

● 李长玲, 刘非凡, 魏绪秋 (山东理工大学 科技信息研究所, 山东 淄博 255049)

基于3-mode 网络的领域主题演化规律分析^{*}

——以知识网络研究领域为例

摘 要: 对某学科、某领域的主题演化轨迹进行分析, 反映该学科、领域的发展历史、现状与未来, 是目前情报分析的重要内容。文章以知识网络为例, 检索了该领域的期刊文献, 提取年份、作者和关键词题录数据, 构建了 3-mode 共现关系词组。运用社会网络分析方法从合著网络、共词网络以及时间—作者—关键词 3-mode 共现网络 3 个角度对知识网络领域的演化特征进行了分析。研究发现, 社会网络分析方法和 3-mode 共现分析方法能够应用于领域主题演化分析中, 揭示网络演化特征, 发现不同年份具有代表性的作者及合著团队、经典和热点研究主题, 综合呈现该领域内年份—作者—主题之间的共现关系。

关键词: 知识网络; 学科领域; 主题演化分析; 可视化

Abstract: Analysis of the theme evolution of a certain discipline and domain to reflect its development history, current status and future is the important content of the current intelligence analysis. Taking knowledge network as an example, this paper retrieves the periodical literature in this domain, extracts the bibliographic data of year, author and keyword, and constructs co-occurrence relations phrase of 3-mode. By using the social network analysis method, the paper analyzes the evolution characteristics of knowledge network area from 3 perspectives of collaboration networks, co-word network and time-author-keyword 3-mode co-occurrence network. The research finds that social network analysis method and 3-mode co-occurrence analysis method can be used on the analysis of domain theme evolution, in order to reveal the characteristics of network evolution, discover typical author and co-author team as well as classical and hot research themes for different years, and present the co-occurrence relations of year-author-theme in this area comprehensively.

Keywords: knowledge network; discipline area; theme evolution analysis; visualization

主题演化是指某学科或研究领域在时间维度上的发展变化过程。与空间变化相比, 时间演化体现的是主题的新陈代谢, 展示了某一学科领域的发展态势和未来走势, 是研究学科发展规律的重要手段。学科主题演化通过比较不同时间段的映射网络, 获取对学科主题产生、消亡、分化、融合等方面的趋势性认识^[1]。

Barabási 等对科学家合著网络的无标度特征进行分析, 通过一个简单算法即时提取演化网络结构图, 发现网络的内部联系在决定标度和网络拓扑^[2]。Csárdi 提出了研究科学引文和合作网络的一种新方法, 这种方法依赖于 KERNEL 和节点属性的标量功能, 可随机增减顶点和边,

可用于理解演化网络的动态性^[3]。L. Leydesdorff 以 M. Callon 教授发表的论文为实例, 研究了增加网络中的异质属性的必要性, 并绘制了 3-mode 可视化网络图。他认为科学计量学不仅仅要关注网络中的行动者, 还应该通过研究网络中多属性之间的联系, 从语义层面上理解科学知识图谱的内涵^[4]。李金华等将复杂网络理论引入到创新网络领域, 定量描绘了创新网络的拓扑结构, 以小世界性程度为指标比较分析了不同的演化模型^[5]。万江平分析了 1991—2005 年间我国软件质量研究主题的演化过程, 讨论了未来的研究趋势: 软件控制、软件开发以及软件测试^[6]。王福生等建立一个基于科研论文作者合作方式的网络模型, 并对 2001—2006 年发表在《图书情报工作》期刊上的科研论文进行统计, 描述了作者合作网络的演化过程^[7]。韩真介绍了学科主题类型划分原理和方法, 认为只有关联强度算法能很好地体现主题网络中核心主题和次要主题的本质特征^[1]。刘向从拓扑结构和演化机制两个方面

^{*} 本文为国家社会科学基金项目“基于社会网络分析的学科评价研究”(项目编号: 11BTQ020) 和山东理工大学人文社会科学发展基金项目“Web 信息挖掘与智能检索”(项目编号: 2010GGTD05) 的研究成果之一。

