software Repositories	1.2
8	IEEE Software	1.1
9	Computer	0.9
10	Journal of Systems and Software	0.9
11	Computers & Geosciences	0.9
12	International Conference on Software Engineering	0.8
13	Empirical Software Engineering	0.7
14	ACM Technical Symposium on Computer Science Education	0.7

首先可以观察到论文发表的类型以期刊为主,其中排名第一的期刊是 *Journal of Open Source Software*,该期刊主要发表与开源软件相关的论文。作为一个支持开源软件转化为对应论文成果的发表平台,其评审标准综合了开源软件本身的成熟度、参与者数量、代码提交的数量、代码行数以及是否被学术论文引用等多方面因素。因此,从以上因素中可以推测大多数研究者希望自己的开源成果可以发表在针对性更强的开源主题发表平台。此外,表中论文发表期刊/会议的一个重要特点是超过半数的名称都包含“Software”关键词,也体现出开源成果的呈现形式更加偏向于开源软件,需要软件工程相关的专业背景与知识。

为了探索现有论文涉及的潜在主题,根据论文的标题信息生成如图 4 所示的词云图。其中字体越大的词汇出现的次数越多,说明它在文本中的重要性相对更高。通过观察可以发现词云中出现的高频词有“Software”“System”“Framework”等,与论文发表期刊/会议的名称拥有较高相似度;同时,从关键词“Development”“Empirical”“Review”中可以发现,当前开源成果可以被分为开源软件设计开发与开源领域实证研究两种类型。从研究动机角度分析,可以推测随着开源思想被社会接纳程度的提高,开源社区的发展与活跃现象逐渐被学术研究人员关注,激发了他们挖掘与分析开源社区生态模式的形成动因与发展驱动力;从研究数据角度分析,可以推测出开源思想的传播与影响下催生了更多开源软件的诞生,甚至从闭源转为开源,从而为这些研究者提供了更可靠的

了合作网络中的超边度分布。超边度(Hyperedge degree)^[19],指一条超边所拥有的节点数量。如图6所示,超边度的分布主要集中在1到5之间,其中平均每篇论文有3.18位作者,表明大部分论文的合作者数量少于5位。值得注意的是,175篇论文拥有多于10位合作者,且拥有最多合作者的论文高达30位合作者,本文推测该现象背后的潜在原因与开源软件的复杂度有关:开源软件的搭建往往涉及到多个软件开发流程,每个流程可能都需要有相应的负责人员,甚至跨领域的专业人士参与,因此合作者的数量可能会随着开源软件规模或者复杂程度的提高而增多。

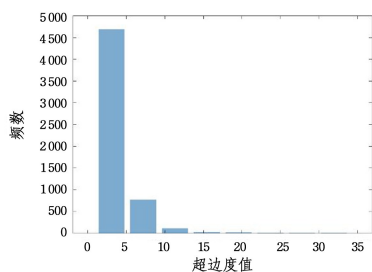


图6 研究者合作网络的超边度分布情况

Fig. 6 Distribution of hyperedge degree in researchers' collaborative network

4.3 国家网络分析

除了建模学者之间的合作网络,本文还对国家间的合作关系进行了探索。由于本文可以获取到每个作者单位对应的国家代码,因此首先对所有研究者的国家分布情况进行分析。如图7所示,颜色越深的地区代表来自该国家或地区的研究者数量越多。可以直观地观察到,来自美国的研究者数量在全球各国中最多;而在欧洲地区,多个国家的相关研究者数量超过1000人,体现出欧洲在开源领域的研究积累。相较之下,南美和亚洲地区只有少数国家的研究者数量与其他国家形成明显对比。例如,在南美洲,巴西的研究者数量显著高于其他国家;在亚洲,中国的研究者具有更高的参与度。

