

论科学计量学在科学知识管理中的应用¹

邱均平¹ 董克¹ 马瑞敏²

(1. 武汉大学 信息管理学院, 武汉 430072; 2. 山西大学 管理与决策研究所, 太原 030006)

摘 要: 本文对科学计量学进行定义并阐述了科学计量学的研究内容, 归纳了应用于科学知识管理的科学计量学方法和工具的发展。科学计量学在科学知识管理中的应用主要包括科学知识结构分析和科学知识演化分析; 从作者网络和共词网络的角度论述了科学计量学如何实现科学知识结构分析, 从引文时序网络和作者、词等其它特征项网络角度论述了当前科学计量学实现科学知识演化分析的主要内容, 认为利用科学计量学能够有效实现科学知识管理。

关键词: 科学计量学; 科学知识管理; 知识结构; 知识演化

中图分类号: G350 **文献标识码:** A **文章编号:** (2014) 02-00091-106

0 引言

定量化和计算化的是社会科学研究未来发展的重要趋势^[1]。微软研究院副主席 Tony Hey 认为, 数据密集型科学发现作为科学研究的第四范式, 将对科学研究产生巨大影响^[2]。在海量数据环境下, 科研管理所采取的具体方式也必然会发生相应变化。科学的发展过程充满不确定性, 科研管理人员和科技政策专家一直都在不断探索科学知识的发展规律; 实现对科学知识的管理是科研管理工作的核心任务。在海量数据环境下, 利用科学计量学的理论和方法管理科学知识, 是制定科学、合理、有效的科技政策的前提, 也是提升科研管理效率的必然途径。

1 科学计量学概述

1.1 科学计量学的概念和研究内容

科学研究系统发展到高度分化和综合的阶段, 逐渐形成从科学研究到技术开发再到生产应用的多层次、多序列的复杂结构, 采用定量化的方法对科学事业进行管理能够最大限度地发挥科学系统的功能。1969 年, 前苏联学者纳里莫夫和穆里钦科首次提出科学计量学这一称谓用以表述采用定量方法对科学事业进行的研究^[3]。在由文献计量学、科学计量学、信息计量学、网络计量学和知识计量学所

¹基金项目: 国家社会科学基金重大项目 (11&ZD152), 中国博士后基金资助项目 (2014M562069), 国家社科基金青年项目 (12CTQ026)。

作者简介: 邱均平 (1947-), 男, 湖南涟源人, 学士, 武汉大学信息管理学院教授, 博士生导师, 研究方向: 科学评价与科学计量, 竞争情报与知识管理, Email: jpqiu@whu.edu.cn; 董克 (1986-), 男, 江苏宜兴人, 博士, 武汉大学信息管理学院博士后, 研究方向: 科学计量与科学评价, 社会网络与引文分析, Email: dk8047@163.com; 马瑞敏 (1983-), 男, 山西榆次人, 博士, 山西大学管理与决策研究所副教授, 硕士生导师, 研究方向: 信息计量与科学评价。

构成的“五计学”学科体系中^[4]，科学计量学相对特殊，是直接以目的命名的计量学。可以认为，科学计量学是一门以描述科学发展过程，揭示科学发展内在机理，预测科学发展趋势，以科学管理工作提供支持依据为目的，通过定量分析方法研究科学事业及其主体和客体的应用性学科。

科学计量学领域最高级别的学术组织国际科学计量学与信息计量学学会（International Society for Scientometrics and Informetrics, ISSI）在其宗旨中认为，科学计量学的内容包括：科学技术和其他重要学术情报的定量研究；科学学、技术学、社会科学、艺术学和人文科学的定量研究；情报的产生、传播和使用的定量研究；图书馆、档案和数据库等定量研究，以及情报方面的数学模型研究^[5]。这种表述比较全面地概括了科学计量学的研究内容，国内一些学者对于科学计量学研究内容的解释也基本上被涵盖在 ISSI 上述宗旨中^[6]。

从本质上看，科学计量学的研究对象就是科学知识，可以从科学知识载体角度分析科学计量学的研究对象，主要包括学术论文、专利文献和其他形式的科学信息。例如，早期科学文献是科学计量学研究的主要对象^[7]；Narin 等学者则利用科学计量学的相关方法对专利文献进行了研究^[8]；网络环境下，越来越多的科学交流活动以实时或者非实时的方式在线进行，网络上的科学信息往往被归入信息计量学和网络计量学的研究范畴，但在线环境下产生的科学交流记录实际上也是科学计量学研究的重要对象。

1.2 应用于科学知识管理的科学计量学方法和工具

传统的科学计量方法主要有出版物统计、著者统计、词频分析、内容分析等，这些方法已经相对成熟。随着研究的深入和学科的不断交叉，共词分析、共被引分析、复杂网络和社会网络分析、多元统计分析、信息可视化等许多新的方法逐渐被引入到科学计量学研究中。

（1）共词分析。共词分析是通过分析在同一文本中的词或词对共同出现的形式，确定文本代表的学科领域中相关主题的关系，进而探索学科的发展的一种科学计量学方法。Calon 等人在 1983 年将共词分析法引入图书情报研究领域进行内容分析，通过高频主题词的聚类发现学科的研究热点^[9]。国外的共词分析研究分成了基于包容指数和临近指数，基于战略坐标和以数据库内容结构为特征的共词分析三个不同的阶段^[10]。Swanson 和 Smalheiser 利用单词词频统计方法挖掘隐藏于医学文献中的有价值的知识^[11]，开创了非相关文献的知识发现。目前，共词分析正逐步与自然语言处理等技术相结合拓展到语义分析层面^{[12]-[13]}。

（2）共被引分析。同时被引用的文献一般被认为在主题上具有一定的相似性，因此共被引强度可以用来测度文献内容主题的相似度^[14]。在文献的基础上，作者共被引^[15]、期刊共被引^[16]等方法也逐渐被提出。在采用此方法分析过程中，一般首先选定一定数量的文献（或作者、期刊等）作为分析对象，采集数据，构建共现矩阵，然后利用元统计分析方法将分析对象之间错综复杂的网络关系简化较少的若干类群之间的关系，最终以图形的方式展示^[17]；近年来，全文本分析技术也逐渐在共被引分析中推广应用^[18]。

（3）多元统计分析。科学计量学中应用的多元统计分析主要是降维分析技术，如主成分分析^[19]、因子分析^[20]和多维尺度分析^[21]。其基本思路都是从较为复杂的数据维度中抽取少数能够反映整体情况的主要维度进行结果解释。目前，主成分分析、多维尺度分析或因子分析的结果常与聚类分析和社

区发现结果相结合^[22], 这些方法在科学知识结构研究中应用广泛。

(4) 复杂网络和社会网络分析。复杂网络分析和社会网络分析方法的主要分析对象为网络数据, 由于表达网络数据的矩阵和科学计量中共现矩阵表现形式一致, 因此社会网络分析被广泛应用于科学计量学。社会网络分析可分为个体, 二方关系对, 三方关系组, 子图与多关系结构分析, 整体网络等五个层次^[23], 能够从宏观、中观和微观多个层面对科学知识结构进行揭示^[24]。

(5) 信息可视化分析。信息可视化是指利用计算机实现对抽象数据的图像化表示, 从而增强人们对这些抽象信息的认知, 其结果也常被称作科学地图或知识图谱^[25]。可视化分析的一个重要的突出特点在于, 对于具有一定规模的数据而言, 以图像的形式展现分析结果具有比其他方式更加良好的效果, 能够有效地为科学知识管理提供支持。