综述了知识网络的理论研究进展^[8]。沈思等分析了文献在不同研究阶段用词特征的基础上,提出了一种基于主题模型的研究发展阶段识别方法^[9]。

因此,当前学者主要通过构建网络演化模型,并运用相关数据,以统计和仿真方法来对所提出的模型进行验证。也有学者对学科主题演化理论和方法进行综述。但是,鲜有学者应用 3-mode 共现分析方法对某一研究主题的演化规律进行研究。本文结合社会网络分析方法,以知识网络领域的研究文献为例,对该领域的动态演化规律进行可视化分析,以找出该研究领域的发展过程,揭示其演化规律。

1 共现分析方法及社群演化分析法

共现分析方法利用文献集中词组共同出现的情况,确定该文献集中各词组之间的关系。一般认为词对在文献中出现的次数越多,这两个词之间的关系越紧密。静态的共现网络可以反映学科的研究现状,而动态的共现网络可反映学科的发展过程。随着新发表文献的不断出现,新的知识点与研究主题也在不断变化,词汇及其共现关系的变化导致共现网络处于不断的演化过程中^[10]。本研究定义 3-mode 共现分析方法为由文献的任意 3 个字段共现的分析方法,以文章发表年份、作者、关键词字段及其共现关系形成的网络为例,分析某研究领域的现状与变化趋势。与合著网络和共词网络相比,3-mode 共现分析方法不仅能够从单一层面揭示科研团队或领域知识的结构关系,还能从年份—作者—关键词共现关系的角度综合分析领域主题的时序演化特征。将领域主题中的年份、作者和关键词及其之间的共现关系用节点和连线来表示,就可以应用社会网络分析软件工具绘制作者合著、相关主题和作者—主题关系等演化网络图,进而对演化规律做进一步探讨和分析。

G. Palla 根据在社群生命周期中可能出现的事件将网络的演化过程分为 6 种:产生、扩张、合并、分裂、收缩和消亡^[11],见图 1。

产生:指 t 时间段不存在的社区,在 $t+1$ 时间段产生;扩张:前 t 时间段存在的社区,在 $t+1$ 时间段继续存在,但规模扩大;合并:前 t 时间段的两个或者多个社区,在 $t+1$ 时间段合成一个新的社区;分裂:前 t 时间段的社区,在 $t+1$ 时间段分化成为两个或多个新的社区;收缩:前 t 时间段存在的社区,在 $t+1$ 时间段继续存在,但规模缩小;消亡:前 t 时间段存在的社区,在 $t+1$ 时间段没有存在^[12]。

某学科或研究领域的合著网络和关键词网络的演化过程能够反映受到社会环境的变化、科研人员的替代、合著

团队的发展、研究重点的深入、研究热点的转移等因素的影响。因此,通过对这些过程的可视化分析,判断网络演化过程中社群不同的事件就能够了解该领域研究的演化规律。

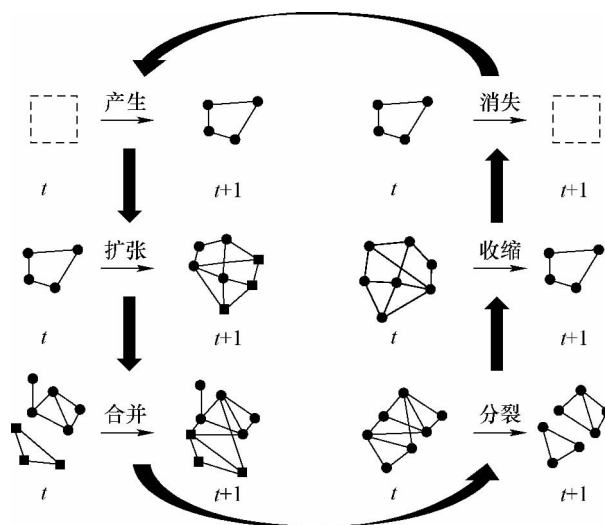


图1 社群网络生命周期演化过程

按照复杂网络理论,网络的结构特征(静态)决定了网络的演化特征(动态),网络的演化结果在一定程度上反作用于网络的结构,网络的结构和演化从某种意义上说是关联在一起的,结构分析和演化分析是两种互补的策略^[13]。

