

共被引分析方法迭代创新路径研究

宋歌^{1,2}

(1. 东南大学图书馆情报科学技术研究所, 南京 210096; 2. 东南大学经济管理学院, 南京 210096)

摘要 科研领域的迭代创新是一项原始创新得以扩散和发展的主要创新形式。该研究以共被引分析方法为典型案例, 探索科研领域迭代创新的基本特征。通过对共被引分析方法迭代创新的整体趋势、创新阶段、迭代路径和跨学科迭代创新的分析, 发现其特征包括对原始创新核心内容的不断调用以及对相关知识的引进与整合; 面向问题的反复调整迭代; 微创新的累积催生原始创新; 在迭代创新形成规模之前存在“沉寂期”; 学科间的广泛扩散不仅促进迭代创新, 而且不同学科在迭代创新中的“角色”不同, 存在创新源发学科、创新助力学科和创新采纳学科; 另外, 持续的迭代创新有赖于外部环境的支撑。

关键词 科研创新; 共被引分析; 迭代创新; 创新路径

Research on Iterative Innovation Path of Co-citation Analysis Method

Song Ge^{1,2}

(1. *The Institution of Information Science & Technology, Southeast University, Nanjing 210096;*
2. *School of Economics & Management, Southeast University, Nanjing 210096*)

Abstract: In the field of scientific research, iterative innovation is the principal method for spreading and developing original innovation. This study takes co-citation analysis as a typical case to explore the basic characteristics of iterative innovation in scientific research. From the analysis of the overall trend, innovation stage, iteration path, and interdisciplinary iteration innovation of the co-citation analysis method, it is found that its characteristics include the continuous call for the core content of original innovation and the introduction and integration of relevant knowledge; problem-oriented iterative adjustment; the accumulation of micro-innovation, which promotes the original innovation; and the existence of a “Silent period” before the scale of iteration innovation is formed. The wide spread of disciplines not only promotes iterative innovation, different disciplines such as innovation-occurred disciplines, innovation-assisted disciplines, and innovative adoption disciplines also have different “roles” in iterative innovation. It was also found that sustained iterative innovation depends on the support of the external environment.

Key words: scientific research innovation; co-citation analysis; iterative innovation; innovation path

1 引言

“用于情报研究的文献计量方法中最具影响力的首推共引分析方法”^[1]。共被引分析方法自1973

年提出以来, 已经历了近半个世纪的发展。起初, Small受到Kessler提出的“文献耦合(bibliographic coupling)”^[2]的启发, 逆向思维提出了一种新型耦合, 即“文献共被引(document co-citation)”^[3]。

收稿日期: 2019-01-30; 修回日期: 2019-09-09

基金项目: 国家社会科学基金项目“学术创新扩散过程及创新力测度研究”(15CTQ027)。

作者简介: 宋歌, 女, 1980年生, 副研究馆员, 硕士生导师, 主要研究领域为科学计量与学术评价, E-mail: songge.2000@163.com。

这一转换的创造性在于突破了文献耦合的静态特征，使得文献之间的关联关系及关联强度不再一成不变。文献共被引是伴随科研成果产出，通过新成果对不同已有成果的引用，构成的已有成果之间不断变换的知识组合，其组合关系由新产生的知识内容决定。由于新知识的加入与已有知识的不断重组正是科研发展的主要形式，因此文献共被引相比文献耦合更适合描述研究主题的演变与科学结构的变化。至今，对共被引分析方法的研究与应用仍然十分活跃，同样得益于该项创新的内在创新力不断被新的科研成果和新的知识组合所激发。这一过程可被称为科研领域的迭代创新。

“迭代创新”从数学概念“迭代计算”和软件开发“迭代模式”演变而来，并首先出现在企业实践中。在数学中，迭代函数是重复的与自身复合的函数，其过程叫做迭代。在计算机科学中，迭代算法是对一段代码进行重复执行，在每次执行这段代码时，都从变量的原值推出它的一个新值。与重复调用函数自身实现循环的递归不同，迭代是对函数内某段代码实现循环调用。在社会实践中，伴随互联网带来的高度动态性，技术创新以快速迭代与试验的方式展开^[4]，迭代创新的概念应运而生。Fitzgerald 等^[5]于 2011 年指出，迭代创新是运用迭代法将创意市场化并基于市场反馈快速调整出新。如今，“迭代”已经由一种算法逐步升级为一种方法、理念和思维模式，其核心是不断用变量的旧值递推新值的过程。当前，科研周期显著缩短，科研产出呈指数增长，科研创新也以迭代创新的方式呈现出来。研究热点、新兴领域的兴起往往是某项原始创

新或曰突破性创新通过科学家们的信息觅食或信息偶遇所选定、采纳，从而触发的迭代创新过程。科研领域的迭代创新就是在原始创新的基础上，对该创新的核心部分一再调用，并面向不同的研究问题与应用场景吸纳相关理论、方法与技术，做出不断改进的创新过程。

科研领域的迭代创新是一项原始创新得以扩散和发展的主要创新形式，而迭代创新的成果数量和活跃程度直接反映了原始创新的理论和实践价值的高低。迭代创新有以下特征：①调用原始创新核心原理算法；②面向特定问题反复试错调整；③累积微小创新产生大变革。科研领域的迭代创新是否也具有类似特征？本文以共被引分析方法为典型案例，重现该项原始创新的迭代创新过程与路径，探讨科研领域迭代创新的基本特征。

2 实证与结果

在 Web of Science (WoS) 核心合集中以“cocitation”或“co-citation”为主题词进行检索，限定文献类型为 Article 或 Review，时间跨度为 1973-2017。清洗检索结果数据，排除 3 篇系统误检论文，最终获得 1143 篇相关成果。根据论文题录信息中的“Research Areas”字段，标注每篇成果所属研究领域。

2.1 迭代创新整体趋势

共被引分析方法的迭代创新整体趋势如图 1 和图 2 所示。如果把图 1 旋转 180°，则会呈现 Crane 所

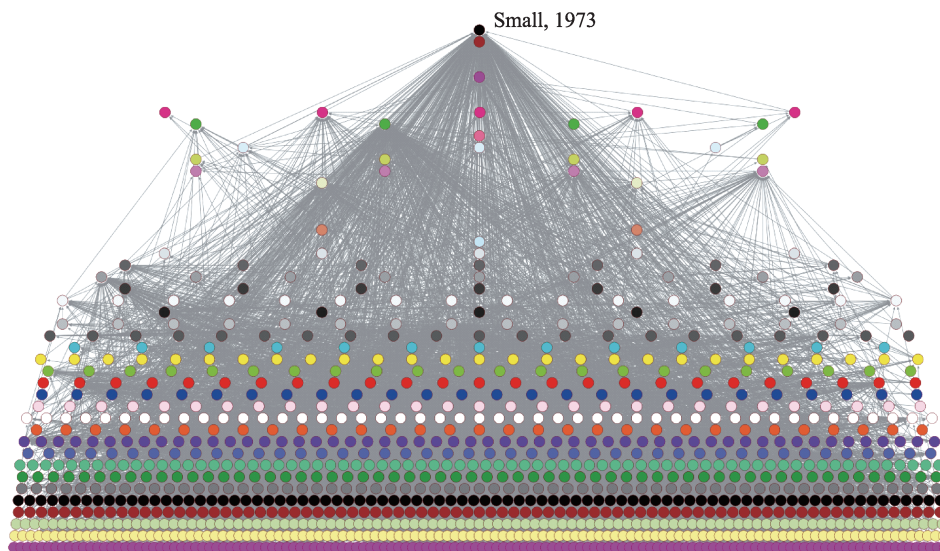


图 1 共被引分析方法扩散网络



图2 共被引分析方法历年成果数量

说的创新与科学增长中的“倒金字塔现象”^[6]。即虽然主要的创新在一个领域的全部历史中不断出现，然而这些成果是通过对这一领域早期成果的引证而联系起来的。大量论文都直接或间接地与处于塔顶的几篇论文发生联系。在共被引分析方法的研究领域中，Small于1973年提出“cocitation”的文章充当了“塔尖”的角色。在WoS核心合集中，至2017年，直接引用该文的论文有442篇，占成果总数的39%；若加上二级引证论文，为87%；若包含三级引证，则达到93%。以此类推，形成迭代创新过程。