(6) 新的评价指标和 Altmetrics。2005 年, 美国物理学家提出 h 指数用来测度个人的科研绩效^[26], h 指数改善了以前的计量指标只关注数量而不关注质量的不足, 引发了新一波的评价指标研究热潮^[27]。同时, 针对传统学术成果重要性评价存在的缺陷, 社交网络环境下的 Altmetrics 成为一项重要的发展方向^[28]。Altmetrics 植根于社交网络, 其对象包括人、期刊、图书、数据集、报告、视频、源代码库和网页等各种类型的知识载体, 重点关注作品本身产生的影响。目前关于 Altmetrics 的研究已经引起了学术界的重视, 相关研究也将为在线环境下的科学知识管理提供重要支持^[29]。

科学计量学研究过程中, 作为分析对象的数据来源广泛, 如 ISI Web of Knowledge^[30]、工程索引 (Engineering Index, EI)^[31], Scopus^[32], 中文社会科学引文索引 (Chinese Social Sciences Citation Index, CSSCI)^[33]等, 在这些系统平台初步的统计分析结果基础上, 结合非嵌入性的分析工具所得到的科学计量学研究结果能够有效支持科学知识管理。

科学计量学中的常用工具依据专业性不断提升大致上可以分为四种类型。第一种类型是数据存储和简单的表格工具, 如 Microsoft Excel、Access 和 SQL 等, 由于具备一定的编程功能, 通过这些工具能够实现科学计量分析, 例如储节旺、郭春侠等阐述了利用 Excel VBA 实现共词分析的基本原理和实现方法^[34]。第二种是在科学计量学过程中需要对数据进行进一步的统计分析所采用的工具, 最常见的如 SPSS、SAS 和 Matlab 等。

第三类工具是交叉学科领域的专用软件, 如社会网络分析工具 Ucinet^[35], Pajek^[36]等。这类工具最初主要应用于人际网络分析, 由于科学知识交流中的大量网络和社交网络采取同样的数学表达形式, 因此这类软件也在科学计量学研究中被广泛使用。第四类软件主要包括了一些专门为进行科学计量学研究目的所设计的软件, 这些软件往往由专业的科学计量学研究人员开发, 例如 2011 年普赖斯奖得主, 瑞典计量学家 Olle Persson 开发的 Bibexcel^[37], Thomson Reuters 公司开发的文本挖掘软件 Thomson Data Analyzer^[38], 荷兰阿姆斯特丹大学的科学计量学教授 Loet Leydesdorff 开发的系列专门软件^[39], 美国印第安纳大学 Katy Börner 教授及其团队所开发的 Science of Science (Sci2)^[40]等, 这些工具可以在多个层面进行计量分析, 从而有效支持科学知识的管理。

2 利用科学计量学实现科学知识管理的途径一：科学知识结构分析

要实现科学知识管理，必须需要对科学发展有足够的认识。利用科学计量来进行科学知识结构的研究，是科学计量学服务于科学知识管理的一项重要内容。科学文献是科学发展的客观记录^[41]，科学文献的作者、关键词等特征项实体从更加精确的角度为科学知识的结构研究提供了途径。作为科学知识产生的主体，作者之间的知识关联网络类型包括合作、引用、共被引、耦合等多种类型，这些网络类型基本上涵盖了其他特征项实体如机构、期刊、国家等其他特征项的可能构建的网络类型；关键词以及在关键词基础上进一步提炼出来的主题则直观地反映了科学知识。本部分主要从这两种特征项来讨论利用科学计量进行科学知识结构的分析。

2.1 作者知识关联网络分析

作者是科学研究的主体，也是科学知识的创造者。作者之间因交流所产生的网络是知识体系的重要表现方式。如果将特征项之间的关联网络按照其是否加权和是否有方向对所有的网络进行分类，作者之间的关联网络基本上涵盖了可能出现的所有类型。作者之间的网络主要包括合作、引用、共被引和耦合等四种类型。

科学合作是大科学时代科学研究的一项重要特征，科学家的研究工作并不孤立，每个研究人员都是整个科学研究团体的成员，他们或许处于不同的地区，或隶属于不同的机构，但都在共同探索科学^[42]。进入 21 世纪以后，复杂网络的研究进一步推动了对作者基于大范围合作网络的研究，合作网络被广泛的用于挖掘科学的智力结构，许多新技术也被应用于研究合作网络的结构动力学^[43]，相关成果极大推动了对于科研创作规律以及学术交流模式的认知^[44]。美国情报学家 White 和加拿大贝尔实验室的学者 Nazer 对作者之间的社会网络和知识网络之间的相关性进行了研究^[45]，他们采集了一个名为 Globalnet 的科学家团体的社会交往数据^[46]，结合他们之间的合作、引用、共被引（统称为知识关联）数据发现，研究人员之间知识关联的产生主要依赖于知识内容的共享和交流，而非社会关联。Ding 以作者合作和引用网络进行为载体，深入探讨了科学合作和科学支持的问题^[47]，其研究结果表明高产作者的确偏好直接与具有相同研究兴趣点的人进行合作，但是在引用上面却没有对这个群体具有明显的偏好；高被引作者间和彼此合作的兴趣较低，但是在引用上却更加具有倾向。

作者之间的引用关联和文献之间的引用关联所形成的网络间最大的不同点在于，科学文献之间的引文网络是二值的，而作者之间的引用网络则是累计的。作者之间的引用关联网络是作者知识关联中唯一的有向关系网络，其中更包含了计量研究中另外两种重要的关联类型：共被引关联和耦合关联^[48]。作者之间的共被引关联是指两个作者同时出现与某篇文献的参考文献集合中，而作者之间的文献耦合关系是指两个作者同时参考了某一篇文章。这两种关联被从单元结构拓展到更大范围的群体结构时，就形成作者之间共被引网络和作者文献耦合网络。

在基于引用所产生的三类网络中，作者之间的互引网络被关注的较少，邱均平、王菲菲等曾利用作者之间的互引网络对《Scientometrics》杂志 1978 年至 2011 年的数据进行分析，通过网络结构挖掘发现，作者的互引关系能够有效地对作者进行分类，总结研究规律、寻找研究同行、发现研究热点^[49]。

作者共被引自产生以来就被广泛用于科学知识结构的探测,例如 White 和 McCain 的图情领域知识结构分析^[50], Chen 和 Lien 进行的 e-learning 领域的知识结构分析^[51]等。此外,对作者共被引分析方法本身的研究也一直都为学者们所关注,例如共被引分析中的相似性测度问题^[52],可视化问题^[53],网络环境下作者共被引分析^[54],全部作者共被引分析^[55]等。

作者之间的耦合网络是基于引用的另一种重要网络,赵党志和马瑞敏分别同时提出了利用作者之间的文献耦合网络进行科学结构分析的设想,并进行了实证研究^{[56]-[57]}。从网络形成的角度来看,耦合关系是最具拓展性的一类关系,构成两个作者之间耦合的衔接要素可以是多种资源实体属性。传统的作者文献耦合是作者双方基于利用同一篇文献,作者之间使用同一个词汇也可以产生耦合关联,作者同在某些出版物上发表自己创作成果也能够产生作者彼此之间的耦合关联;前者可以称之为作者关键词耦合或基于关键词的作者耦合,后者则可以称之为作者出版物耦合或基于出版物的作者耦合,不同类型的耦合关联研究是未来作者网络关联研究的重要发展方向^[58]。