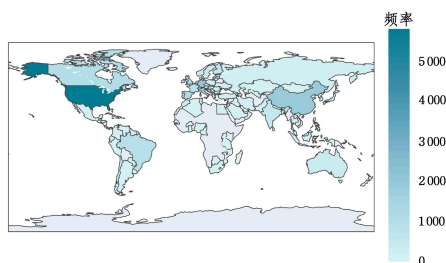


图7 研究者的国家分布情况

Fig. 7 Country distribution of researchers

为了进一步理解这些国家之间的合作情况,以国家为单位构建国家合作网络超图。其中,节点代表作者单位对应的国家,超边代表这些国家之间的合作关系。表4中对每个国家按照论文数量进行降序排序,对于一篇论文,主要将第一作者的国家纳入统计。从表4中可以发现,美国、德国以及意大利的发表论文数量位于前三。同时可以发现,以上国家都属于发达国家,在经济发展水平、科技实力和教育质量等多方面拥有较深厚的积累,因此可以推测这些国家在推动开源成果的发展方面有较丰富的经验,甚至可能拥有来自政府、工业界

和学术界的多方面支持。另外,中国在开源领域的研究也位于前列,说明以中国为首的发展中国家对开源领域研究的重视程度逐渐提高。为了说明国家间的合作情况,表4中同时统计了每个国家的国内外合作比例。可以发现,大多数国家更倾向于与本国的研究者合作,跨国合作的形式并没有成为主流,其中潜在的原因可能与开源成果完成过程中的沟通效率、协作难度与社交关系有关。

表4 论文发表数量位于前10位的国家

Table 4 Top 10 countries in publications

序号	国家缩写	论文数	国内合作比例/%	国际合作比例/%
1	US	1486	87.5	12.5
2	DE	498	80.9	19.1
3	IT	377	74.8	25.2
4	CN	374	76.4	23.6
5	CA	274	76.2	23.8
6	GB	254	68.1	31.9
7	ES	246	73.9	26.1
8	BR	211	81.0	19.0
9	FR	185	75.1	24.9
10	JP	158	82.2	17.8

将国家合作网络超图团展开^[20]为普通图,随后进行有效规模(Effective size)的计算。有效规模大小是与结构洞理论相关联的重要指标之一^[21]。有效规模的原理是描述某个节点冗余连接的程度,一个节点获得的冗余连接越少,其有效规模越大,这类研究者越有可能是结构洞占据者。在合作网络中,结构洞占据者可能是连接不同学术团体的桥梁,是关键信息扩散路径上的关键节点。如表5所列,有效规模前2的国家是US和DE,说明这两个国家的研究者不仅拥有紧密的内部合作,而且拥有广泛的跨国合作,因此更有可能在国家网络中占据重要的位置。表中的亚洲国家仅仅包括中国和印度,说明这两个亚洲国家寻求对外合作的意向相对于其他亚洲国家而言更加明显。

表5 有效规模位于前10位的国家

Table 5 Top 10 countries in effectiveness

序号	国家缩写	有效规模
1	US	56.8
2	DE	34.0
3	ES	30.5
4	CA	29.7
5	GB	29.6
6	IT	28.1
7	FR	27.4
8	CN	25.8
9	NL	20.2
10	IN	18.6

结束语 开源已成为当今软件开发领域的重要趋势,极大地推动了技术创新与进步。本文通过分析1998—2023年间DBLP数据库收录的5990篇开源主题相关论文,揭示了开源从提出至今,开源领域的整体发展趋势。通过分析论文标题与高影响力论文可以发现,当前开源学术成果分为两类,一类是与开源软件相关的研究,另一类是与开源相关的实证研究,其中前者在数量上明显占优。此外,为了更有效地描述研究者之间的协作关系及国家间的合作模式,本文基于超图对开源领域的研究者合作关系进行建模,从不同角度深入分析

高阶网络结构特征及重要节点。结果显示,在开源领域拥有较多发表成果的研究者大多数来自高校,研究兴趣主要集中在软件工程和开源软件领域。最后,通过研究者构建国家合作网络,发现目前开源成果的合作主要集中在单一国家内,并且网络中占据重要地位的国家大多为发达国家,这些发现为进一步理解开源领域的合作动态提供了重要视角。