图2显示，共被引分析方法的开山之作在其发表后的一二十年中并没有得到广泛扩散，最初的3篇论文都是由Small发表的。1990年以后，相关研究逐渐升温，截至数据采集结点年，研究热度持续攀升的趋势没有改变。因此，在迭代创新整体爆发之前，存在仅有少量迭代创新成果的“沉寂期”。

2.2 迭代创新阶段

文献[7]提出的创新扩散速度、扩散加速度、扩散广度、扩散强度和扩散延时系列指标可用于描述学术创新扩散过程的基本特征。共被引分析方法的扩散速度和加速度指标测度结果如表1所示，据其绘制的共被引分析方法扩散曲线如图3所示。结合表1和图3可见，2010年的扩散加速度最高，为

27，是该曲线上的二阶拐点。根据创新扩散理论(diffusion of innovations)^[8]，扩散在此时达到了临界值，即2010年以后，可以判定共被引分析方法作为一项科研创新将成功扩散，其累积成果数在时间轴上将形成S形曲线。至于扩散阶段，由于2006年加速度明显增加，因此1973—2005年为该迭代创新的起步阶段，时长33年；自2006年始为起飞阶段，至2017年已有12年。以下针对不同阶段进行分析，其中起飞阶段以二阶拐点为界分为两个部分论述。

2.2.1 起步阶段：1973—2005年，迭代伊始

起步阶段的成果占总成果数的22.48%，而被引量占总被引量的51.46%。共被引分析方法的主要技

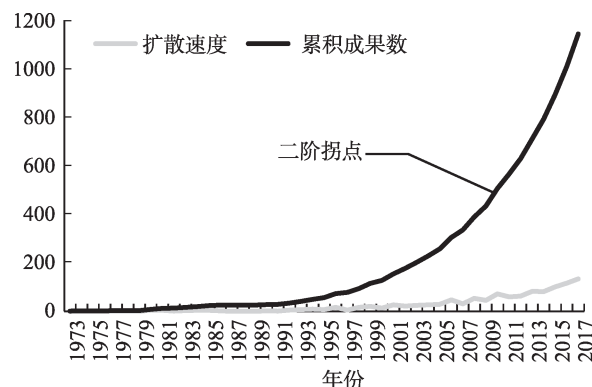


图3 共被引分析方法扩散曲线

表 1 共被引分析方法扩散基础指标测度结果

年份	扩散速度	扩散加速度	扩散广度	年份	扩散速度	扩散加速度	扩散广度	年份	扩散速度	扩散加速度	扩散广度
1973	1	0	2	1988	0	0	9	2003	25	3	26
1974	1	0	4	1989	0	0	9	2004	27	2	32
1975	0	-1	4	1990	2	2	9	2005	29	2	33
1976	0	0	4	1991	1	-1	9	2006	47	18	37
1977	1	1	5	1992	5	4	12	2007	31	-16	39
1978	0	-1	5	1993	7	2	12	2008	54	23	44
1979	0	0	5	1994	9	2	14	2009	45	-9	47
1980	5	5	6	1995	7	-2	15	2010	72	27	53
1981	4	-1	6	1996	16	9	16	2011	60	-12	54
1982	1	-3	6	1997	5	-11	17	2012	63	3	59
1983	3	2	6	1998	16	11	21	2013	83	20	66
1984	4	1	6	1999	21	5	22	2014	81	-2	68
1985	4	0	6	2000	13	-8	22	2015	101	20	69
1986	2	-2	8	2001	27	14	23	2016	117	16	71
1987	0	-2	8	2002	22	-5	25	2017	134	17	81

术、主要应用领域和主要研究分支在起步阶段已基本形成。图 4 中最大的几个节点依次是共被引、科学结构、作者共被引、知识结构、引文分析和文献计量学。从研究主题来看，首先“引文分析”和“文献计量学”既是共被引分析方法形成的根基，也是被该方法所拓展的领域。至此，引文分析和文献计量学进入以共被引分析为主导的新阶段，而 1963 年提出的“文献耦合分析”有所沉寂。其次，国际上，如何更准确、更大范围地揭示“科学结构”和“知识结构”是文献计量学最重要的研究问

题，而“共被引分析”为此提供了新方法。最后，“作者共被引”在共被引分析方法的起步阶段就成为了显著的研究领域。

从学科扩散广度来看，截至 2005 年，共被引分析方法已扩散至 33 个学科。其中，图书情报学和计算机科学是共被引分析方法产生的源头，Small 提出共被引分析的文献就是这两个学科的跨学科研究成果。其次，共被引分析在商业与经济学、运筹学与管理学的应用也较多。此外，相关研究在迭代创新的起步阶段还进入科学史、医学、心理学、工程、社会学、数学、通信、药理学、遗传学、化学、农学、生物学、语言学和人文艺术等领域。从扩散广度的增长来看，共被引分析方法在 2003 年以后进入学科扩散的快速长时期，这正是进入起飞阶段的前两年。

2.2.2 起飞阶段：2006—2010 年，跨越拐点

该阶段的成果占总成果数的 21.70%，被引量占总被引量的 29.49%。图 5 显示的主要研究领域除了包含图 4 中的 6 大节点以外，增加了网络和信息可视化 2 个较大的节点。新增的次要节点有：科学计量学、科研合作、社会网络分析、寻径网络、学科、领域、主题。结合新增的主要和次要节点，即可看出起飞阶段最初 5 年的基本演变方向。

首先，社会网络分析方法的引入为共被引分析带来了广阔的发展空间。2002 年，Otte 等^[9]明确提出，社会网络分析是研究信息科学的有力策略，可以成功地应用于引用和共被引网络、合作网络、社