许多新的分析技术也在作者网络分析中被广泛使用,例如,有学者将 Google Maps 等技术将作者标注于全球地图上,图 1 展示了信息科学领域地区之间的合作网络,从中可以直观地看到该领域的研究在全球的分布情况^[59]。网络结构分析法也常被用于科学合作的结构研究,常用的指标或分析包括合作网络的密度、可达性、中心性、小团体、核心-边缘结构等^[60]。这些指标在网络结构分析上的功能无可置疑,但是对于分析对象所处的环境和结果的解释,一些原理性的问题容易被忽视,例如距离作为一类重要的测度依据,在社会网络分析中最初主要针对社会关系的远近,作者之间的知识关联与社会关系存在明显区别,距离作为一类指标,其有效性就需要甄别^[61]。



图 1 全球信息科学领域合作网络

Fig.1 The global collaborative network of information science

2.2 共词网络分析

科学词汇一直都是科学计量学研究的重要内容,其中即包含了词频分布的研究,也包含了由词网络所反映的科学知识结构的研究。作为科学计量学三大基础理论之一的齐普夫定律深入揭示了科学文献词频的分布规律^[62], Newman 曾将齐普夫定律纳入幂率分布的框架下进行进一步的解释^[63]。通过统

计分析表达科学知识和核心内容的专业术语在某一研究领域的频次,能够有效反映该领域研究的主题和重点。

单纯的词频统计能够反映的内容相对有限,通过词与词之间共同出现产生的共词网络进行的科学知识结构研究受到学者们更多的关注。目前对共词网络的研究主要采用多元统计分析、网络结构分析和战略坐标图三种方法^[64],图2是中国图情领域博士论文关键词的共词网络,其中节点大小代表中心度的大小,连线的粗细代表共现强度,词之间的距离则代表了其关系的亲疏^[65]。

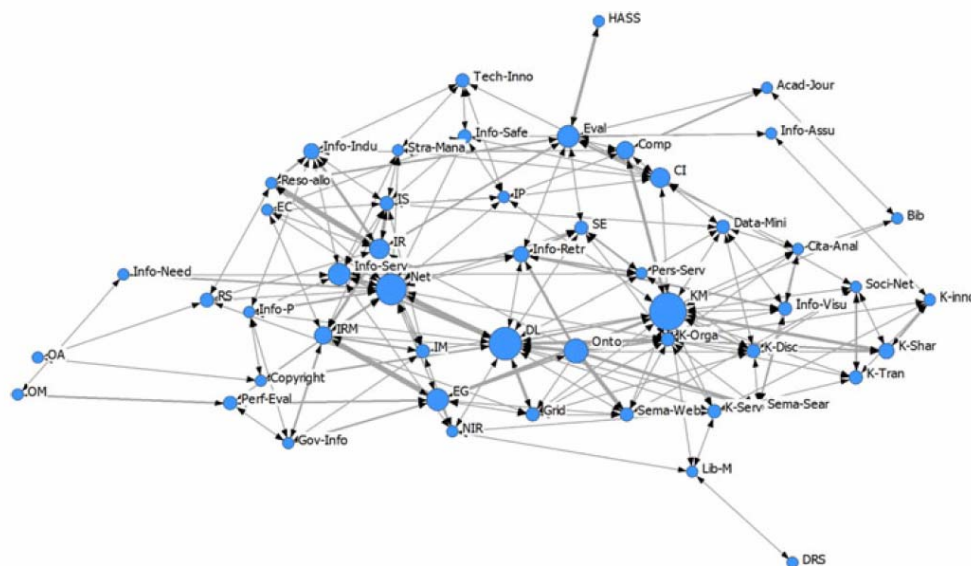


图2 图书情报领域博士论文关键词共词网络

Fig.2 The co-words network of library and information science dissertations' keywords

单独从词出发的分析无法判断其包含的具体语义,而共词分析可以从一定程度上解决这个问题;通过将词以及与它共现的词归结为一类,分析这个类反映的共同主题,可以确定词在分析对象集合中的具体含义。词共现只是一种简单的表现形式,可以结合其他方法进行更多的语义分析,例如采用因子分析方法将词进行聚类后进行因子解释,用因子来描述词集所包含的语义。目前应用于科学知识结构分析的共词分析主要存在如下问题:

首先,是分析基本单元的选择问题;共词分析对语义的揭示是通过和其它词之间的关系来进行判断的,因此词最终会划分到一个类中或者极为少量的主维度上解释,因此容易造成反应某些主题的语义无法识别和非重要的主题语义被识别的情况。

其次,词孤立的问题;对于某些含义较泛的词来说,比较容易出现该词与出现频次较少,语义内涵较弱的词共现,会造成所处语义环境相对薄弱,将词进行分门别类的时候,很容易由于其语义环境的薄弱性而归入不是很相关的类别中。

第三,词和词之间的距离问题;由于共现是以共同出现为基本分析假设,并没有考虑词和词之间是如何共现的,在同一个文档或资源集中,某些词之间是直接共现的,例如两个词直接在一起;某些则是间接共现的,例如出现在不同段落中,单纯的共现频率统计会造成频次相同而实际的语义关联强度不同的问题,因此无法反应词所包含的真实语义内涵。

采用主题图等方式加强共词分析的语义能力已经被提出^[66], 其基本思想是在抽取高频关键词对的基础上, 对词对所包含的主题内涵进行分析, 进一步将关键词集成为主题词, 利用主题词表达语义。主题图提供了一种良好的思想, 即在词的维度基础上加入主题的维度, 用主题来表达词的语义, 这也是未来利用共词网络进行科学知识结构分析的一个重要方向。

3 利用科学计量学实现科学知识管理的途径二: 科学知识演化分析

通过科学文献的数量增长和变化能够认识科学发展的过程, 尤其是一些数量上的突变更反映了科学发展过程中的重要事件。无论是从文献所构成的引文网络, 还是作者、机构、关键词等角度, 在科学知识结构分析基础上, 都可以进一步进行科学知识的演化分析^[6, 7]。科学作为一种系统, 其结构具有层次特征和动态特征, 科学研究发展的每一步都在不断继承和发展知识。科学知识的演化以时间为轴分为三个部分: 对过去发展变化的梳理, 对当前最新发展的监测, 对未来发展趋势的预测^[68]。

利用科学计量学进行科学知识演化的研究大致上有两种类型的研究, 一种是依据引文网络所进行的分析, 另一种则是将时间维度加到作者、词等其他特征项的网络分析中。引文关系所具有的时序特征在揭示科学知识的演化进程上具有十分突出的效果; 而后一种分析方法有时采取对整个时间段的作者、词进行网络分析, 然后进行分时间段分析, 有时则预先进行时间段的切分, 然后进一步分析每个时间段的内容, 进一步得到科学知识演化的整体结果。

3.1 基于引文网络的知识演化分析

引文关系具有明显的时序性, 通过对这种文献引文网络的时序关系进行分析, 可以展示科学发展的来龙去脉, 从而为厘清知识传承的过程和描述科学发展历史提供定量依据。早在 1960 年, 美国学者 Allen 就利用引文时间分布网络结构图给出了核酸染色方法在 20 世纪 40-60 年代的发展过程^[69]。科学引文索引的创始人 Garfield 从 1964 年开始相关的研究, 他在最初的研究中以著作《遗传密码》为基本出发点, 绘制了生物学家 Asimov 描述的关键事件图, 进一步从 Asimov 提及的描述事件的文献开始, 依据文献的参考关系编制了一个引文编年图, 通过引文编年图和节点事件的对比发现两者具有很高的一致性^[70]。Garfield 认为引文编年图对科学发展过程具有良好的揭示作用, 并在此基础上开发了 HistCite 软件^[71]。在 HistCite 软件中, 被引次数被用来衡量文献重要性的程度, 即在编年图中利用被引次数的大小标识文献节点的大小, 使引文编年图的可视化效果更为合理, 其可视化效果如图 3 所示。