本文检索了国内知识网络领域的相关题录数据,构建 3-mode 共现关系词组,运用社会网络分析软件工具对该领域演化规律进行研究,探索为学科或领域演化分析的方法研究提供新的思路和途径。

2 运用 3-mode 数据分析知识网络领域演化规律

2.1 数据采集和预处理

在《中国学术期刊全文数据库》中,时间跨度选择 2003—2012 年,检索以“知识网络”为主题的核心期刊文献的题录数据,检索时间为 2013 年 10 月 6 日。数据清洗过程中发现,教育学领域同样有关于知识网络的概念,但这一概念大多数围绕教学当中知识点和教学手段等方面,与本文所研究情报学等相关学科领域的知识网络概念差别很大,因此去除了与教育学相关的题录数据,共剩余有效文献数量总计 477 篇。确定文献范围后,用 Bibexcel^[14]进行关键词处理与统计分析,共得到研究人员 697 位,文献关键词 1111 个。

首先构建年份、作者和关键词之间的共现关系词组,为进一步可视化分析做准备。具体步骤如下:

1) 构建年份—作者和年份—关键词 2-mode 关系数据。用 Bibexcel 对搜集到的文献题录信息中年份和作者、年份和关键词分别进行共现分析,得到年份—作者和年份—关键词共现关系词组,部分数据见表 1。

表 1 知识网络领域年份—作者、年份—关键词
共现关系词频表 (部分)

年份	作者	共现频次	年份	关键词	共现频次
2009	顾新	12	2009	知识管理	13
2009	肖冬平	10	2010	知识管理	13
2011	顾新	8	2008	知识管理	10
2007	赵蓉英	5	2005	知识管理	9
2011	汪涛	4	2007	知识管理	9
2011	刘向	4	2004	知识管理	8
2010	周晓英	4	2006	知识管理	8
2007	邱均平	4	2011	知识管理	7
2011	马费成	4	2012	知识管理	7
2011	吴绍波	4	2011	模型	6
2011	姜永常	3	2012	社会网络	6
2008	马德辉	3	2011	社会网络分析	6
...
2010	彭宁波	1	2009	作者共引图谱	1

2) 构建 3-mode 共现词组。将文献题录信息按照年份分为 10 个文件,再分别用 Bibexcel 统计文献题录信息中作者和关键词共现频次,得到作者—关键词共现关系词组。为共现词组添加上年份项并按照共现频次降序排列,即得到年份—作者—关键词 3-mode 共现词组,部分数据见表 2。

表 2 知识网络领域年份—作者—关键词
3-mode 共现关系词组 (部分)

年份	作者	主题	共现频次	年份	作者	主题	共现频次
2009	顾新	知识网络	12	2011	马费成	增长	3
2009	肖冬平	知识网络	10	2011	刘向	增长	3
2011	顾新	知识网络	8	2011	王涛	知识网络	3
2007	赵蓉英	知识网络	5	2011	陈健	知识网络	3
2011	吴绍波	知识网络	4	2010	顾新	知识网络	3
2010	周晓英	知识链接	4	2008	马德辉	知识管理	3
2007	邱均平	知识网络	4	2008	包昌火	知识管理	3
2011	顾新	公平感	3	2008	马德辉	知识网络	3
2011	吴绍波	公平感	3	2008	包昌火	知识网络	3
2011	陈健	公平感	3	2007	赵蓉英	知识节点	3
2011	马费成	老化	3	2012	刘向	知识网络	2
2011	刘向	老化	3	2012	王斌	知识网络	2
2011	马费成	模型	3
2011	刘向	模型	3	2003	邱均平	组织绩效	1

