

当代知识图谱的构建方法研究

梁晓婷 奉国和

摘 要 对构建知识图谱常用的共现分析、引文分析、多元统计分析、社会网络分析和可视化分析方法的基本原理、使用工具、优劣势和适用性进行阐述,以期引导学者选择合理的方法及工具进行学术分析,提升知识图谱应用水平。

关键词 知识图谱 构建方法 综述

Review of Knowledge Mapping Construction Method

Liang Xiaoting, Feng Guohe

Abstract This paper studies the commonly used knowledge map construction methods including co-occurrence analysis, citation analysis, multivariate statistical analysis, social network analysis, and visual analysis method. It describes the basic principles, the tools, their advantages and disadvantages, and their applicability in the hope of providing reference for scholars to select the most suitable method and toolkits for academic study.

Key words Knowledge mapping, Construction method, Research review

0 引言

笔者于2012年7月3日以“知识图谱”为检索词,在CNKI全文数据库进行题名检索,得到2005-2012年相关文献199篇。从文献数量来看,近几年发表的文献占据了大多数,并呈现逐年增长的趋势。从学科来源看,文献主要分布在科学学、图情学和计算机学三大学科。从知识图谱的发展起源来看这种学科分布并不足为奇,科学计量学、文献计量学和信息计量学属于同一门学科——科学计量学。^[1]在对知识图谱如火如荼研究之时,必须看到研究中存在的问题:国内学者主要利用已有的工具和方法对某个学科或主题进行应用研究,而对知识图谱构建方法的研究寥寥无几,没有形成系统清晰的理论基础,导致学者在研究中存在大量问题。针对这种情况,对知识图谱进行深层次的理论研究便显得迫切和必要。国内不少学者已对知识图谱发展历史、基本概念、应用领域、关键技术、绘制方法和相关软件进行了综述^[1-4],概述全面但缺乏针对性。论文侧重对常用的共现分析、引文分析、多元统计分析、

社会网络分析和可视化分析方法的相关理论进行梳理,就各种方法的优劣势和适用性进行详细阐述,丰富知识图谱理论知识,引导研究者选择合理的方法及工具进行恰当的学术分析,提高国内知识图谱研究质量。

1 常用的知识图谱构建方法

1.1 引文分析方法

1.1.1 引文分析方法基础理论

引文分析是利用各种数学及统计学的方法、比较、归纳等逻辑方法,分析科学期刊、论文、著者等引用与被引用情况,以便揭示其数量特征与内在规律的一种文献计量法。引文分析主要用于^[5]:评价期刊、论文、机构、作者及地区的科研水平;揭示学科的发展与联系并展望未来前景等;揭示科学的结构和进行文献检索。与引文分析相关的概念还有同被引、文献耦合、作者耦合和作者关键词耦合等。

(1) 同被引分析。同被引分析通过研究对象之间形成的相对位置信息和相互关系的亲疏

程度，来研究其发展规律，预测其发展趋势。根据共引对象的不同，分为论文同被引、期刊同被引、学科同被引、作者同被引等。论文同被引揭示一个学科专业发展的来龙去脉；期刊同被引揭示各学科期刊之间的相互关系和结构特征；学科同被引分析可以揭示某些学科间的相互交叉和依赖关系；作者同被引揭示学科专业人员之间的联系和结构特点，并反映出他们所从事的学科专业之间的联系及其变化趋势，是使用最频繁的一种方法。

(2) 文献耦合。美国学者 M.M.Kessler 发现越是学科或者专业内容相近的论文，其参考文献中包含的相同文献数量就越多。他将共同的参考文献称为耦合论文，并把它们之间的这种关系称为文献耦合。用耦合强度或耦合频次（即两篇论文共同引用文献的篇数）度量两篇论文的相似性程度，耦合强度越强，越相似。^[5] 文献耦合在研究学科前沿、发现学科知识结构、信息检索等方面都有重要应用。