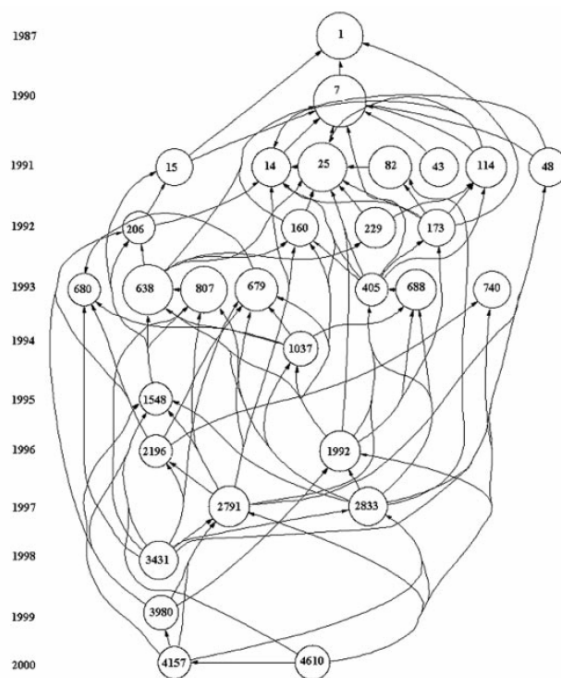


图 3 引文网络时序分析图

Fig. 3 The historiographs of citation network

然而，引文时序分析存在的局限也比较明显。在实际的分析过程中，衡量文献之间引用关系时往往采用二值方法，即“有-没有”引用，有则计 1，无则计 0。这种定义存在两个方面的问题：第一，对于文献之间存在引用关系，单纯的二分计数方法对强度的衡量存在局限性；对于某一篇文章而言，某些文献对于它是相对次要的，例如作为背景而被引用的文献，某些文献则是相对重要的，比如多次提及的文献，或者借鉴核心思想、方法迁移而引用的文献，事实上传统的引文分析是一种参考文献分析，而不是针对文献中引用本身的分析，引用行文本本身可能是多对多的关系，要解决这个问题需要对引文出现在施引文献的具体部分和引用的原因进行区分^[72]。第二，传统引用次数在作为衡量文献重要性时的另一个问题是无法区分施引文献的重要性，在实际的科学研究过程中，某一篇文章被不同的文献引用都会计一次，但被一篇重要性很低的文献和一篇重要性程度很高的文献引用对整个知识系统而言重要性是不一样的，想要区分不同文献的重要性，则又回到了是否用被引次数来衡量重要性的问题原点上，导致从局部角度基本无法解决这个问题。

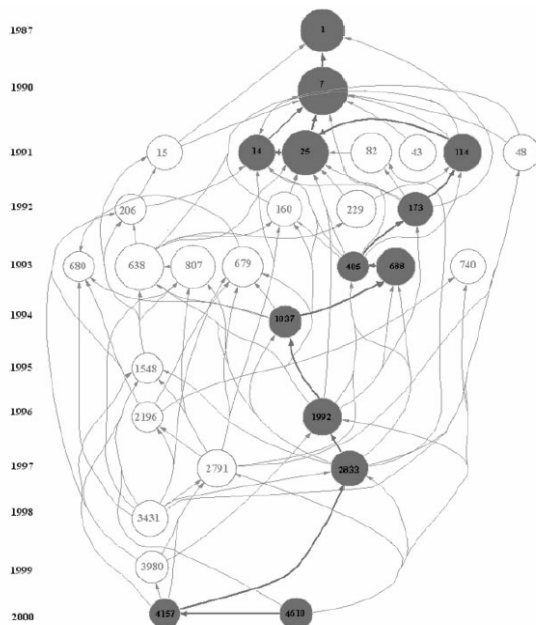


图 4 结合主路径的引文网络时序图

Fig.4 The main path combined historiographs of citation network

解决上述问题缺陷的一个重要途径是利用新的方法通过整体网络进行重要性评价,主路径分析法是其中的一个典型代表^[73]。主路径分析的基本思想是假设整个科学知识发展的过程存在路径,这个路径中的“驿站”即科学知识最初发展开始到当前发表的所有相关文献,其中的通道即按照时间顺序在这些“驿站”间通过引文关系所产生的文献链。从所有的原点出发,以引文为通道可以得到从知识的最初产生到发展现状的全部过程,在此基础上选择最重要的文献节点,连接而成的链将能够反映整个知识发展过程中最为重要的通道^[74]。结合主路径的引文时序分析能够有效地刻画出科学知识演化的主要过程,其可视化结果如图 4 所示。

由于引文网络具有复杂网络的一般特性,因此对引文网络的研究也可以分成宏观的统计特征揭示,中观的结构分析和围观的局部动机研究,但从科学知识的 management 角度而言,宏观的统计特征揭示所得到的结果更倾向于表现为一种现象,微观的局部动机研究则视野过于狭窄,不具备普遍的意义;对表现科学知识演化过程的网络进行结构分析和主路径抽取是一类典型的中观层次的研究,其结果能够有效对科学知识的演化过程进行有效的描述,明确在整个科学知识体系发展过程中哪些研究更加处于核心位置,进一步的,可以依据这些核心研究定位到相应的科学家,从而为科学知识的 management 和决策提供支持。

3.2 基于其他特征项网络的知识演化分析

由于不同类型的特征项网络分析能够揭示一定时间段内的科学知识结构,那么,通过对比不同时间段科学知识结构并揭示其中的变化过程,就可以实现科学知识的演化分析。Palla 等人基于 t 和 $t+1$ 两个时段社区的变化将演化归纳为六种模式:产生、消亡、分化、融合、扩散、收缩^[75]。科学知识的产生是一种 t 到 $t+1$ 时段从无到有的过程,消亡是 t 到 $t+1$ 时段从有到无的过程,分化是 t 到 $t+1$ 时段由一变多的过程,融合则是 t 到 $t+1$ 时段由多变一的过程,扩散 t 到 $t+1$ 时段规模变大的过程,收

缩则是 t 到 $t+1$ 时段规模变小的过程。这 6 种类型基本上涵盖了科学知识演化的全部过程。在进行两个不同时段比较式, 则可以从节点重合度、关系重合度、核心节点重合度等三个方面进行综合判断[76]。

基于分时段数据进行科学知识演化分析的典型案例是美国情报学家 White 和 McCain 的一项研究。White 和 McCain 利用作者共被引网络对 1972 年-1995 年期间国际图书情报领域科学知识演化的分析, 通过对比 1972-1979, 1980-1987, 1988-1995 年这三个时间段高被引学者所组成的共被引聚类发现, 从 1972 年到 1995 年的关于文献与科学交流的研究越来越多的被关注, 逐渐与情报检索研究趋于平衡状态[77]。图 5 是西班牙学者 Gonzalez-Alcaide 等人从作者网络角度进行黑热病领域科学知识演化研究的一个例子[78]。在这个图中以 10 年为界, 分别构建 1981-1990, 1991-2000, 2001-2010 这三个时间段的合作网路, 然后将之置于同一个时间轴上, 进一步分析每个时间段内最主要的作者合作网络关注的核心主题, 然后通过这种主题的变化来说明科学知识的演化过程。Gonzalez-Alcaide 等人研究基本上采用了与 White 和 McCain 相同的思路, 两者之间最大的区别在于 White 等人对于核心群体的区分依赖于聚类分析, 而 Gonzalez-Alcaide 等人则依赖于网络中社区发现的结果。