以此为研究对象,运用社会网络分析软件工具,对检索结果进行处理、分析与可视化,更形象、直观地研究知识网络领域的演化规律。

2.2 数据结果分析

首先分析知识网络领域的文献信息特征,再分别从时间序列的角度对知识网络的合著、共词和作者—主题演化网络进行可视化分析,以全面揭示出该主题的动态演化特征。

2.2.1 知识网络领域研究文献信息特征 通过统计历年知识网络的文献数量,计算发文累积量对数分布,结果见图 2。从图中可以看出“知识网络”研究文献增长呈指数增长规律,累积量—年序号对数分布近似一条直线,拟合高,达到 0.9968,从另一个角度再次说明该领域的文献增长符合普赖斯提出的科技文献指数增长规律^[13],同时也表明“知识网络”领域在这 10 年间正处于快速良好的发展阶段。根据图 2 中“发文量累计对数分布”模型公式: $y = 1.7485x + 2.135$,可以预测,2013 年“知识网络”累积量为 560 篇左右,即这一领域在 2013 年发文量可能为 80 篇左右。

知识网络领域文献的增长可能是由于以下多方面的原因共同造成的,如科研经费和科技人员数量的增长(2003 年该领域研究人员共 13 人,2010 年增加到 98 人),研究范围的扩大和细分(2003 年涉及该领域关键词共 14 个,2010 年增加到 277 个)或学科之间相互渗透的结果(例如,知识网络领域中涉及社会网络相关分析方法的关键词共有 124 个)。

文献计量方法仅能反映出知识网络领域的文献信息演化特征,为了能从内容上揭示知识网络领域更多的演化特征,本文将借助社会网络分析软件 Netdraw 对知识网络合著网络、共词网络和作者—关键词演化网络进行可视化分析和探讨。

2.2.2 知识网络研究领域合著情况演化分析 通过学者共同署名发表科研论文建立的关系称为合著关系,基于合著关系构建的网络称为合著网络^[16]。合著论文的作者是社会关系网络中的节点,作者之间的合著关系为网络的边。合著网络是最具代表性的社会网络之一,某研究领域合著网络的演化过程,可以发现随着时间的演进,该领域具有代表性的科研人员和学术合作社团之间的相互联系、变化和发展。

将表 1 中第 1~3 列的知识网络作者共现关系数据表达成成为社会网络分析软件 Ucinet^[17] 可识别的文件,再运用可视化分析软件 Netdraw 绘制出网络图,见图 3。