(3) 作者耦合。作者耦合是在文献耦合的基础上建立的，两个作者共同引用的文献越多，说明他们的研究兴趣越相近。作者之间的耦合强度会随着时间的改变而改变，反映的是作者之间的动态结构。对作者耦合的研究可以揭示一个学科领域的学者研究兴趣的发展变化。与作者同被引相比，作者耦合能够较好地分析一个学科的知识结构，尤其是挖掘一个学科的“前沿知识结构”；而作者同被引更好地分析一个学科的“基础知识结构”，这是由于作者耦合和作者同被引分析所分析的基础文献时限不同所致。作者耦合分析的文献相较作者同被引更具有时效性，两者结合能够更好地发现一个学科的知识结构。^[6]

(4) 作者关键词耦合。作者关键词耦合利用作者作品的关键词耦合强度建立作者之间的关系，其耦合强度是指两个作者拥有的相同关键词的数量。使用相同术语的作者所形成的网络可以看作是一个研究领域，因此通过共同关键词而联系到一起的作者，也就可以被看作一个研究领域了。^[7] 与作者同被引方法相比，其是利用作者现在所发表的文献进行分析，而不需要等到文献被引用才建立起关系，所以它更

能反映出目前的研究现状。与作者耦合相比，其分析的粒度更细，分析的结果更切合实际。

1.1.2 引文分析法应用研究

文献 [8-9] 利用引文分析法分析期刊馆藏资源建设，确保期刊资源保障率。在实际应用中，通常通过构建共被引矩阵进行多元统计分析、社会网络分析。文献 [10-11] 从作者共被引的角度入手，运用多元统计分析方法分析国内竞争情报、制浆造纸领域的研究状况。文献 [12] 利用关键词检索与专利引文检索相结合的方法分析 DVD 激光头技术发展中的主要竞争对手及其技术优势。文献 [13] 从引文分析法的角度分析图书情报和管理学被引半衰期。文献 [14] 通过对“211”工程高校图书馆的论文被引情况的统计，揭示“211”工程高校图书馆的学术研究状况。从上述应用研究来看，国内主要通过共引分析方法对某一专业领域进行实证研究，与国外相比：国外研究侧重于引文分析理论研究，如新方法、工具、系统等探索，国内研究停留在利用传统的方法进行实证研究，如多元统计方法。

在引文分析法中，以下问题争议比较大：

(1) 高频词选择问题。一种是结合研究者的经验在选词个数和词频高度上予以平衡，大多是作者根据个人经验设定的，理论上没有一个标准的阈值确定规则。另一种是结合齐普夫第二定律低频词分布规律判定高频词的界限。(2) 共被引矩阵对角线取值问题。目前多采用最大共被引次数+1。^[15] (3) 相似系数问题。相似性计算方法众多，每种方法适于何种情况，哪种方法测度更准确，这些还都有待商榷。国内学者在进行矩阵标准化时，大多采用皮尔逊系数。皮尔逊系数考察的是两个变量间的相关程度，适用于两个变量的总体是正态分布，或接近正态的单峰分布情况，但共引矩阵的数据并非正态分布，有学者建议采用 Cosine 和 Chi-Squared 两种方法代替皮尔逊系数。^[16]

1.2 共现分析方法

1.2.1 共现分析方法基础理论

共现分析是将各种信息载体中的共现信息定量化的分析方法。通过共现分析，人们可以发现研究对象之间的亲疏关系，挖掘隐含的或

潜在的知识,并揭示某一学科或学科群的结构与演化。根据研究对象的不同,有作者共现,机构共现,词频共现等方法。作者共现主要用于分析研究者之间的合作现象,机构共现是在作者共现的基础上建立的用以分析机构之间的合作现象以揭示学科之间的交叉现象。为了揭示科学内在的结构演变状况,主要对内容层面的特征进行共现分析,即共词分析。