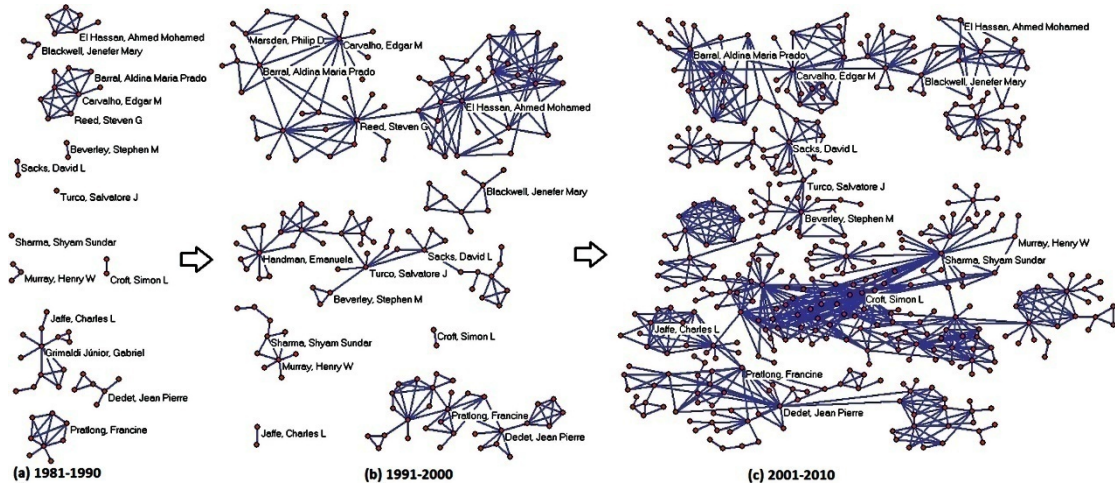


图 5 基于合作的黑热病研究领域的科学知识演化分析结果

Fig 5. Evolution of coauthorship networks for global research on leishmaniasis

依据作者等特征项知识关联网络分析的结果解读主要依赖于分析者对于特征项所代表内容的认知, 直接对反映内容的词进行分析来描述科学知识的演化具有一定的优势。词直观地反映了科学知识内容, 利用词进行的知识演化分析一般而言主要包括词频的时序分布和共词网络的时序分布两个方面。

研究人员通过关键词在不同时间段的出现频次能够监测科学研究主题随时间的变化, 这种变化利用可视化手段表现出来较为直观, 例如美国麻省大学 Swan 和 Jensen 开发的 TimeMines 系统[79], 该系统可以从数据样本中抽取有意义的关键词, 进一步通过 Timeline 视图去展示关键词的演化情况。随着技术的不断发展, 许多新的可视化效果逐渐出现, 例如 Havre 等学者利用 ThemeRiver 的方式占线关键词在时间维度上的演化, 不同颜色的河流代表不同意义的关键词, 图 6 展示了 ThemeRiver 的可视化效果[80]。

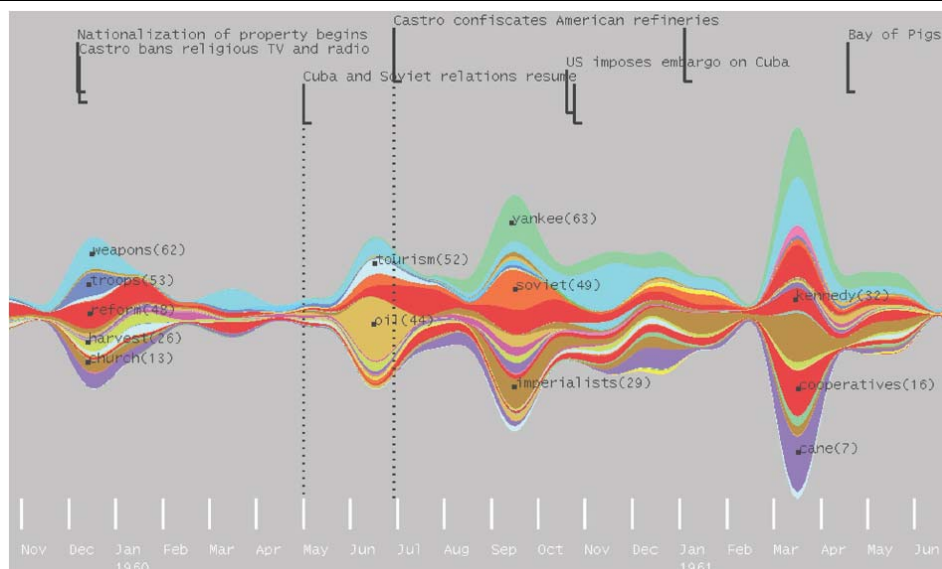


图 6 基于词频的科学知识演化 Themeriver 图

Fig.6 Themeriver graph on evolution of science based on words' frequency

基于共词网络的演化分析往往和主题演化结合在一起。共词能够在一定层面上解决歧义问题，因此利用共词网络来分析主题语义可以取得较好的效果，进一步得到的科学知识演化分析结果也更为准确。西班牙学者 Cobo 等将基于共词网络的演化分析步骤概括为数据抽取、网络表示、时序数据生成、社区发现、主题标识、演化分析和可视化结果等多个内容^[81]。

从分析步骤上来看，利用共词网络进行的科学知识演化分析中最核心的内容包括社区发现、主题标识和社团演化的可视化。对于社区发现而言，目前的研究中主要与社会网络和复杂网络分析方法相结合，例如 Newman&Girvan 算法^[82]，Blondel 等人提出的算法^[83]，Van Eck 的 VOS 算法^[84]等。共词网络主题的标识是社区发现后需要开展的一项重要工作，由于共词网络中的节点是关键词，因此确定社区所反映主题的过程就是定位和标识社区中核心节点的过程，利用少数的核心节点所组成的集合能够代表社区所对应的科学知识主题。但是，目前大量的指标在衡量节点重要性时精确度不高，针对这种缺陷，Guimera 等学者提出利用 Z-Value 来衡量节点之间的紧密性，在局部层面衡量节点的重要性^[85]。

在演化结果的可视化方面，典型的方法包括聚类链^[86]、冲积图^[87]、Timeline 和 Timezone 等^[88]。图 7 是 2001-2007 年分三个时间段的信息检索领域知识演化二分图可视化结果^[89]，从中可以很容易地发现每个时期不同的研究内容和热点，并且明确地刻画了科学知识的演化过程。

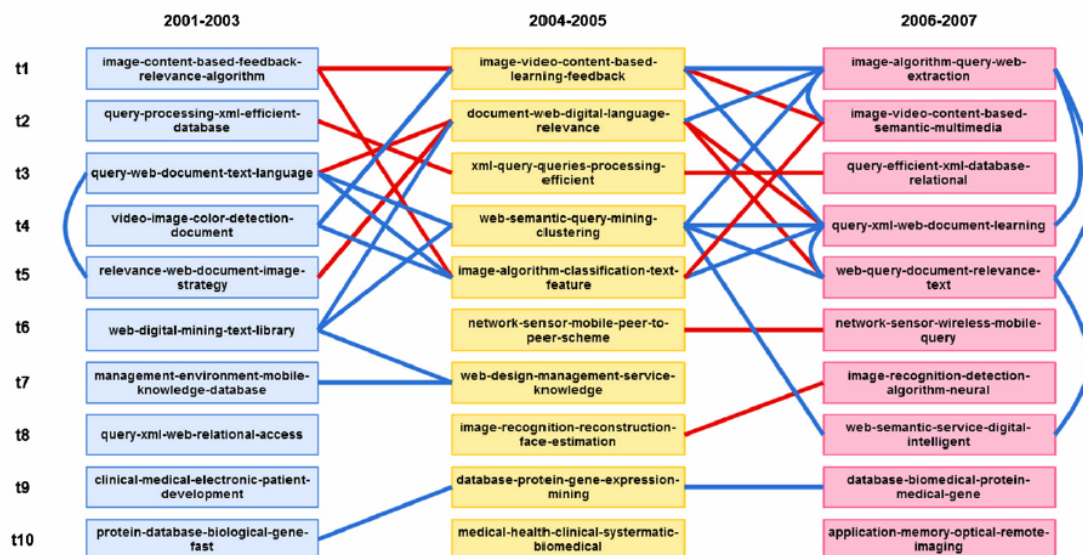


图 7 2001-2007 年信息检索领域的知识演化

Fig. 7 The evolution of knowledge about information retrieval during 2001 to 2007