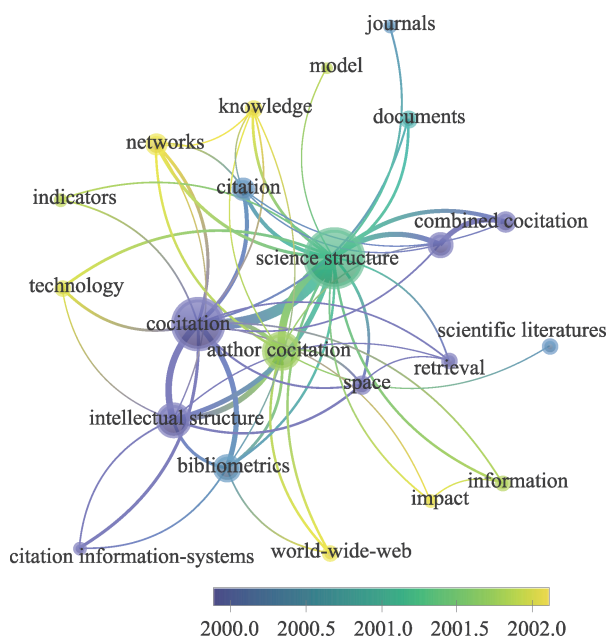


图 4 共被引分析方法起步阶段研究主题

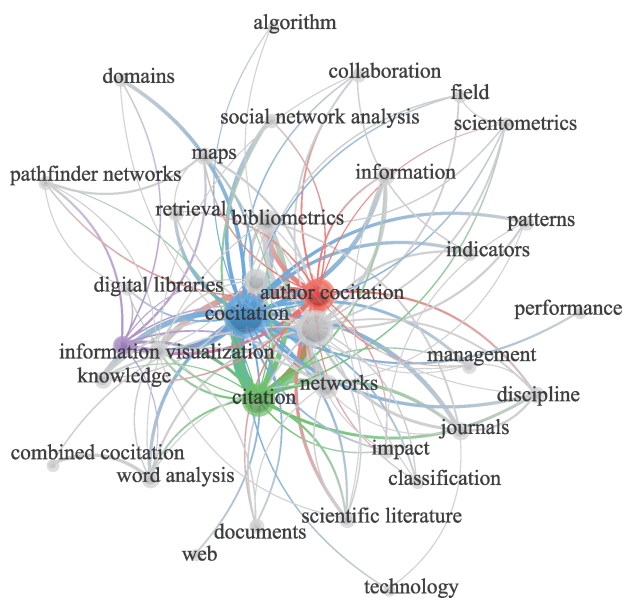


图 5 共被引分析方法起飞阶段(2006—2010)研究主题

会网络和互联网。在共被引分析迭代创新的起步阶段仅有个别研究运用了社会网络分析方法，而就在刚刚进入起飞阶段的 2006 年，利用社会网络分析研究共被引网络的成果开始明显增多。相关研究主要描绘知识结构、作者群落，绘制主题地图，追踪发展路径、学科分化，并将如寻径网络等的修剪算法结合起来，用来解释科学的宏观和微观结构。

其次，可视化技术成为共被引分析领域的一大研究热点。陈超美成为该领域的代表人物，其成果

不断改进知识领域可视化技术，并具象化“研究前沿”和“知识基础”这一对概念，通过它们随着时间而变化甚至相互影响的情况来达到显示学科发展趋势的目的，并将相关技术集成于 CiteSpace II^[10]。

最后是科学计量学转向。除了作为共被引分析上位类的文献计量学，以科学计量学为研究主题的成果逐渐增多。这一现象意味着，共被引分析逐渐从产生该方法的基本载体脱离，开始以解决科研发展过程出现的实际问题为导向开展研究。从图 5 可见，科研合作、针对学科、主题和领域的分析已经成为新的细分领域。在社会网络分析和可视化技术的助力下，共被引分析对科学发展和研究中遇到的具体问题具备了更强、更直观的分析能力、手段和解释力，成为解构和解决科研产出、人员和过程问题的有力工具。

从学科领域来看，2006—2010 年，共被引分析扩散广度增加了 20 个学科，涉及教育学、环境科学生态学、地理、自动化控制、政府法律、材料科学、冶金工程、哲学、生命科学、生物医学、国际关系等。多学科的采纳、方法的引入与技术的迭代更新，促使共被引分析研究跨越拐点，达到了成功扩散的临界值。

2.2.3 起飞阶段:2011—2017 年,加速迭代

该阶段的成果占总成果数的 55.82%，被引量占总被引量的 19.04%。由图 6 可见，这一阶段涌现出

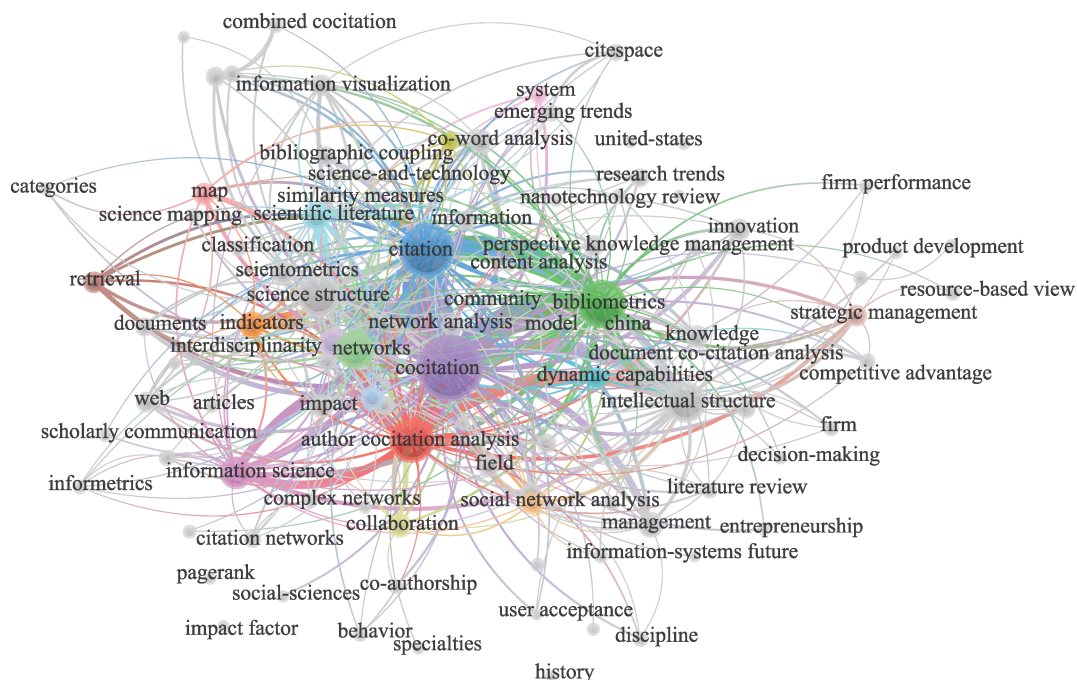


图 6 共被引分析方法起飞阶段(2011—2017)研究主题

丰富的研究主题，可称之为起飞阶段的加速迭代期。然而迭代创新的基础始终非常稳定，共被引领域的研究始终围绕 6 大主题展开，它们是共被引分析方法的母领域，即文献计量学和引文分析；主要研究问题，即揭示知识结构和学科结构；方法本身，即共被引及更为专门的作者共被引分析。基于共被引分析的 6 大基本问题，新涌现的研究主题主要在两个方面。首先，研究更加关注与研究前沿、创新、预测、新兴趋势、交叉学科有关的问题。此阶段，对于共被引分析方法的应用已不局限于领域热点分析，而是通过分析方法的改进、与其他方法的融合等方式，越来越多地被应用于对学科研究前沿和新兴领域的探索与预测，并且关注到交叉学科领域与科研创新的关系。其次，在方法层面上，较多地出现了与共词分析、其他引文网络分析、复杂网络和社会网络分析、内容分析、文献耦合等方法的融合研究。

2011—2017 年，共被引分析扩散广度增加了 28 个学科，涉及的学科包括大气科学、神经科学、公共环境卫生、物理、地质、水资源、能源、细胞生物学、电信、历史、文学、兽医学、体育科学、遥