共词分析是一种内容分析方法,它通过对一组词两两统计它们在同一文献中出现的次数,并以此为基础对这些词进行聚类分析,从而反映出这些词之间的亲疏关系,进而分析这些词所代表的学科研究热点、主题的结构变化和转移趋势。其是对当前发表文献的直接统计,所寻找的是当前论文所集中关注的主题,适合寻找前沿领域。因为前沿领域的研究往往人数众多而不集中,作品比较分散,被引用情况不稳定,而关键词却很好地体现了该学科的研究热点和发展方向。^[17]而共引分析得出的结论往往不能反映新文献所代表的研究主题,从而其更适于研究成熟学科的研究范式,不适于学科的研究热点研究。

1.2.2 共现分析方法应用研究

在实际应用中,通常利用关键词共词矩阵进行多元统计分析和社交网络分析。文献[18-23]利用共词探讨了国内数字图书馆、国内开放存取、Web2.0环境下知识共享、信息生命周期、国内自然语言处理和国外电子政务领域的研究热点。文献[24]通过因子分析、聚类分析和多维尺度分析提炼出国外高校图书馆核心价值的体系结构和关键要素。文献[25-33]利用共词分析并结合多元统计方法揭示国内竞争情报的演进、信息资源管理研究现状、国内知识服务研究现状、国外图书馆学情报学研究现状、国内发展心理学研究热点、食品安全研究结构、数字图书馆研究进展和用户信息行为研究进展。文献[34]采用第三代共词分析技术进行技术研发主题分析,引入密度指标和中央指标并绘制战略图以反映技术主题的研究地位。

共现分析方法在应用中应注意:(1)词语类型选择问题。关键词的选择具有很强的主观性和随意性,汉语词中普遍存在的同义词和多

义词,导致结果并不真实客观。(2)高频孤立词问题。高频孤立词是指当一个高频词与多个低频词相关,由于低频词出现次数较低,未被选作为聚类分析的词语,导致只有这个高频词出现的情况,而在聚类分析时,所有的词又必须被归为某类,使得该词被归入到了一个不相关的类的情况。(3)共现词对的“同量不同质”的问题。一篇文献的多个关键词之间的相关度并不相同,这就是词对的“同量不同质”的问题。文献[35]针对上述问题,提出了一种基于语义的共词分析法,解决了共词分析中的问题。文献[36]提出了一种基于关键词加权的共词分析方法,通过在关键词词频统计和词对相似度计算两个步骤中使用的改进的加权算法,从而实现了基于关键词重要性的加权。

1.3 多元统计分析方法

1.3.1 多元统计分析方法基础理论

多元统计分析是从经典统计学中发展起来的一个分支,通常采用因子分析、聚类分析、多维尺度分析和战略坐标图分析等方法。

(1)因子分析的基本原理是以较少几个因子反映原始资料的大部分信息,从而达到数据精简的目的。其基本过程需要注意的是考察原始变量之间是否存在较强的相关性,是否适合进行因子分析,通常用KMO和球形Bartlett进行因子分析的适用性检验^[37]。优点:它不是对原有变量的取舍,而是根据原始变量的信息进行重新组合,找出影响变量的共同因子,化简数据;它通过旋转使得因子变量更具有可解释性。缺点:数据降维必然会丢失一部分信息。

(2)聚类分析的基本思想是利用变量间不同程度的相似性把具有相似属性和特征的对象归为一类,可以通过SPSS、SAS、Ucinet、TreeView、Matlab等来实现聚类作图^[38]。其优点在于:它利用高频词反映某学科或主题的研究热点,能定量和客观地反映出词与词之间的亲疏关系,进而反映这些词所代表的主题内容的结构。其弊端在于:没有中心概念。聚类成团的主题词,由于相互间的密切程度高而聚集起来,在聚集的过程中没有形成中心词。聚在一起的主题词未必都能表达同一个内容。由于词之间距离传递作用的存在导致聚类时一