参 考 文 献

- [1] SCACCHI W. Free/open source software development [C]// Proceedings of the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering, 2007:459-468.
- [2] HAUGE Ø, AYALA C, CONRADI R. Adoption of open source software in software-intensive organizations — A systematic literature review[J]. Information and Software Technology, 2010, 52(11):1133-1154.
- [3] ZHAO S, XIA X, FITZGERALD B, et al. OpenRank Leaderboard: Motivating Open Source Collaborations Through Social Network Evaluation in Alibaba[C]// Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Practice, 2024:346-357.
- [4] HUANG W, XIA X, ZHOU A, et al. OSGraph: A Data Visualization Insight Platform for Open Source Community[C]// International Conference on Database Systems for Advanced Applications. Singapore: Springer Nature Singapore, 2024:476-479.
- [5] 王洁. 加快开源生态系统建设降低开发门槛[N]. 中国信息化周报, 2024-09-30(022).
- [6] LIU L, SUN T, HUANG G B. Review of Foreign Research Progress on Open-source Software from 2007 to 2008 [J]. Library Development, 2009(4):100-106.
- [7] MILLER C, COHEN S, KLUG D, et al. “Did you miss my comment or what?” understanding toxicity in open source discussions[C]// Proceedings of the 44th International Conference on Software Engineering, 2022:710-722.
- [8] LI Q F, LI B. Evolution of Contributors in Open Source Software Development[J]. Computer Science, 2015, 42(12):43-46.
- [9] HE P, LI B, CHENG C, et al. Analysis of the Co-evolution of Socio-technical Networks in Open-source Software Systems [J]. Complex Systems and Complexity Science, 2015, 12(2):38-45.
- [10] XIA H X, ZHANG X, ZHANG X Z. An analysis of developer collaboration networks in the OpenStack open-source software project [J]. Systems Engineering-Theory & Practice, 2017, 37(5):1373-1382.
- [11] LEY M. The DBLP computer science bibliography: Evolution, research issues, perspectives[C]// International Symposium on String Processing and Information Retrieval. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002:1-10.
- [12] GAO J, WU J, DING J. Heterogeneous Graph Condensation[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(7):3126-3138.
- [13] ALI Z, QI G, ULLAH I, et al. GLAMOR: Graph-based Language Model embedding for citation Recommendation[C]// Proceedings of the 18th ACM Conference on Recommender Systems, 2024:929-933.
- [14] PRIEM J, PIWOWAR H, ORR R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts [J]. arXiv:2205.01833, 2022.
- [15] KIM S, LEE S Y, GAO Y, et al. A survey on hypergraph neural networks: an in-depth and step-by-step guide[C]// Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024:6534-6544.
- [16] FENG Y, YOU H, ZHANG Z, et al. Hypergraph neural networks[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2019:3558-3565.
- [17] GAO M, LI Z, LI R, et al. EasyGraph: A multifunctional, cross-platform, and effective library for interdisciplinary network analysis[J/OL]. Patterns, 2023, 4(10):100839.
- [18] MANCASTROPPA M, IACOPINI I, PETRI G, et al. Hypercores promote localization and efficient seeding in higher-order processes[J]. Nature Communications, 2023, 14(1):6223.
- [19] LIU B, ZENG Y J, YANG R M, et al. Fundamental statistics of higher-order networks: a survey[J]. Acta Physica Sinica, 2024, 73(12):128901.
- [20] SUN L, JI S, YE J. Hypergraph spectral learning for multi-label classification[C]// Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008:668-676.
- [21] LIN Z, ZHANG Y, GONG Q, et al. Structural hole theory in social network analysis: A review[J]. IEEE Transactions on Computational Social Systems, 2022, 9(3):724-739.



YE Bodian, born in 2000, master. Her main research interests include social computing and complex networks.



CHEN Yang, born in 1981, Ph.D, professor. His main research interests include social computing, intelligent networks and systems, and open source big data.

(责任编辑:柯颖)