感、建筑技术、热力学、动物学等，学科扩散稳步加速。

2.3 迭代创新路径

利用主路径分析算法（main path analysis）获得共被引分析方法迭代创新路径，如图 7 所示。主路径分析测算单向无环网络中，每条路径和节点被网络中其他路径与节点需要的程度，所得结果被称为某条弧或节点的遍历权值。在引文网络中，节点是研究成果，弧是成果间的引证关系。因此，弧表达了知识传输的路径与方向。提取所含的弧的遍历权值之和最高的路径，就是创新扩散的主路径。由于主路径只是创新扩散的骨架结构（backbone），所能提供的信息过于单薄，因此本研究以主路径上略低于最低遍历权值的值作为阈值，兼顾路径清晰性与内容丰富性，从原始引文网络中提取迭代创新路径。以下根据该路径，按照迭代创新的类型，梳理共被引分析迭代创新的发生过程。

1) 理论奠基

Small^[11-16]不但提出了共被引分析方法，而且在该领域持续发表成果，对于该方法的方法论、认识

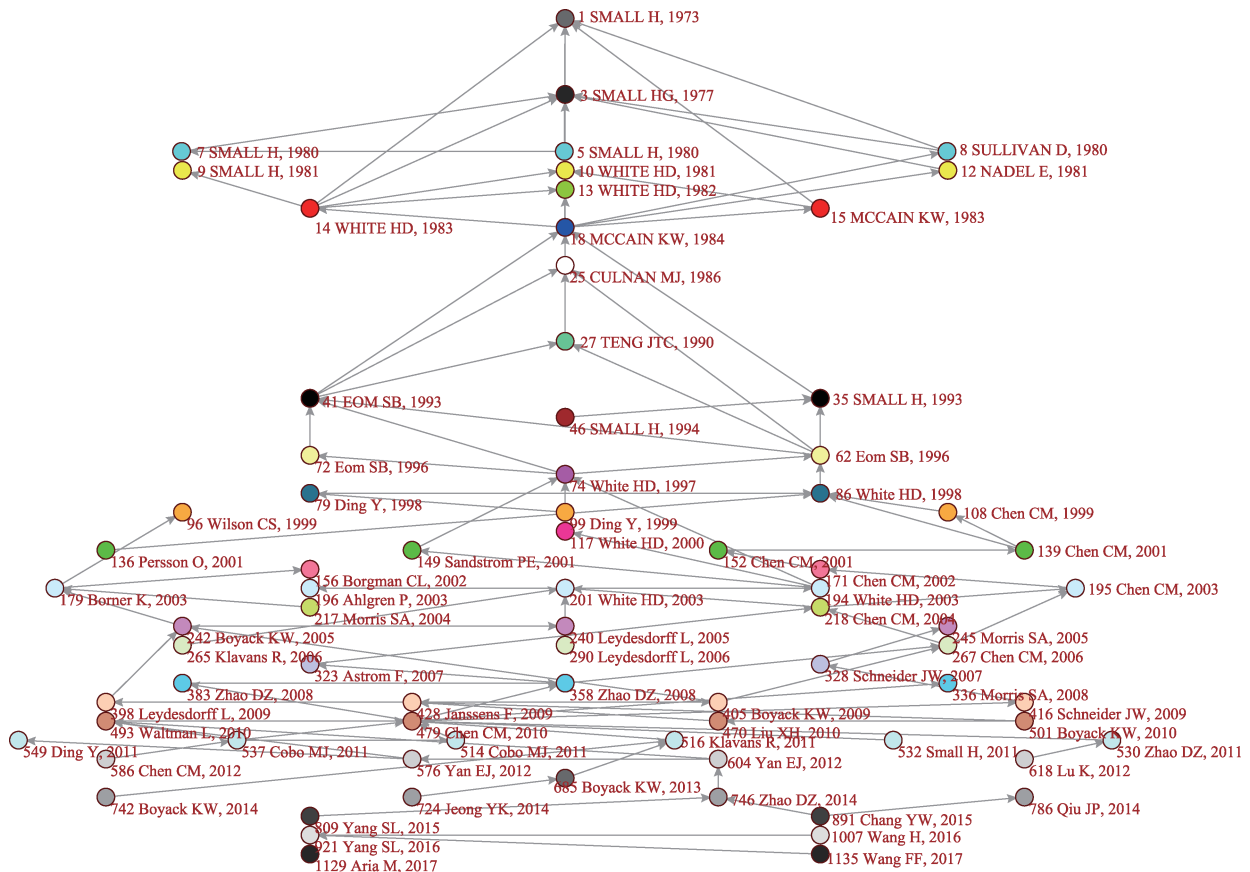


图 7 共被引分析方法迭代创新路径

论、概念、模型、图谱、作用、应用，与其他技术的结合，以及有效性验证等方面做出了基础性研究。他进一步阐述了共被引分析的概念基础；应用共被引分析识别出无形学院；通过问卷调查发现，运用共被引模型得出的某领域重要成果和作者与专家的感知大致相符；提出库恩的范式革命将可以通过引文数据表现出来，并通过个案进行实证；解析共被引图谱，认为较为松散的部分意味着较为开放的主题，知识拓展有着较多的可能性，致密的网络结构则可能代表完结或已被充分讨论的知识结构；首次尝试将共被引聚类技术与引文上下文分析结合，验证利用该技术能够在没有专家参与的情况下得到对专业知识的表征，因此认为文献计量能够真正成为认识论的分支。

2) 理论迭代创新

不同节点类型的共被引分析被提出。作者共被引分析 (author co-citation analysis, ACA)^[17]、期刊共被引分析 (journal co-citation analysis)^[18]、专利共被引分析 (patent co-citation analysis)^[19]、主题和类的共被引 (cocitation of classes and categories)^[20]、国家共被引 (national co-citation)^[21]、关键词共被引 (keyword co-citation)^[22] 分别于 1981 年、1991 年、1999 年、2004 年、2009 年、2010 年被相继提出。与文献共被引分析的提出时间间隔 8~37 年。其中，作者共被引分析由 White 和 Griffith 提出，两人运用该方法可视化信息科学领域作者群落，认为作者共被引分析将“作者”作为与“文献”一样的分析学科领域的基本单元，为理解学科的知识结构提供了新技术；1982 年，他们采集多学科数据进行实证，证实了作者作为知识空间标识的稳定性^[23]。这一研究为作者共被引分析在不同学科领域的应用提供了范例。随后，作者共被引分析成为了共被引分析中比重最大的子领域，并延续至今，其成果数占总成果数的三分之一强。发文量自 1993 年开始阶段性升高，而被引量自 1995 年以后持续升高，1998 年以后，每年成果数都占据总成果数的相当比重。

共被引分析概念与原理向超链接网络等其他领域移植。受到共被引概念的启发，共链分析 (co-linking analysis)^[24] 和共提及网络 (co-mention network)^[25] 分别于 2003 年和 2016 年被提出。前者利用两个网站共同被其他网站链接的关系，揭示网站间的集群结构，后者用于分析公司网络。在此之前，Prime 等^[26] 于 2002 年仿照“co-citations”提出

“co-citations”，指出在新出现的网络计量学领域，学者们正在将引文分析中的概念体系移植到超链接网络分析中，并认为 Web 共被引似乎是描述 Web 主题的有效途径；但也提醒同行，在移植过程中会有很多误解，因此在解释分析结果时需要采取一些防范措施。