些关联不太大的主题词聚集在一起。对类团间的相互影响没有说明。无法识别哪些是主要类团,哪些是次要类团。

(3) 多维尺度分析是通过低维空间展示对象之间的联系,并利用平面距离来反映对象之间的相似程度。图谱中高度相似的词聚集在一起,形成一个学科研究类别,越在中间的词表明与它有联系的词越多,在学科里的位置也就越核心;反之,越在外围,则越边缘。与聚类树图相比,多维尺度图谱利用平面距离展示出词间亲疏关系,能够容易判断出某研究领域在学科内的位置。缺点是多维尺度二维显示往往并不能准确地表达多维问题,同时图谱中词间界限模糊,难以分类。

(4) 聚类、因子分析、多维尺度只是对研究结构静态特征的描述,而战略坐标图能更好地展现研究热点和研究结构的发展动态。它是向心度和密度为参数绘制成的二维坐标图,一般X轴为向心度,Y轴为密度,原点为二者的均值。向心度是用来度量一个类团和其他类团相互影响的程度。密度是该类的内部强度,它表示该类维持自己和发展自己的能力。该方法显著的优点在于:能判断热点主题的核心度和成熟度,展示学科结构演变的过程及原因。但由于战略坐标图是在聚类的基础上进行的,所以前期聚类的效果好坏将直接影响坐标图的效果。^[39]

1.3.2 多元统计分析方法应用研究

文献[40]将共引分析法与战略坐标相结合对重点学科发展进行战略情报分析,通过引文共引聚类描述重点学科领域的研究历史,“引文战略坐标图”进一步了解各研究热点的新颖度和关注度。文献[41-47]利用多元统计方法对搜索引擎、国内推荐引擎、国外开放存取、我国图书馆学、国外图书情报学、情报与档案管理学科和数字信息资源进行热点分析。文献[48]将医学文献的主要主题词和次要主题词进行加权处理,在主题遴选阶段采用改进信息熵算法对主题进行遴选,最后利用共词分析中的战略坐标图对热点主题进行描绘。文献[49]利用多元统计与社会网络分析结合进行医学知识图谱绘制。在引文分析和共现分析中对多元统计分析方法的应用研究已进行总结,在

此不再重复。

多元统计分析法在应用中需注意:(1)因子分析中的类别命名,通常根据关键词贡献大小进行命名;(2)聚类数目的确定,需结合专家知识和图谱共同确定聚类数目;(3)多维尺度的类别划分,需结合聚类和因子分析确定;(4)战略坐标图的计算方法。

1.4 社会网络分析方法

1.4.1 社会网络分析方法基础理论

社会网络分析就是通过分析社会网络内部成员之间的关系和交互,发现它们的组织结构、组织特点、行为方式、个性特征等。在网络中用节点代表人或组织,用连线表示他们之间的关系。关系强弱以连线的粗细表示,关系越强,连接它们之间的线条越粗。借助社会网络分析的紧密性、中介性、中心性、桥、簇、团、丛等概念,可以分析出占重要地位的作品、作者、学科力量和群体分布等^[50]。社会网络图谱中词团间联系的强弱,通过网络关系图一目了然。以直观形象的方式反映词团间的关系,能快速定位核心词团与边缘词团。其不足之处在于:不能反映词团的成熟度,难以判定某研究领域的成长趋势,与战略坐标图相结合能够全面地分析某一主题。

1.4.2 社会网络分析方法应用研究

文献[51]提出将社会网络分析引入到引文评价中,以改进评价方法,补充评价指标。文献[50]从研究者合作发文网络、共词网络、引文网络、小团队不同时期合作网络几个方面,分析我国的情报学研究的不足之处。文献[52]提出基于语义的社会网络分析应用模型。文献[53]指出目前大部分社会网络搜索算法基于最短路径原则,忽视社会网络上主体之间社会关系的强度和方向,提出最大影响强度的路径优化原则并基于该原则构造搜索算法。文献[54]分析了组织社会网络中存在的结构洞所能带来的结构位置利益及其对组织隐性知识共享的影响。文献[55]利用社会网络对国内竞争情报进行比较研究。文献[56]利用社会网络对我国竞争情报论文进行统计分析从中间中心性的角度揭示了武汉大学作为国内竞争情报研究的中心研究地位。