图 3 中节点的不同形状代表节点的不同属性:圆形为“年份”,三角形为“作者”。节点大小代表该点网络中心

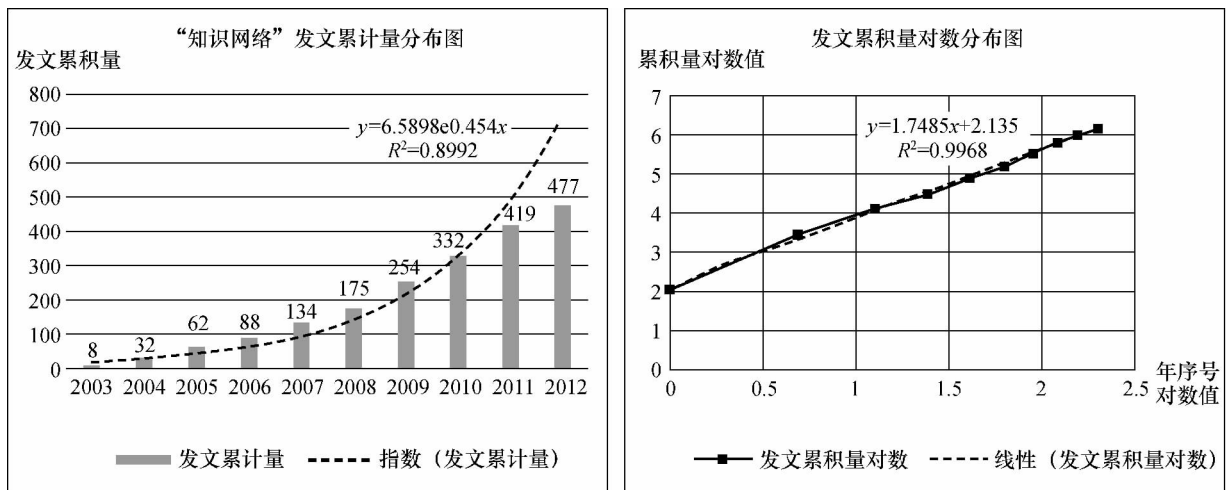


图2 知识网络文献信息特征分布图

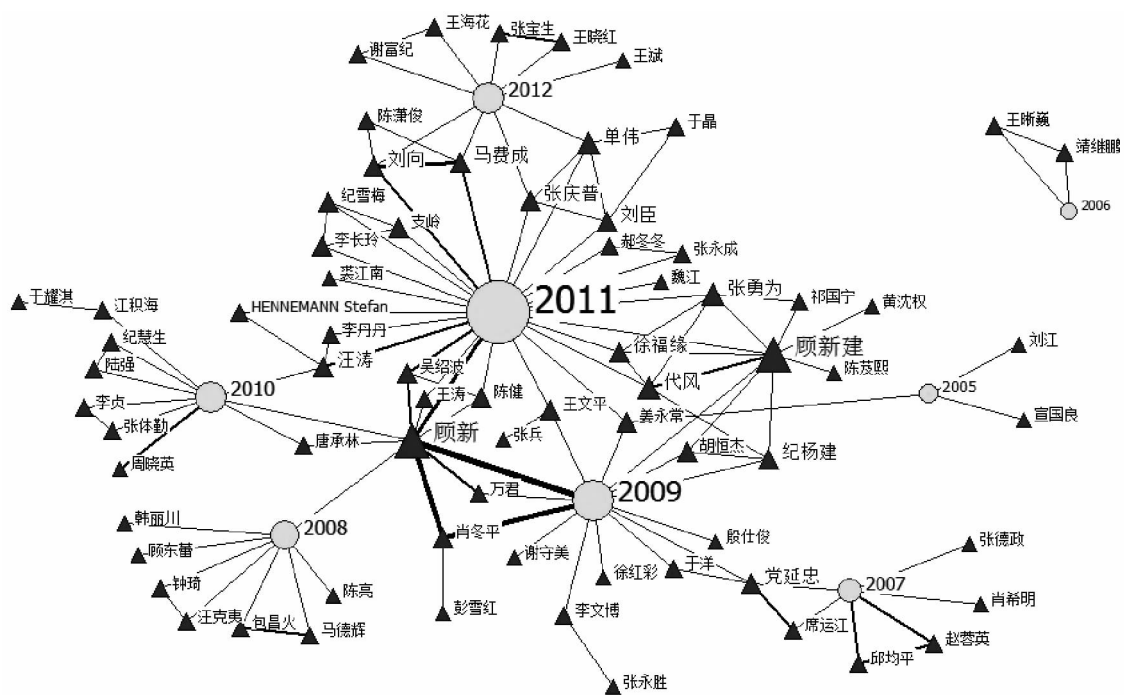


图3 知识网络合著演化网络图 (阈值=2)

度的大小, 节点之间的连线代表两两共现关系, 连线粗细代表共现频次的高低。从图中可以看出:

1) 知识网络研究领域的主要研究者与合作团队基本形成, 并且研究人员与合著团队逐年产生和扩张过程明显。图3的阈值为2, 即共现次数达到2次及以上的节点, 才会在网络图中出现。2003年与2004年是知识网络领域研究的起始年份, 发文数量较少, 基本没有作者发表2篇及以上文章, 所有2003与2004年的节点没有出现在网络图中。2005与2006年有少数作者一年内发表2篇及以上论文, 但研究规模较小。党延忠2007年与席运江合作过2

篇以上的文章, 2009年又与于洋产生合作。顾新在2008年还没有共同发文两次以上的合作者, 到了2009年有了稳定团队成员肖冬平、万君和彭雪红, 2010年成员变为唐承林, 2011年又与陈健、王涛和吴绍波合作, 反映出合著团队的发展和成员的更迭交替的过程。顾新建、代风和纪杨建在2009年一起合作过, 在2011年与新的成员徐福缘、张勇等为共同发文, 体现出这一团队的不断发展和扩张。马费成和刘向团队、张庆普、单伟和刘臣团队是知识网络研究领域新产生的合作团队, 2011、2012年连续发文2篇以上。因此, 可以看出: 党延忠、顾新、顾新