3) 方法迭代创新

(1) 分析流程与分析单元的讨论。McCain^[27] 于 1990 年将作者共被引分析总结为 6 个步骤。此后相关研究大都遵循该基本流程。而新世纪出现了有关作者共被引分析方法上的争论。Zhao^[28] 于 2006 年对仅利用第一作者进行共被引分析的传统方法与全作者共被引分析进行对比，发现二者的主要区别是，全作者共被引分析结果包含更连贯的作者群体，因此可获得更清晰的可视化图谱，而第一作者共被引分析可以细分更多的专业领域。Eom^[29] 于 2008 年通过实证认为全作者共被引分析优于第一作者共被引分析，因其可以捕获领域内所有有影响力的作者。为提供更多的方法和经验证据，Schneider 等^[30] 于 2009 年进一步比较了第一作者和全作者分析单元，以及传统的多变量方法和 Drexel 方法所形成的不同组合对图谱结果的影响；结果表明，全作者的基于 Drexel 的方法使得作者共被引在地图中更易集聚成组。

(2) 相关研究方法的融合。2009 年，Chang 等^[31] 证明了基于专利、引用、共被引和聚类的方法工具在识别新兴、高影响力技术集群和趋势方面的有用性。2014 年，Jeong 等^[32] 提出基于文本内容的作者共被引分析 (content-based author co-citation analysis)。2014 年，Qiu 等^[33] 归纳了 5 种作者共现网络，采用社会网络分析和层次聚类方法识别子网络，QAP 相关性检测发现，虽然这 5 种网络都具有揭示科学知识结构的能力，但所揭示的结构是不同的。2016 年，Bu 等^[34] 利用引文发表时间和引文关键词提高作者共被引分析及可视化的精度。2017 年，Wang 等^[35] 通过对比作者共被引、合作、作者文献耦合、作者直接引用和作者关键词耦合网络之间的差异，挖掘潜在的合作者。

4) 技术迭代创新

(1) 技术迭代对方法的简化。例如，在迭代创新起步阶段，共被引分析依赖多维尺度、聚类分析和因子分析等多元统计分析进行聚类 and 可视化，计算中存在如何处理矩阵对角线数值、采用何种相关系数和距离测度方式等问题，复杂的步骤和降维过

程容易导致信息失真。1999年,Chen^[36]将认知心理学中用于分析语义关系的寻径网络技术(pathfinder network scaling, PFNETs)^[37]用于作者共被引分析。共被引网络中作者对之间的边权,仅用二者最高的共被引次数表示,即用最突出的网络关系简化了作者共被引网络,并以高中心度作者定义关键作者,用作者与关键作者之间的连接定义专业,用关键作者之间的连接把专业连成学科。由于该技术利用的是原始计数矩阵,无需利用Pearson相关系数等生成相关或距离矩阵,因此这一技术迭代简化了ACA分析步骤,并且能够分析更大规模的共被引网络,分析结果也更为可信。

(2) 可视化技术的不断改进。White等^[38]于1997年以“Visualization of literatures”为标题明确提出了基于文献的可视化技术,认为基于术语共现和出现,新的模型可以给出文献间关联的视觉线索,并于次年与McCain一起采用可视化的方式分析了信息科学领域作者的结构及分布情况^[39]。紧随其后,Small^[40]于1999年发表“Visualizing science by citation mapping”。此3人共同开启了共被引分析的可视化研究。丁鹰和陈超美基于White等^[38-39]和Small^[40]提出的文献可视化研究,在可视化领域发表了一系列成果。其中,陈超美在1999年的文献中运用三维虚拟现实演示了引文及共被引逐年增长的过程^[41],这一研究及相关研究促成了CiteSpace的诞生。

(3) 共被引聚类技术的迭代创新。为揭示更大规模的学科结构,Small等^[42]于1985年提出可变水平聚类(variable level clustering),以不断调整聚类临界值的方法来消除不同学科文献被引率不同造成的分析结构不平衡,提出以类聚类(clustering of clusters)以便形成学科的大类、超类。1997年,Lin^[43]尝试将源自计算机科学与神经科学的自组织映射技术(self-organized mapping, SOM)^[44]应用到作者共被引分析中,通过自学习过程进行降维和自动聚类。2010年,Chen等^[45]提出的多视角共被引分析技术(multiple-perspective cocitation analysis)通过集成网络可视化、谱聚类、自动聚类标记和文本摘要提高了作者共被引和文献共被引结果的可解释性和可探索性,说明了集成的和交互式的共被引可视化技术有利于进行探索性分析。2012年,Boyaack等^[46]利用全文本分析(full-text analysis)将参考文献间的邻近性纳入模型,提高共被引聚类的准确性。2014年,Jeong等^[47]利用引文上下文(content-

based)开展基于文献内容的作者共被引分析,测度作者之间的相似性,实证发现该方法在揭示信息科学知识结构的子领域时比传统的ACA能够提供更多细节。

5) 应用领域拓展

(1) 在图书情报学已有领域加入共被引分析技术,是生成共被引分析新的增长点的有效方式。例如,利用作者共被引改进信息检索策略;利用共被引聚类对检索结果过滤;采用作者共被引分析和可视化技术,结合引用和共被引分析方法,建立、更新和维护术语词表,或设计能够生成实时交互作者地图的AuthorLink系统^[48];利用越来越先进的共被引聚类技术确定科学前沿,对信息科学家范式重新描绘等。

(2) 相关学科对共被引分析方法的采纳不但大大增加了创新扩散广度,而且加速了迭代创新周期。例如,在起步阶段,运筹管理学领域主要运用共被引分析方法揭示管理信息系统(MIS)和决策支持系统(DSS)的研究现状和存在问题^[49-50];而在起飞阶段,共被引分析主要被应用于对企业动态能力(dynamic capabilities)、企业吸收能力(absorptive-capacity)理论的分析^[51-52]。该领域是资源和战略管理理论框架中一个充满活力的研究领域,对于这样一个复杂而重要的领域,研究者们需要借助共被引分析方法构建一个清晰的概念地图。再如,2015年,Zupic等^[53]结合引文分析、共被引分析、文献耦合、共词分析等文献计量学方法为管理和组织学者提供了一个文献计量研究的工作流程,以补充元分析和定性评价方法的不足,增强该领域论文评价的严谨性,减轻同行评议的偏见。

(3) 应用研究对理论探索的推动。例如,2009—2010年有一批成果较为集中地研究跨学科和大型科学网络,在一定程度上揭示了科学发展与创新机制。Leydesdorff^[54]认为利用科学文献将学科和亚学科结构解析出来是科学计量学的核心目的之一,通过实证证实了科学映射地图可以是大规模的和可靠的。Porter等^[21]采用国家共被引、文献共被引等方法验证了人们认为科学研究正在变得越来越跨学科的猜测,其科学地图提示跨学科研究主要发生在相邻学科领域之间,并有适度的远距离认知领域的链接。Small^[55]论述了共被引上下文和类比在基于科学地图的跨学科研究中的作用。研究表明,学科之间的联系往往基于不同学科作者对科学领域中相似问题的看法,共被引上下文提示这些跨学科链接提供