4 总结

科学计量学自诞生以来至今已经历了近半个世纪的发展,随着科学研究系统发展到高度分化和综合的阶段,对科学知识的管理逐渐成为科研管理人员和决策机构关注的重点。作为一种针对科学事业的定量分析方法,科学计量学在科学知识管理和科研事业管理中的所起作用也越来越重要。针对科学计量学的定义众说纷纭的现状,本文认为科学计量学是一门以描述科学发展过程,揭示科学发展内机理,预测科学发展趋势,以科学管理工作提供支持依据为目的,通过定量分析方法研究科学事业及其主体和客体的应用性学科;论述了应用于科学知识管理的科学计量方法和工具的新发展。科学计量学的应用性特征在科学知识的管理过程中表现尤其突出。

利用科学计量学进行科学知识的管理主要包括两个部分,从静态的角度对科学知识的结构研究和从动态的角度对科学知识的演化研究。结构分析能够揭示科学知识的学科、地理分布,学科内部的主题、学派结构,跨学科交叉、融合情况等内容;对于研究人员来说,这些分析结果对科学知识的管理具有重要意义;对于科研政策制定者和管理者来说,这种布局结果则是他们进行决策的重要参考依据。科学知识的演化分析在以往的研究中被广泛用于描述科学知识的发展史,这种发展过程的刻画能够实现更具普遍意义的科学知识发展规律的归纳和总结,依据这些结论则能够实现对未来发展趋势的分析,有效支持科研决策。对于科学知识结构的分析,本研究主要选择了从作者和词两类网络的角度进行了论述,其原因是从上述两个角度的研究基本上涵盖了可能会用到的方法。虽然利用引文或者其衍生的共被引、耦合等方法也能够实现结构分析,但引文的本质是代表了科学知识的扩散和传承,因此本文将引文网络作为演化分析的重要途径进行了单独的论述;而当其他特征项网络结构分析结果加入时间维度时,同样可以实现演化分析的目的,因此演化分析部分也主要分成两个部分。

目前的研究中,越来越多的新指标和新方法被逐渐应用于科学计量分析从而实现科学知识的管

理,随之而来也产生了许多需要注意的内容,例如,大规模数据环境下,大多数研究样本都会存在覆盖范围和数据质量的问题,数据清洗工作应该引起足够的重视;许多新的测度方法和指标的应用缺乏机理性的解释,例如社会网络分析中的中心性目前已经被广泛应用于计量分析中,但这类指标大多产生于人际网络分析的研究中,科学知识网络所具有的特点能否保证这些指标的适用性需要给出更加全面的机理性解释。随着研究的不断深入,这些问题将会逐步得到解答,最终促进科学计量学在科学知识管理中的有效应用。

参考文献

- [1] Lazer D., Pentland A. S., Adamic L., et al. Life in the network: the coming age of computational social science[J]. *Science*, 2009, 323(5915): 721-723.
- [2] Hey T. The Fourth Paradigm-Data-Intensive Scientific Discovery[M]//*E-Science and Information Management*. Springer Berlin Heidelberg, 2012: 1.
- [3] 宋兆杰, 王续琨. 纳里莫夫: 苏联科学计量学的创始人-纪念纳里莫夫诞辰 100 周年[J]. *科学学研究*, 2010(3): 333-338.
Song, Z. J., Wang, X. K. Nalimov: the founder of Soviet Union Scientometrics[J]. *Studies in Science of Science*, 2010(3): 333-338. (in Chinese)
- [4] 邱均平, 文庭孝, 宋艳辉等. 知识计量学[M]. 北京: 科学出版社, 2014: 1-7.
Qiu J. P., Wen T. X. Song Y. H. Knowledgometrics[M]. *Science Press*, Beijing, 2014: 1-7. (in Chinese)
- [5] ISSI. Mission[EB/OL]. [2014-6-10]. <http://www.issi-society.org/mission.html>.
- [6] 侯海燕, 刘则渊, 栾春娟. 基于知识图谱的国际科学计量学研究前沿计量分析[J]. *科研管理*, 2009, 01: 164-170.
Hou, H. Y., Liu, Z. Y., Luan C. J. Quantitative analysis on the research front of international scientometrics based on mapping of knowledge[J]. *Science Research Management*, 2009, 01: 164-170. (in Chinese)
- [7] 邱均平. 信息计量学[M]. 武汉: 武汉大学出版社, 2007: 16.
Qiu, J. P. Informetrics[M]. *Wuhan University Press*, Wuhan, 2007: 16. (in Chinese)
- [8] Narin, F. Patent bibliometrics[J]. *Scientometrics*, 1994, 30(1): 147-155.
- [9] Callon M., Courtial J P., Turner W A, et al. From translations to problematic networks: An introduction to co-word analysis[J]. *Social science information*, 1983, 22(2): 191-235.
- [10] 冯璐, 冷伏海. 共词分析方法理论进展[J]. *中国图书馆学报*, 2006, 32(2): 88-92.
Feng, L., Leng, F. H. Development of Theoretical Studies of Co-Word Analysis[J]. *Journal of Library Science in China*, 2006, 32(2): 88-92.
- [11] Swanson D. R. Fish oil, raynaud's syndrome, and undiscovered public knowledge[J]. *Perspectives in biology and medicine*, 1986, 30(1): 7-18.
- [12] 董克. 数字文献资源多元深度聚合研究[D]. 武汉: 武汉大学博士学位论文, 2014: 49.
Dong K. Research on Diversified and Deep Aggregation of Digital Literature Resources[D]. Wuhan: Wuhan University, 2014: 49. (in Chinese)
- [13] Leydesdorff L., Welbers K. The semantic mapping of words and co-words in contexts[J]. *Journal of Informetrics*, 2011, 5(3): 469-475.
- [14] Small H. Co-citation in the 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.
- [15] White H.D., Griffith B.C. Author cocitation: A literature measure of intellectual structure[J]. *Journal of the American Society for Information Science*, 1980(May): 163-171.