建、马费成、张庆普等是知识网络研究领域的核心研究者，他们所带领的团队是该领域的核心研究团队。

2) 稳定的、持续合作的研究团队没有形成。从图3中可以看出：核心作者节点所在合著团队成员更新频繁，合著团队融合和分裂现象不突出。核心作者能够持续关注该领域，但其团队成员研究持续性不高。另外，本文数据共有697位作者，571位作者撰写了1篇知识网络方面的研究文献，约占到作者总数的82%。并且，图3中，大多数发表2篇及以上的作者也仅与其中一个年份相连，表明大多数作者并没有持续研究该领域，但这些作者为丰富

和扩展该领域的研究内容起到了重要作用。

2.2.3 知识网络研究领域共词演化分析 在动态复杂的领域主题网络中，每个时期都对应着一定数量规模的关键词网络。动态跟踪关键词演进，可以更有效揭示该领域知识发展变化及其相互作用的特征与规律性，能更好地发现可能的新知识增长点，追溯学科发展轨迹，对学科未来发展进行超前思维和战略决策^[18]。

将表1中第4~6列的知识网络相关年份和关键词共现关系数据制作成Ucinet可识别的文件，用Netdraw软件进行可视化，阈值取3，网络图见图4。

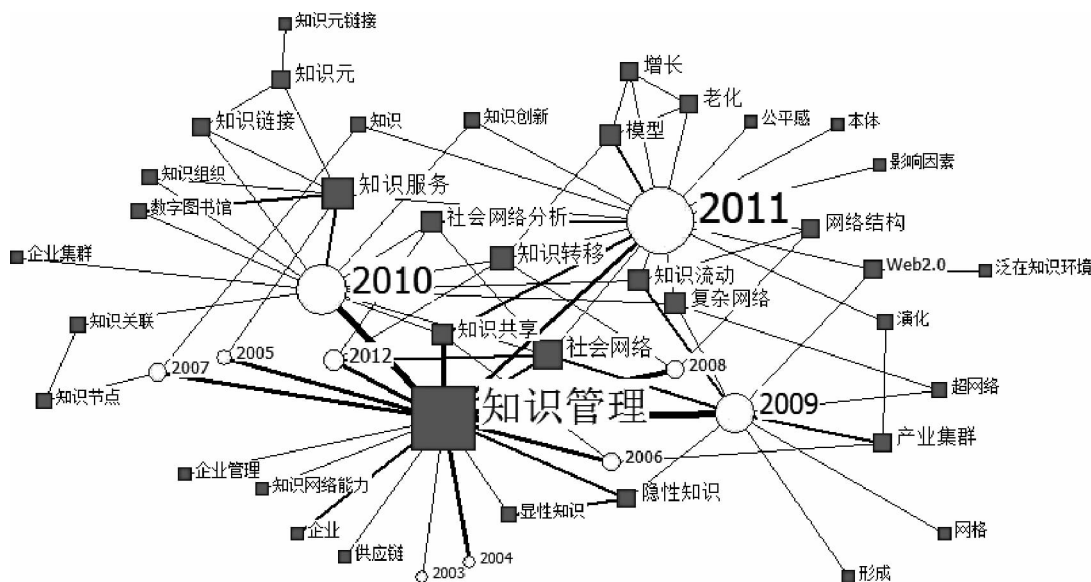


图4 知识网络主题演化网络图 (阈值=3)

图4中圆形为“年份”，方形为“关键词”，节点和连线的含义同图3。总的来看，这10年间知识管理、知识共享等一直是知识网络领域研究的经典问题，社会网络及分析方法、知识网络的演化机制处于网络中的中心位置，是近几年的研究热点。具体来说：

1) 知识网络研究内容不断深入。知识网络是知识管理的深入研究，是其分支内容，所以在知识网络研究初期的2003—2004年，知识网络的相关文献只与知识管理密切联系；2005—2008年，研究内容逐步深入到知识服务、知识共享、知识节点与关联、知识转移与知识网络结构；2009—2012年，社会网络、复杂网络的理论与方法渐渐应用到知识网络研究的各个领域，包括2009年的产业集群、2010年的知识组织与知识链接、2011年的网络结构分析等。