了科研创新的重要而成熟的机会。

2.4 跨学科迭代创新

共被引分析方法的相关研究分布在81个研究领域,即2017年的扩散广度为81。图8为扩散广度增长曲线。表2展示了共被引分析方法主要的扩散领域、采纳时间及扩散延时。从“占总成果数百分比”可知,图书情报学和计算机科学不仅是共被引分析方法产生的源头,也是该领域迭代创新的活跃学科。另外值得注意的是,这两个学科的成果大部分为跨学科研究,只属于图书情报学的论文仅占两学科论文总量的14%,仅属于计算机科学的占9%,而跨学科论文占77%。此外,商业与经济学的研究也较多,其他相关研究广布于科学技术、社会科学和人文艺术领域。

将跨学科引用量不小于10的学科领域作为节点,可视化共被引分析方法跨学科迭代创新网络。由于图书情报学和计算机科学的结果大部分重合,因此将这两个学科合并,标注为节点ISLS,CS,如图9所示。该网络规模为15,箭头方向为创新扩散

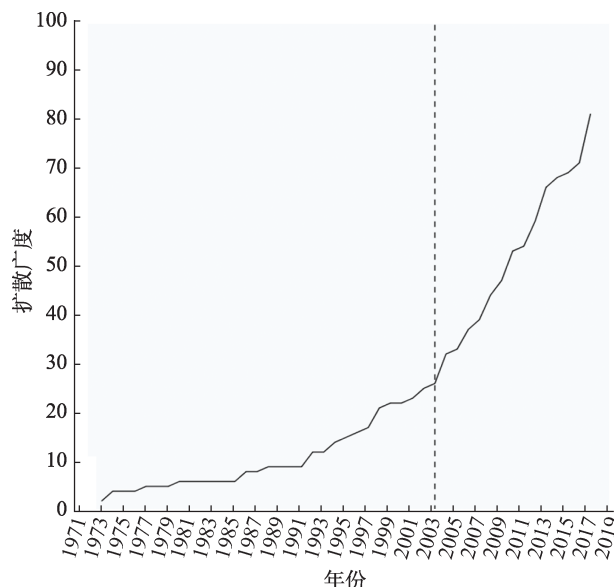


图8 共被引分析方法扩散广度增长曲线

方向。为明确各学科在迭代创新中发挥的具体作用,将所有学科分为3种类型:创新源发学科、创新助力学科和创新采纳学科。类型划分依据以下规则:①出度大于入度的节点为创新源发学科;②出

表2 共被引分析方法学科领域扩散情况表

序号	研究领域	成果数	占总成果数百分比/%	研究领域	采纳年份	扩散延时/年
1	Information Science Library Science	643	56.26	Information Science Library Science	1973	0
2	Computer Science	638	55.82	Computer Science	1973	0
3	Business Economics	192	16.80	Science Technology Other Topics	1974	1
4	Engineering	61	5.34	Social Sciences Other Topics	1974	1
5	Science Technology Other Topics	33	2.89	History Philosophy of Science	1977	4
6	Social Sciences Other Topics	32	2.80	Psychiatry	1980	7
7	Public Administration	25	2.19	Business Economics	1986	13
8	Operations Research Management Science	24	2.10	Operations Research Management Science	1986	13
9	Education Educational Research	20	1.75	Mathematical Methods in Social Sciences	1988	15
10	Environmental Sciences Ecology	18	1.58	Psychology	1992	19
11	Psychology	18	1.58	General Internal Medicine	1992	19
12	Communication	12	1.05	Research Experimental Medicine	1992	19
13	Mathematical Computational Biology	12	1.05	Public Administration	1994	21
14	Biochemistry Molecular Biology	11	0.96	Fisheries	1994	21
15	Biotechnology Applied Microbiology	11	0.96	Engineering	1995	22
16	Health Care Sciences Services	11	0.96	Sociology	1996	23
17	Mathematics	10	0.88	Mathematics	1997	24
18	Sociology	8	0.70	Communication	1998	25
19	Chemistry	7	0.61	Health Care Sciences Services	1998	25
20	Energy Fuels	7	0.61	Medical Informatics	1998	25
21	Medical Informatics	7	0.61	Pharmacology Pharmacy	1998	25
22	Neurosciences Neurology	7	0.61	Gastroenterology Hepatology	1999	26

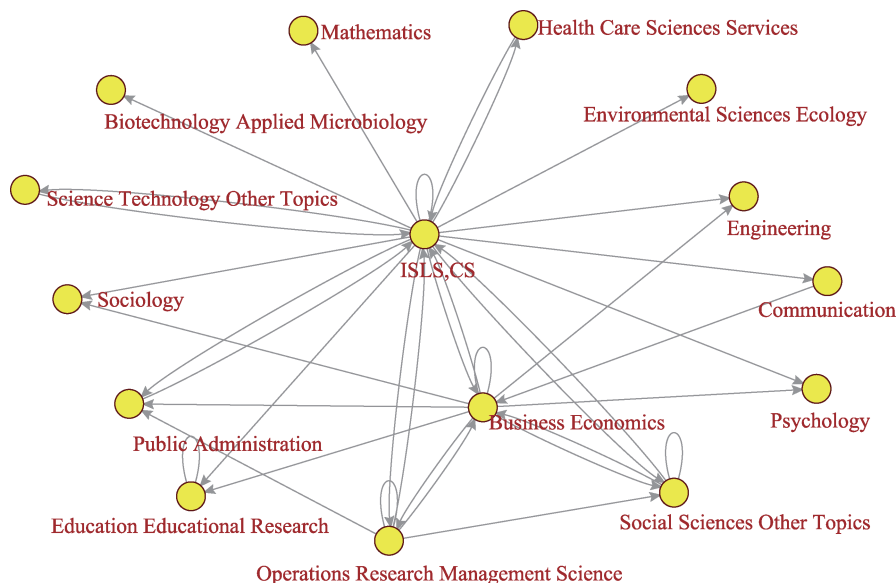


图 9 共被引分析方法跨学科迭代创新网络

表 3 共被引分析方法学科领域创新类型分布

创新类型	学科领域
创新源发学科	ISLS/CS; Business Economics; Operations Research Management Science
创新助力学科	Science Technology Other Topics; Social Sciences Other Topics; Communication; Health Care Sciences Services; Public Administration
创新采纳学科	Engineering; Education Educational Research; Environmental Sciences Ecology; Psychology, Biotechnology Applied Microbiology; Mathematics; Sociology

度为零的节点为创新采纳学科；③其余为创新助力学科。据此，得出 15 个学科领域的创新类型分布，如表 3 所示。

创新源发学科除了图书情报学和计算机科学，还有商业与经济学和运筹与管理科学。这 3 组学科领域的创新扩散路径形成双向闭环，迭代创新在这 3 个学科领域之间不断发生，并向其他学科输出新的理论、方法与技术。其中，图书情报学和计算机科学是创新扩散网络的核心，其成果被所有其他 14 个学科领域所采纳，另有 6 个学科的迭代创新成果反哺该领域，其出度入度比为 14/6。商业与经济学和运筹与管理科学为创新扩散网络的亚核心，出度入度比分别为 8/4 和 4/2。

创新助力学科有科学技术其他主题、社会科学其他主题、公共管理、传播学、保健科学与服务 5 个。其中，前 3 个学科是与创新源发学科密切相关的学科领域，二者之间拥有很多共同的研究主题，如学科的兴起与发展、科技创新与涌现、专利共被引方法与技术等。而传播学所涉及的传播网络、学术传播过程及学术传播媒介研究，以及保健科学对医学信息学、生物信息学、医药情报，医学文献检