社会网络分析法在应用中需注意:(1)节点的选择问题,过多的节点或线条会掩盖真实有用的信息,而过少则无法包含所研究主题的信息。(2)图谱的解释问题,不少学者注重研究分析社会网络关系的形式,较少分析关系的实质。我们应该结合实际具体深入地分析数据,得出相应的结论,只有这样社会网络分析才能在情报学中得到应用和发展。

1.5 信息可视化分析方法

1.5.1 信息可视化方法基础理论

信息可视化是利用计算机支撑的、交互的、对抽象数据的可视表示。可视化技术不仅在揭示信息资源的广度与深度上有很大的优势,还能够将隐藏在信息资源内部的、复杂的、抽象的语义以直观的图形方式呈现给用户,为用户直观、方便获取、过滤、理解大规模数据和信息提供了有效途径,从而发现信息之间的关系特征和规律。陈超美教授对可视化技术的研究加速了国内知识图谱的研究,为知识图谱研究引入新鲜血液。2004年,其开发出CiteSpace软件,通过展现“研究前沿术语的共现”、“知识基础文章的同被引”、“研究前沿术语引用知识基础文章”等三个网络随着时间演变的情况来寻找研究热点及趋势,并以可视化的方式展示出来,一直对软件进行更新增添新功能,最新版本Citespace 3.1 R3。信息可视化方法能够生动形象地展示一个学科领域的发展演进、研究热点和研究前沿,以“快照”的形式将其发展记录下来。

1.5.2 可视化方法应用研究

关于可视化的研究,目前较多。主要分为:(1)进行专利分析。文献[57]借助分析工具CiteSpaceII,通过文献共引和共词分析的方法,对未来专利研究的热点进行预测。文献[58]运用信息可视化技术和社会网络分析方法对世界500强企业进行技术竞争的专利地图分析,挖掘技术竞争网络中的关键企业。(2)用于学科与主题分析。文献[59]利用CiteSpaceII工具对知识管理研究的现状进行可视化研究,揭示知识管理研究发展的现状。文献[60]通过CiteSpace软件梳理了图书馆学的重要学术文献及学术代表人物。文献[61]对国际范围内

图书情报学领域的研究热点和前沿进行辨识和追踪。文献[62]对国内外企业知识共享的研究热点、前沿及发展趋势进行探讨,为我国企业知识共享的研究提出建设性意见和建议。文献[63]运用知识图谱方法描绘国际网络搜索引擎领域该领域近年来的研究现状及热点。文献[64]mashup利用信息可视化软件CiteSpace绘制出mashup研究的关键期刊、文献的知识图谱,并通过分析文献关键词分布和检测词频变动趋势确定出mashup的研究热点。文献[65]利用信息可视化工具Citespace研究国内机构知识库的演进过程、各阶段研究热点及其变化情况。

从目前的应用中发现,存在下列问题:研究缺乏标准;在数据库的选择、检索式的构造、时间跨度大小、样本大小、样本代表性、结果有效性评估、科学知识图谱解释方面还没有统一的定论[66]。研究过程粗糙,结果趋于浅显,如关键字合并问题,工具中提供了主关键词和次关键词进行合并,却很少有文献使用。引文数据库不完善导致分析困难大。