2) 知识网络相关新研究主题产生频繁。一方面，知识元链接、超网络、供应链、知识网络能力、社会网络分析、公平感和本体等新的概念随着时间的演进被引入到知

识网络研究领域当中，使得知识网络研究研究内容不断丰富，研究范围更加广泛，方法更加多元。另一方面，由于知识网络应用环境和领域的变化，泛在知识环境和数字图书馆也出现在网络中。

3) 知识网络核心主题融合和分裂现象突出。图4中，多为2个关键词共现，但也出现了少数3个关键词共现的现象，例如显性知识、隐性知识与知识管理共现；知识元、知识链接与知识服务共现；增长、老化与模型共现，这说明知识管理中的显性与隐性知识共享问题研究、知识元的知识链接在知识服务中的应用研究，以及知识增长与老化的模型研究，是知识网络领域的主要核心研究内容。

2.2.4 知识网络领域年份—作者—关键词3-mode共现演化网络分析 将表2的知识网络3-mode共现关系数据文件导入Netdraw中，绘制演化网络图，取阈值为3，网络图见图5。

通过图5，结合表2中年份、作者和关键词的3-mode共现关系和源数据，探讨知识网络时间序列演化情况，可

创新网络中知识优势问题，基于超循环的知识网络演化机理，知识网络知识优势的种群生态学模型。江积海研究了企业网络知识传导绩效和知识市场运行因素，周晓英就知识链接理论进行了深入研究，发表相关论文 4 篇。2011 年顾新、陈健和吴绍波讨论了知识网络公平感形成及影响因素。王涛和顾新研究了知识网络组织之间相互信任的构建和影响因素等相关问题。姜永常等提出了知识构建的 E-knowledge 机制，基于知识元本体语义链接的知识网络实现流程。刘向和马费成在《情报学报》上连续发表了 3 篇论文，深入探讨了知识网络的增长、老化和链接机制等演化问题。顾新建和代风等构建了面向发明问题解决理论的知识网络概念模型。汪涛和李丹丹运用社会网络分析和地理信息系统分析软件研究了航空航天科学技术领域知识流动的社会网络和空间网络。2012 年，一方面，社会网络及分析方法被科研人员广泛应用于知识网络领域，如农业产业集群、研发团队创造力、知识社区等；另一方面，知识转移机制、创新路径和均衡模型等内容进入学者的视野。

值得一提的是,2012年,这一领域的发文数量、合著网络和研究主题的规模增加的趋势都有所减缓,这或许是研究人员在这一领域研究的热度下降,知识网络研究主题可能进入“非常时期”所致。通常情况下,到了一定阶段,某一研究领域的发展可能出现“饱和现象”,科研成果数量下降,增速减缓,文献信息特征似乎开始不符合

— 109 —

指数增长规律,这一时期被称为领域科学主题发展的“非常时期”。而从另一个角度看,这或许也是因为该领域主题的科研水平达到了一定深度,发展进入相对成熟的时期^[16]。

总之,通过分析领域主题文献信息特征和演化特性可以发现:领域主题团队科研合作和主题发展都与论文产出之间呈正相关关系;核心作者持续关注该领域和所在团队成员的不断更新也是保持团队稳定、主题不断持续发展、充满活力的重要因素;已有的领域主题概念是新主题产生和发展基础,领域主题的发展具有创新性和递进性。

3 结束语

本文在前人研究的基础上,运用社会网络分析方法和3-mode 共现分析,对知识网络领域的研究演化规律进行了分析。研究发现,时间序列演化分析能够从宏观的角度发现历年具有代表性的作者及合著团队、经典研究主题和新的知识增长点等,并能揭示学科主题的演化轨迹。

社会网络分析方法常用于核心作者识别,高频词共现网络绘制,科研团队凝聚子群分析,而很少有学者将其应用于综合分析作者及研究主题随时间演进的发展变化规律。本文以知识网络为例,构建了年份、作者和关键词3-mode 共现词组。结合社会网络可视化分析方法,从合著网络、关键词网络和年份—作者—关键词3-mode 共现关系网络3个角度对知识网络演化规律进行了揭示。研究认为,这种方法能够推广到整个学科中,综合分析学科主题中研究人员和研究主题关系在时间维度上的发展、变化趋势以及不同主题之间的交叉重叠内容。但本文缺乏因阈值的设置导致数据样本损失的补充方法,并没有对主题演化过程中的分裂现象做过多的讨论。主题演化规律的分析还应当包括对网络拓扑结构的形成机理的解释及对未来演化方向的预测,因此网络演化机制探究是本文后续研究的主要内容。□