索等主题的研究，为共被引分析提供了相关学科理论与分析视角、新的数据来源、分析单元、网络类型和新的研究问题。

创新采纳学科有 7 个：工程学、教育学、环境科学与生态学、心理学、生物技术与应用微生物学、数学、社会学。这些学科领域采用共被引分析方法与技术帮助解决核心研究领域转移、研究主题可见度、不同流派的交叉情况、作者分群、学科专业的演变与衰落、判别新动向和新进展、网页排名算法、节点相似度算法、空间优先连接模型构建等问题。

3 结 论

通过以上分析可知，在科研领域，迭代创新除了具有其一般特征，还有一些独特之处，归并总结为以下几点。

(1) 对原始创新核心内容的不断调用，以及对相关知识的引进与整合。共被引分析方法迭代创新的基础非常稳定，以文献计量学、引文分析、共被引分析、作者共被引分析的概念、原理、方法为迭

代基础,面向揭示科学、知识结构,探索科研规律等研究问题,通过吸纳、整合已有或新的相关理论、方法与技术推进迭代创新。例如,社会网络分析方法的引入使得共被引分析能够回到基于共被引网络研究的初衷,而不再主要依赖多元统计方法,实现方法论上的跃迁。

(2) 面向问题反复调整试错迭代。从起飞阶段开始,共被引分析研究的焦点不再是单纯的理论与方法,而是越来越多地面向解决科研过程中不断凸显的问题,包括挖掘潜在合作者、判定新兴领域、预测研究趋势、改进科研评价方法等。例如,为了更清晰、准确地可视化学科领域,一方面,共被引分析本身的方法与技术不断改进,另一方面,共被引分析方法与其他相关方法的融合也被不断尝试。

(3) 累积微创新产生大变革。持续的迭代创新和更多的应用场景使得原始创新的理论原理和应用边界愈加清晰、明确,一旦有新的相关问题亟待解决,该项创新就易于催生新的理论与方法。例如,网络计量学的基础理论之一即是共被引分析,它是共被引分析概念、原理在网络技术迅猛发展和网络资源激增的情景下催生的一种新型计量学。

(4) 在迭代创新形成规模之前存在一个“沉寂期”。在此期间,迭代创新成果较少,但是其中包含重要创新,见图7中1993年之前的“塔顶”部分。该部分包含 Small 和 White 等的多篇重要成果,为共被引分析设定了在理论、方法、技术和应用方面不断迭代的基本方向。从图3的累积成果量来看,沉寂期是迭代创新的开始部分,此时只有个别研究人员选定该项创新,其迭代创新成果增长曲线接近线性增长或指数增长的初始阶段。然而,虽然沉寂期在形态上不显著,相关成果也未充分体现其影响力,但确是一项科研创新能否顺利扩散和迭代的关键期。分析起步阶段的特征可知,该阶段的成果数约为总成果数的1/5,而其被引量占总被引量的1/2以上。

(5) 学科扩散促进迭代创新。2003年以后共被引分析的学科扩散明显加快,而迭代创新在2006年进入起飞阶段,并于2010年达到成功扩散的理论临界值。至于创新成果在学科领域间的加速扩散是否是该项创新能够顺利进行迭代的原因还需要更深入的研究,但是可以推断,学科扩散能够促进迭代创新。多学科理论与方法的引入,以及应用场景的增加都会为迭代创新提供新的动力。

(6) 不同学科领域在迭代创新中的“角色”不

同。共被引分析迭代创新有3个创新源发学科、5个创新助力学科和7个创新采纳学科。3种不同类型的学科领域在跨学科迭代的拓扑结构中占有不同地位,分别拥有扩散中心、亚中心,以及互惠和单向接收的位置与关系。学科领域间多样的交流形式,保证了迭代创新的可持续性。例如,图书情报学和计算机科学是共被引分析的源发学科。其中,图书情报学提出问题并设计解决问题的方法与指标,而计算机科学提供算法与技术,整合方法步骤,实现分析平台。

(7) 持续的迭代创新有赖于外部环境的支撑。1964年,科学引文索引SCI的诞生为共被引分析提供了数据与检索平台。2004年问世的Scopus数据库为全作者共被引分析提供了便利。学术文献搜索引擎Google Scholar和CiteSeer的出现进一步开放了引文数据的获取途径,开辟了网络环境下共被引分析的研究空间。近年来,随着全文数据资源的可获得性以及自然语言处理技术的不断提高,基于文本内容的共被引分析^[47,56]研究正逐步展开。

4 展望

对科学发展构成关键事件的原始创新在科研成果中仅占很小比例,往往可遇而不可求,而迭代创新包括以原始创新为基础的各种类型创新,是科研创新的主要形式。大量的迭代创新不仅促进科学研究的深入发展,而且为解决新问题提供了不同知识组合的尝试,并进一步激发原始创新的产生。因此,开展科研领域迭代创新研究,是探索科研创新规律的重要内容。本文仅以案例研究对科研领域迭代创新的基本特征进行了初步探索,更具一般性的结论有待于未来更充分和深入的研究。

参 考 文 献

- [1] 王建芳,冷伏海. 共引分析理论与实践进展[J]. 中国图书馆学报, 2006, 32(1): 85-88.
- [2] Kessler M M. Bibliographic coupling between scientific papers [J]. American Documentation, 1963, 14(1): 10-25.
- [3] Small H. Co-citation in scientific literature: A new measure of the relationship between two documents[J]. Journal of the American Society for Information Science, 1973, 24(4): 265-269.
- [4] 张腾,王迎军. 迭代式创新的研究与实践发展[J]. 现代管理科学, 2016(10): 100-102.
- [5] Fitzgerald E, Wankerl A, Schramm C. Inside real innovation: How the right approach can move ideas from R&D to market—and get the economy moving[M]. Singapore: World Scientific