2 总结

论文针对近几年知识图谱理论研究不足、实证研究过于盲目的现象,及时地对构建知识图谱常用的共现分析、引文分析、多元统计分析、社会网络分析和可视化分析方法的基本原理、使用工具、优劣势及适用性进行阐述以期指导学者选择合理的方法及工具进行恰当的学术分析,提高国内知识图谱论文的质量。各种知识图谱方法的构建原理和特点不同,分析得到的结果自然不同。因此,在进行数据分析时必须根据研究目的来选择知识图谱构建方法。如只想了解某学科热点的主题结构,可用共词的聚类分析得到树状图来展示;如果想判断出某研究领域在学科中所处位置,可以结合多维尺度分析;若要判定学科的成熟度和核心性以及该学科演化情况,可以制作战略坐标来分析;如果想知道学科中各主题之间的关系,可以通过共词网络来展现,并且还可以加入社会网络分析,来挖掘其核心和边缘主题。当然,也可以多种可视化方法相结合,从不同的角度来分析,多方位展现学科结构的发展过程。

参考文献

- [1] 秦长江, 侯汉清. 知识图谱——信息管理与知识管理的新领域[J]. 大学图书馆学报, 2009(1):30-37.
- [2] 梁秀娟. 科学知识图谱研究综述[J]. 图书馆杂志, 2009, 28(6):58-62.
- [3] 廖胜姣, 肖仙桃. 科学知识图谱应用研究概述[J]. 情报理论与实践, 2009, 32(1):122-125.
- [4] 陈悦, 刘则渊, 陈劲等. 科学知识图谱的发展历程[J]. 科学学研究, 2008, 26(3):450-460.
- [5] 邱均平. 信息计量学[M]. 武汉:武汉大学出版社, 2007:315-427.
- [6] 马瑞敏. 作者耦合分析:一种新的学科知识结构发现方法的探索性研究[J/OL]. 中国图书馆学. DOI: CNKI:11-2746/G2. 20111026. 1932. 001. [2012-07-03]. <http://www.cnki.net/kcms/detail/11.2746.G2.20111026.1932.001.html>.
- [7] 刘志辉, 张志强. 作者关键词耦合分析方法及实证研究[J]. 情报学报, 2010, 29(2):268-275.
- [8] 李海霞, 刘贵吾. 引文分析法在图书馆期刊资源建设中的应用研究——以哈尔滨工程大学为例[J]. 情报理论与实践, 2012, 35(1):99-102.
- [9] 李峰. 使用引文分析法考察图书馆文献保障情况[J]. 大学图书馆学报, 2011(5):104-108.
- [10] 宋艳辉, 邱均平. 作者共被引视角下的我国竞争情报研究[J]. 情报资料工作, 2011(5):17-20.
- [11] 邱均平, 秦鹏飞. 基于作者共被引分析方法的知識图谱实证研究——以国内制浆造纸领域为例[J]. 情报理论与实践, 2010, 33(10):53-56.
- [12] 孙涛涛, 刘云云. 基于专利耦合的企业技术竞争情报分析[J]. 科研管理, 2011, 32(9):141-146.
- [13] 黄利平. 基于引文分析的文献老化研究——以图书情报和管理学为例[J]. 情报杂志, 2011, 30(10):31-35.
- [14] 鄂丽君, 李茂. 基于引文分析的高校图书馆学术研究[J]. 情报科学, 2010, 28(10):1518-1521.
- [15] 邱均平, 马瑞敏, 李晔君. 关于共被引分析方法的再认识和再思考[J]. 情报学报, 2008, 27(1):69-74.
- [16] 王玲玉. 文献共引分析中的关联强度和相关系数改进研究[D]. 郑州:郑州大学, 2011.
- [17] 伍若梅, 孔悦凡. 共词分析与共引分析方法比较研究[J]. 情报资料工作, 2010(1):25-28.
- [18] 董伟. 国内近十年数字图书馆领域研究热点分析——基于共词分析[J]. 图书情报知识, 2009(5):59-63.
- [19] 李武, 董伟. 国内开放存取的研究热点:基于共词分析的文献计量研究[J]. 中国图书馆学报, 2010, 36(190):105-115.