- [16] 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): 290-296.
- [17] 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.
- [18] Boyack, K.W., Small, H., Klavans, R. Improving the accuracy of co-citation clustering using full text[J]. *Journal of American Society for Information Science and Technology*, 2013, 64(9): 1759-1767.
- [19] 袁军鹏. 科学计量学高级教程[M]. 北京: 科学技术文献出版社, 2010: 9.
Yuan J. P. Advanced Course in Scientometrics[M]. *Scientific and Technical Documentation Press*, Beijing, 2010: 9. (in Chinese)
- [20] 朱星宇, 陈勇强. SPSS 多元统计分析及方法应用[M]. 北京: 清华大学出版社, 2011: 241.
Zhu X. Y., Chen, Y. Q. Methods and Application of Multivariate Statistical Analysis in SPSS[M]. *Tsinghua University Press*, Beijing, 2011: 241. (in Chinese)
- [21] Tijssen I.R.J.W, Van Raan, A.F.J. Mapping Co-word Structures: A Comparison of Multidimensional- scaling And Leximappe[J]. *Scientometrics*, 1989, 15(3-4): 283-295.
- [22] Small H, Sweeney E. Clustering the Science Citation Index Using Co-citations: A Comparison Of Methods[J]. *Scientometrics*, 1985, 7(3-6): 391-409.
- [23] Wasserman S., Faust K. Social Network Analysis: Methods and Application[M]. *Cambridge University Press*, New York, 1994: 59.
- [24] 董克, 刘德洪, 江洪. 基于三方关系组的引用网络结构分析[J]. 情报理论与实践, 2010, 33(11): 50-53.
Dong K. Liu, D. H., Jiang H. Structural Analysis of Citation Network Based on Triad Census[J]. *Information Studies: Theory & Application*, 2010, 33(11): 50-53. (in Chinese)
- [25] Klavans R., Boyack K. W. Identifying a better measure of relatedness for mapping science[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(2): 251-263.
- [26] Hirsch J. E. An index to quantify an individual's scientific research output[J]. *Proceedings of the National academy of Sciences of the United States of America*, 2005, 102(46): 16569-16572.
- [27] Alonso S., Cabrerizo F. J., Herrera-Viedma E., et al. h-Index: A review focused in its variants, computation and standardization for different scientific fields[J]. *Journal of Informetrics*, 2009, 3(4): 273-289.
- [28] Priem, J., Taraborelli, D., Groth, P., Neylon, C. altmetrics: a manifesto[EB/OL]. [2014-5-15]. [http: //altmetrics.org/manifesto/](http://altmetrics.org/manifesto/).
- [29] Piwowar H. Altmetrics: Value all research products[J]. *Nature*, 2013, 493(7431): 159-159.
- [30] Web of Knowledge[EB/OL]. [2014-6-25]. www.webofknowledge.com.
- [31] About Engineering Index[EB/OL]. [2014-6-25]. [http: //www.elsevier.com/online-tools/engineering-village](http://www.elsevier.com/online-tools/engineering-village).
- [32] Scopus[EB/OL]. [2014-6-29]. [http: //www.elsevier.com/online-tools/scopus](http://www.elsevier.com/online-tools/scopus).
- [33] 中文社会科学引文索引简介[EB/OL]. [2014-6-29]. [http: //www.lib.whu.edu.cn/web/dzzy/detail.asp?q=IDN=WHU00887&s=detail&full=Y](http://www.lib.whu.edu.cn/web/dzzy/detail.asp?q=IDN=WHU00887&s=detail&full=Y).
Introduction to Chinese Social Science Citation Index[EB/OL]. [2014-6-29]. [http: //www.lib.whu.edu.cn/web/dzzy /detail.asp?q=IDN=WHU00887&s=detail&full=Y](http://www.lib.whu.edu.cn/web/dzzy/detail.asp?q=IDN=WHU00887&s=detail&full=Y). (in Chinese).
- [34] 储节旺, 郭春侠. 共词分析法的基本原理及 EXCEL 实现[J]. 情报科学, 2011, 06: 931-934.
Chu, J. W. Guo, C. X. General Principles of Co-Words analysis Method and Achieved by EXCEL[J]. *Information Science* 2011, 6: 931-934. (in Chinese)
- [35] Borgatti, S. P., Everett M. G., Freeman L. C. About UCINET[EB/OL].[2014-6-11]. [http: //www.analytictech.c om/ucinet/description.htm](http://www.analytictech.com/ucinet/description.htm)
- [36] Batagelj V, Mrvar A. Pajek-program for large network analysis[J]. *Connections*, 1998, 21(2): 47-57.
- [37] Bibexcel[EB/OL]. [2014-7-2]. [http: //www8.umu.se/inforsk/Bibexcel/](http://www8.umu.se/inforsk/Bibexcel/).
- [38] Thomson Reuters. Thomson Data Analyzer[EB/OL]. [2014-7-2]. [http: //thomsonreuters.com/thomson-data-analyz](http://thomsonreuters.com/thomson-data-analyz)

er/.

- [39] Leydesdorff's Software[EB/OL]. [2014-7-2]. <http://www.leydesdorff.net/software.htm>.
- [40] Börner K. Science of Science Studies: Sci2 Tool[J]. *Communications of the ACM*, 2011, 54(3): 60-69.
- [41] 邱均平. 信息计量学[M]. 武汉: 武汉大学出版社, 2007: 502.
- Qiu, J. P. Informetrics[M]. *Wuhan University Press*, Wuhan, 2007: 502. (in Chinese)
- [42] Cronin B. Invisible colleges and information transfer A review and commentary with particular reference to the social sciences[J]. *Journal of Documentation*, 1982, 38(3): 212-236.
- [43] Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology*, 60(10), 2107-2118.
- [44] 董克. 数字文献资源多元深度聚合研究[D]. 武汉: 武汉大学博士学位论文, 2014: 42-43.
- Dong K. Research on Diversified and Deep Aggregation of Digital Literature Resources[D]. Wuhan: Wuhan University, 2014: 42-43. (in Chinese)
- [45] White H. D., Wellman B., Nazer N. Does citation reflect social structure?: Longitudinal evidence from the "Globenet" interdisciplinary research group[J]. *Journal of the American Society for information Science and Technology*, 2004, 55(2): 111-126.
- [46] Nazer, N. Operating virtually within a hierarchical framework: How a virtual organization really works[D]. Toronto: University of Toronto, 2001.
- [47] Ding Y. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks[J]. *Journal of informetrics*, 2011, 5(1): 187-203.
- [48] White H.D., Griffith B.C. Author cocitation: A literature measure of intellectual structure[J]. *Journal of the American Society for Information Science*, 1980(May): 163-171.
- [49] 邱均平, 王菲菲. 基于作者互引分析的科学结构研究探析—以科学计量学为例[J]. *科学学研究*, 2012, 06: 829-840.
- Qiu J. P., Wang F. F. An exploration of scientific structure research based on author cross-citation analysis-taking scientometrics as an example[J]. *Studies in Science of Science*, 2012, 06: 829-840. (in Chinese)
- [50] 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.
- [51] Chen L. C., Lien Y. H. Using author co-citation analysis to examine the intellectual structure of e-learning: A MIS perspective[J]. *Scientometrics*, 2011, 89(3): 867-886.
- [52] Egghe L., Leydesdorff L. The relation between Pearson's correlation coefficient r and Salton's cosine measure[J]. *Journal of the American Society for Information Science and Technology*, 2009, 60(5): 1027-1036.
- [53] Chen C, Paul R. J, O'Keefe B. Fitting the jigsaw of citation: Information visualization in domain analysis[J]. *Journal of the American Society for Information Science and Technology*, 2001, 52(4): 315-330.
- [54] Zhao D., Strotmann A. Information science during the first decade of the web: An enriched author cocitation analysis[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59(6): 916-937.
- [55] 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.
- [56] Zhao D., Strotmann A. Evolution of research activities and intellectual influences in information science 1996 – 2005: Introducing author bibliographic - coupling analysis[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59(13): 2070-2086.
- [57] 马瑞敏, 倪超群. 作者耦合分析: 一种新学科知识结构发现方法的探索性研究[J]. *中国图书馆学报*, 2012, 02: 4-11.
- Ma R. M., Ni C. Q. Author Coupling Analysis: An Exploratory Study on a New Approach to Discover Intellectual Structure of a Discipline[J]. *Journal of Library Science in China*, 2012, 02: 4-11. (in Chinese)