- Publishing, 2011: 36-42.
- [6] 黛安娜·克兰. 无形学院——知识在科学共同体的扩散[M]. 刘珺珺, 顾昕, 王德禄, 译. 北京: 华夏出版社, 1988: 63-64.
- [7] 宋歌. 学术创新的扩散过程研究[J]. 中国图书馆学报, 2015, 41(1): 62-75.
- [8] 罗杰斯. 创新的扩散[M]. 辛欣, 译. 北京: 中央编译出版社, 2002: 86-90.
- [9] Otte E, Rousseau R. Social network analysis: a powerful strategy, also for the information sciences[J]. *Journal of Information Science*, 2002, 28(6): 441-453.
- [10] Chen C M. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(3): 359-377.
- [11] Small H. A co-citation model of a scientific specialty: A longitudinal-study of collagen research[J]. *Social Studies of Science*, 1977, 7(2): 139-166.
- [12] Small H, Greenlee E. Citation context analysis of a co-citation cluster: Recombinant-DNA[J]. *Scientometrics*, 1980, 2(4): 277-301.
- [13] Small H. Co-citation context analysis and the structure of paradigms[J]. *Journal of Documentation*, 1980, 36(3): 183-196.
- [14] Small H. The relationship of information-science to the social-sciences: A co-citation analysis[J]. *Information Processing & Management*, 1981, 17(1): 39-50.
- [15] Small H. Macro-level changes in the structure of co-citation clusters: 1983-1989[J]. *Scientometrics*, 1993, 26(1): 5-20.
- [16] Small H. A Sci-Map case-study: Building a map of AIDS research [J]. *Scientometrics*, 1994, 30(1): 229-241.
- [17] White H D, Griffith B C. Author cocitation: A literature measure of intellectual structure[J]. *Journal of the American Society for Information Science*, 1981, 32(3): 163-171.
- [18] McCain K W. Mapping economics through the journal literature: An experiment in journal co-citation analysis[J]. *Journal of the American Society for Information Science*, 1991, 42(4): 290-296.
- [19] Moge M E, Kolar R G. Patent co-citation analysis of Eli Lilly & Co. patents[J]. *Expert Opinion on Therapeutic Patents*, 1999, 9(3): 291-305.
- [20] Moya-Anegón F, Vargas-Quesada B, Herrero-Solana V, et al. A new technique for building maps of large scientific domains based on the cocitation of classes and categories[J]. *Scientometrics*, 2004, 61(1): 129-145.
- [21] Porter A L, Rafols I. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time[J]. *Scientometrics*, 2009, 81(3): 719-745.
- [22] Su Y M, Hsu P Y, Pai N Y. An approach to discover and recommend cross-domain bridge-keywords in document banks[J]. *The Electronic Library*, 2010, 28(5): 669-687.
- [23] White H D, Griffith B C. Authors as markers of intellectual space: Co-citation in studies of science, technology and society [J]. *Journal of Documentation*, 1982, 38(4): 255-272.
- [24] Faba-Pérez C, Guerrero-Bote V P, de Moya-Anegón F. Data mining in a closed Web environment[J]. *Scientometrics*, 2003, 58(3): 623-640.
- [25] Park J, Seok S, Park H W. Web feature and co-mention analyses of open data 500 on education companies[J]. *Journal of the Korean Data Analysis Society*, 2016, 18(4): 2067-2078.
- [26] Prime C, Bassecouard E, Zitt M. Co-citations and co-sitations: A cautionary view on an analogy[J]. *Scientometrics*, 2002, 54(2): 291-308.
- [27] McCain K W. Mapping authors in intellectual space: A technical overview[J]. *Journal of the American Society for Information Science*, 1990, 41(6): 433-443.
- [28] Zhao D Z. Towards all-author co-citation analysis[J]. *Information Processing & Management*, 2006, 42(6): 1578-1591.
- [29] Eom S. All author cocitation analysis and first author cocitation analysis: A comparative empirical investigation[J]. *Journal of Informetrics*, 2008, 2(1): 53-64.
- [30] Schneider J W, Larsen B, Ingwersen P. A comparative study of first and all-author co-citation counting, and two different matrix generation approaches applied for author co-citation analyses[J]. *Scientometrics*, 2009, 80(1): 103-130.
- [31] Chang C K N, Breitzman A. Using patents prospectively to identify emerging, high-impact technological clusters[J]. *Research Evaluation*, 2009, 18(5): 357-364.
- [32] Jeong Y K, Song M, Ding Y. Content-based author co-citation analysis[J]. *Journal of Informetrics*, 2014, 8(1): 197-211.
- [33] Qiu J P, Dong K, Yu H Q. Comparative study on structure and correlation among author co-occurrence networks in bibliometrics [J]. *Scientometrics*, 2014, 101(2): 1345-1360.
- [34] Bu Y, Liu T Y, Huang W B. MACA: a modified author co-citation analysis method combined with general descriptive metadata of citations[J]. *Scientometrics*, 2016, 108(1): 143-166.
- [35] Wang F F, Wang X H, Yang S L. Mining author relationship in scholarly networks based on tripartite citation analysis[J]. *PLoS ONE*, 2017, 12(11): e0187653.
- [36] Chen C M. Visualising semantic spaces and author co-citation networks in digital libraries[J]. *Information Processing & Management*, 1999, 35(3): 401-420.
- [37] Housner L D, Gomez R L, Griffey D C. Pedagogical knowledge structures in prospective teachers: Relationships to performance in a teaching methodology course[J]. *Research Quarterly for Exercise and Sport*, 1993, 64(2): 167-177.
- [38] White H D, McCain K W. Visualization of literatures[J]. *Annual Review of Information Science and Technology*, 1997, 32: 99-168.
- [39] White H D, McCain K W. Visualizing a discipline: An author co-

- citation analysis of information science, 1972-1995[J]. *Journal of the American Society for Information Science*, 1998, 49(4): 327-355.
- [40] Small H. Visualizing science by citation mapping[J]. *Journal of the American Society for Information Science*, 1999, 50(9): 799-813.
- [41] Chen C M. Visualising semantic spaces and author co-citation networks in digital libraries[J]. *Information Processing & Management*, 1999, 35(3): 401-420.
- [42] Small H, Sweeney E, Greenlee E. Clustering the science citation index using co-citations. II. Mapping Science[J]. *Scientometrics*, 1985, 8(5-6): 321-340.
- [43] Lin X. Map displays for information retrieval[J]. *Journal of the American Society for Information Science*, 1997, 48(1): 40-54.
- [44] Kohonen T. Self-organized formation of topologically correct feature maps[J]. *Biological Cybernetics*, 1982, 43(1): 59-69.
- [45] Chen C M, Ibekwe-Sanjuan F, Hou J H. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis[J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(7): 1386-1409.
- [46] Boyack K W, Small H, Klavans R. Improving the accuracy of co-citation clustering using full text[J]. *Journal of the American Society for Information Science and Technology*, 2013, 64(9): 1759-1767.
- [47] Jeong Y K, Song M, Ding Y. Content-based author co-citation analysis[J]. *Journal of Informetrics*, 2014, 8(1): 197-211.
- [48] Lin X, White H D, Buzydlowski J. Real-time author co-citation mapping for online searching[J]. *Information Processing & Management*, 2003, 39(5): 689-706.
- [49] Teng J T C, Galletta D F. MIS research directions: A survey of researchers views[J]. *ACM SIGMIS Database*, 1991, 22(1-2): 53-62.
- [50] Eom S B. Mapping the intellectual structure of research in decision support systems through author cocitation analysis (1971-1993)[J]. *Decision Support Systems*, 1996, 16(4): 315-338.
- [51] Peteraf M, di Stefano G, Verona G. The elephant in the room of dynamic capabilities: Bringing two diverging conversations together[J]. *Strategic Management Journal*, 2013, 34(12): 1389-1410.
- [52] Dagnino G B, Levanti G, Minà A, et al. Interorganizational network and innovation: A bibliometric study and proposed research agenda[J]. *Journal of Business & Industrial Marketing*, 2015, 30(3-4): 354-377.
- [53] Zupic I, Čater T. Bibliometric methods in management and organization[J]. *Organizational Research Methods*, 2015, 18(3): 429-472.
- [54] Leydesdorff L, Rafols I. A global map of science based on the ISI subject categories[J]. *Journal of the American Society for Information Science and Technology*, 2009, 60(2): 348-362.
- [55] Small H. Maps of science as interdisciplinary discourse: Co-citation contexts and the role of analogy[J]. *Scientometrics*, 2010, 83(3): 835-849.
- [56] 步一, 刘天祯, 赵丹群, 等. 国外作者共引分析研究评述[J]. *情报杂志*, 2015, 34(12): 48-53.

(责任编辑 马 兰)