参考文献

- [1] 韩真. 基于共词分析的主题类型划分方法比较研究 [J]. 图书馆, 2009 (2): 46-47, 53.
- [2] BARABÁSI A L, JEONG H, NÉDA Z, et al. Evolution of the social network of scientific collaborations [J]. Physica A: Statistical Mechanics and Its Applications, 2002, 311 (3): 590-614.
- [3] CSÁRDI G, STRANDBURG K, ZALÁNYI L, et al. Estimating the dynamics of kernel-based evolving networks [EB/OL]. [2014-06-12]. <http://www.docin.com/p-112537975.html>.
- [4] LEYDESDORFF L. What can heterogeneity add to the scientometric map? Steps towards algorithmic historiography [EB/OL]. [2014-06-12]. <http://arxiv.org/abs/1002.0532>.
- [5] 李金华, 孙东川. 创新网络的演化模型 [J]. 科学学研究, 2006 (1): 135-140.
- [6] 万江平, 何振辉, 郑楚卫. 基于词频法的我国软件质量研究主题分析 [J]. 情报杂志, 2007 (10): 138-140.
- [7] 王福生, 杨洪勇. 作者科研合作网络模型与实证研究 [J]. 图书情报工作, 2007 (10): 68-71.
- [8] 刘向, 马费成, 陈潇俊, 等. 知识网络的结构与演化——概念与理论进展 [J]. 情报科学, 2011 (6): 801-809.
- [9] 沈思, 王东波, 张祥, 等. 基于文献时间特征的学科主题演化分析方法研究——以图书情报学领域为例 [J]. 图书情报工作, 2013 (7): 81-88.
- [10] 李长玲, 纪雪梅, 支岭, 等. 社会网络分析方法在科技评价中的应用研究 [J]. 科学与管理, 2012 (4): 78-82, 92.
- [11] PALLA G, BARABÁSI A L, VICSEK T. Quantifying social group evolution [J]. Nature, 2007, 446 (7136): 664-667.
- [12] 程齐凯, 王晓光. 一种基于共词网络社区的科研主题演化分析框架 [J]. 图书情报工作, 2013, 57 (8): 91-96.
- [13] SAMOYLENKO I, CHAO T C, LIU W C, et al. Visualizing the scientific world and its evolution [J]. Journal of the American Society for Information Science and Technology, 2006, 57 (11): 1461-1469.
- [14] PERSSON O, DANELL R, SCHNEIDER J W. How to use Bibexcel for various types of bibliometric analysis [C]. Celebrating Scholarly Communication Studies. Leuven, Belgium: International Society for Scientometrics and Informetrics, 2009: 9-24.
- [15] 邱均平. 信息计量学 [M]. 武汉: 武汉大学出版社, 2007: 46-48, 60-62, 93-94.
- [16] NEWMAN M E J. Scientific collaboration networks. I. Network construction and fundamental results [J]. Physical Review E, 2001, 64 (1).
- [17] BORGATTI S P, EVERETT M G, FREEMAN L C. Ucinet for Windows: software for social network analysis [M]. Harvard, MA: Analytic Technologies, 2002.
- [18] 王春秀, 冉美丽. 学科主题演化定量分析的理论基础探析 [J]. 现代情报, 2008 (6): 48-50.
- [19] 赵蓉英. 知识网络及其应用 [M]. 北京: 北京图书馆出版社, 2007: 47-52.

作者简介: 李长玲, 女, 1969年生, 教授, 硕士生导师, 发表论文70余篇。

刘非凡, 男, 1991年生, 硕士生, 发表论文5篇。

魏绪秋, 男, 1990年生, 硕士生, 发表论文2篇。

收稿日期: 2014-05-29