- [58] 董克. 数字文献资源多元深度聚合研究[D]. 武汉: 武汉大学博士学位论文, 2014: 140-141.
Dong K. Research on Diversified and Deep Aggregation of Digital Literature Resources[D]. Wuhan: Wuhan University, 2014: 140-141. (in Chinese)
- [59] Leydesdorff L., Persson O. Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes[J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(8): 1622-1634.
- [60] Nooy, W., Mrvar, A., Batagelj, V. Exploratory social network analysis with Pajek[M]. *Cambridge University Press*, New York, 2005.
- [61] 邱均平, 董克. 作者共现网络的科学研究结构揭示能力比较研究[J]. 中国图书馆学报, 2014, 01: 15-24.
Qiu J. P., Dong K. A Comparative Study on the Ability of Author Co-occurrence Network in Revealing Scientific Structure[J]. *Journal of Library Science in China*, 2014, 01: 15-24. (in Chinese)
- [62] 袁军鹏. 科学计量学高级教程[M]. 北京: 科学技术文献出版社, 2010: 97-106.
Yuan J., P. Advanced Course in Scientometrics[M]. *Scientific and Technical Documentation Press*, Beijing, 2010: 97-106. (in Chinese)
- [63] Newman M. E. J. Power laws, Pareto distributions and Zipf's law[J]. *Contemporary physics*, 2005, 46(5): 323-351.
- [64] Hu C. P., Hu J. M., Deng S. L., et al. A co-word analysis of library and information science in China[J]. *Scientometrics*, 2013, 97(2): 369-382.
- [65] Zong Q. J., Shen H. Z., Yuan Q. J., et al. Doctoral dissertations of Library and Information Science in China: A co-word analysis[J]. *Scientometrics*, 2013, 94(2): 781-799.
- [66] 李纲, 王忠义. 基于语义的共词分析方法研究[J]. 情报杂志, 2011, 30(12): 145-149.
Li G. Wang Z. Y. Research on Semantic Co-words Analysis Mehtod[J]. *Journal of Intelligence*, 2011, 30(12): 145-149. (in Chinese)
- [67] 陈必坤. 学科知识可视化分析研究[M]. 武汉: 武汉大学博士学位论文, 2014: 125.
Chen B. K. Research on Visualization Analysis of Disciplinary Knowledge[D]. Wuhan: Wuhan University, 2014: 125. (in Chinese)
- [68] 化柏林, 武夷山. 动态环境需要演化分析[J]. 情报学报, 2013, 32(9): 1.
Hua B. L., Wu Y. S. Dynamic environment needs evolution analysis[J]. *Journal of the China Society for Scientific and Technical Information*, 2013, 32(9): 1.
- [69] 庞景安. 科学计量研究方法论[M]. 北京: 科学技术文献出版社, 2002: 264.
Pang J. A. Methodology of Scientometric Research[M]. *Scientific and Technical Documentation Press*, Beijing, 2002: 264. (in Chinese)
- [70] 李运景, 侯汉清, 裴新涌. 引文编年可视化软件 HistCite 介绍与评价[J]. 图书情报工作, 2006, 50(12): 135.
Li Y. J., Hou H. Q., Pei X. Y. Introduction and Evaluation to HistCite – a Citation Historiography Visualization Software[J]. *Library and Information Service*, 2006, 50(12): 135. (in Chinese)
- [71] Garfield E. Historiographic mapping of knowledge domains literature[J]. *Journal of Information Science*, 2004, 30(2): 119-145.
- [72] 胡志刚, 陈超美, 刘则渊等. 从基于引文到基于引用——一种统计引文总被引次数的新方法[J]. 图书情报工作, 2013, 57(21): 5-10.
Hu Z. G., Chen, C. M., Liu Z. Y., et al. From Counting References to Counting Citations: A New Way to Calculate the Total Cited Times of References[J]. *Library and Information Service*, 2013, 57(21): 5-10. (in Chinese)
- [73] Hummon N. P., Dereian P. Connectivity in a citation network: The development of DNA theory[J]. *Social Networks*, 1989, 11(1): 39-63.
- [74] Hummon N. P., Carley K. Social networks as normal science[J]. *Social Networks*, 1993, 15(1): 71-106.
- [75] Palla G., Barabási A. L., Vicsek T. Quantifying social group evolution[J]. *Nature*, 2007, 446(7136): 664-667.
- [76] 王晓光, 程齐凯. 基于 NEViewer 的学科主题演化可视化分析[J]. 情报学报, 2013, 32(9): 900-911.

- Wang X. G., Chen Q. K. Analysis on Evolution of Research Topics in a Discipline Based on NEViewer[J]. *Journal of the China Society for Scientific and Technical Information*, 2013, 32(9): 900-911. (in Chinese)
- [77] 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.
- [78] Gonzalez-Alcaide G., Huamani C., Park J., et al. Evolution of coauthorship networks: worldwide scientific production on leishmaniasis[J]. *Revista da Sociedade Brasileira de Medicina Tropical*, 2013, 46(6): 719-727.
- [79] Swan R., Jensen D. Timemines: Constructing timelines with statistical models of word usage[C]//*KDD-2000 Workshop on Text Mining*. 2000: 73-80.
- [80] Havre S., Hetzler E., Whitney P., et al. Themeriver: Visualizing thematic changes in large document collections[J]. *Visualization and Computer Graphics, IEEE Transactions*, 2002, 8(1): 9-20.
- [81] Cobo M. J., López-Herrera A. G., Herrera-Viedma E., et al. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field[J]. *Journal of Informetrics*, 2011, 5(1): 146-166.
- [82] Newman M. E. J., Girvan M. Finding and evaluating community structure in networks[J]. *Physical Review E*, 2004, 69(2): 026113.
- [83] Blondel V. D., ajardo A., Heymans M., et al. A measure of similarity between graph vertices: Applications to synonym extraction and web searching[J]. *SIAM Review*, 2004, 46(4): 647-666.
- [84] Van Eck N. J., Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping[J]. *Scientometrics*, 2010, 84(2): 523-538.
- [85] Milo R., Shen-Orr S., Itzkovitz S., et al. Network motifs: simple building blocks of complex networks[J]. *Science*, 2002, 298(5594): 824-827.
- [86] Small H. Tracking and predicting growth areas in science[J]. *Scientometrics*, 2006, 68(3): 595-610.
- [87] Rosvall M., Bergstrom C. T. Mapping change in large networks[J]. *PloS One*, 2010, 5(1): e8694.
- [88] Chen C. 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.
- [89] Yan E., Ding Y., Milojević S., et al. Topics in dynamic research communities: An exploratory study for the field of information retrieval[J]. *Journal of Informetrics*, 2012, 6(1): 140-153.

Research on the Application of Scientometrics to

Scientific Knowledge Management

Qiu Junping¹, Dong Ke¹, Ma Ruimin²

1. School of Information Management, Wuhan University, Wuhan 430072, China;
2. Institute of Management and Decision, Shanxi University, Taiyuan 030006, China

Abstract: This paper defines the term of scientometrics, illustrates the research scope of scientometrics, and induces new developing trend of scientometric methods and tools adopted for the scientific knowledge management. The practical application of scientometrics to the scientific knowledge management includes two aspects: knowledge structure analysis, and knowledge

evolution analysis. The first aspect is illustrated in detail by author network together with co-words network analysis; the second aspect is discussed from citation network, author network and words network analysis. This paper proves that scientometrics can give efficient support to the scientific knowledge management.

Key words: Scientometrics; Scientific Knowledge Management; Knowledge Structure; Knowledge Evolution