- [20] 国内外Web2.0环境下知识共享研究主题分析——基于共词分析法的研究[J]. 情报科学, 2012, 30(1):126-131.
- [21] 游毅, 索传军. 国内信息生命周期研究主题与趋势分析——基于关键词共词分析与知识图谱[J]. 情报理论与实践, 2011, 34(10):17-21.
- [22] 杨皓东, 江凌, 李国俊. 国内自然语言处理研究热点分析——基于共词分析[J]. 图书情报工作, 2011, 55(10):112-117.
- [23] 董伟, 贾东琴. 国外电子政务研究进展分析——基于共词分析方法的研究[J]. 图书情报工作, 2011, 55(21):125-142.
- [24] 陆宇杰, 张凤仙, 范并思. 基于共词分析的高校图书馆核心价值研究[J]. 大学图书馆学报, 2011(6):34-40.
- [25] 李颖, 贾二鹏. 基于共词分析的国内竞争情报研究演进态势[J]. 现代情报, 2011, 31(4):82-93.
- [26] 张伟, 鲁荣辉, 王宇. 基于共词分析的国内信息资源管理研究现状分析[J]. 情报杂志, 2009, 28(12):84-85.
- [27] 张凌, 周春雷, 寇广增. 基于共词分析的国内知识服务研究[J]. 图书情报工作, 2009, 53(24):64-68.
- [28] 郭璇. 基于共词分析的国外图书馆情报学领域研究现状探析[J]. 情报杂志, 2011, 30(11):38-41.
- [29] 王莉亚, 张志强, 卫军朝. 基于共词分析的近十年国外图书情报学研究主题分析[J]. 情报杂志, 2011, 30(3):51-58.
- [30] 许梅华. 基于共词分析的近年国内发展心理学研究热点分析[J]. 现代情报, 2010, 30(8):171-175.
- [31] 赵美玲, 郭军康. 基于共词分析的近十年我国食品安全研究结构探讨[J]. 情报杂志, 2009, 29(6):23-25.
- [32] 苏娜. 基于共词分析的数字图书馆领域研究主题及进展分析[J]. 情报杂志, 2009, 28(6):16-19.
- [33] 肖明, 杨楠, 李国俊. 基于共词分析的我国信息用户行为研究结构探讨[J]. 情报杂志, 2010, 29(12):13-15.
- [34] 韩红旗, 安小米, 朱东华, 汪雪锋. 专利技术术语共现的战略图分析方法[J]. 计算机应用研究, 2011, 28(2):577-579.
- [35] 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):146-149.
- [36] 李纲, 李轶. 一种基于关键词加权的共词分析方法[J]. 情报科学, 2011, 29(3):322-324.
- [37] 杨颖, 崔雷. 基于共词分析的学科结构可视化表达方法的探讨[J]. 现代情报, 2011, 31(1):91-96.
- [38] 吴亚娟. 基于因子-聚类分析的大学生就业满意度统计及预测[J]. 南京信息工程大学学报:自然科学版, 2010, 2(6):510-513.
- [39] 杨颖, 崔雷. 基于共词分析的学科结构可视化表达方法的探讨[J]. 现代情报, 2011, 31(1):91-96.
- [40] 崔雷, 杨颖, 王孝宁. 重点学科发展战略情报研究(一)——引文战略坐标[J]. 情报理论与实践, 2009, 30(6):8-13.
- [41] 刘阳, 宋余庆. 搜索引擎学术研究知识图谱[J]. 图书情报知识, 2010(6):105-110.
- [42] 奉国和, 梁晓婷. 国内推荐引擎学术研究知识图谱分析[J]. 情报科学, 2012, 30(1):145-160.

- [43] 苏小波, 张新兴. 2001-2009年国外开放存取研究内容分析[J]. 国家图书馆学刊, 2011(1):78-83.
- [44] 宗乾进, 袁勤俭, 沈洪洲. 2001-2010年国内情报学研究回顾与展望——基于知识图谱的当代学科发展动向研究[J]. 情报资料工作, 2012(1):10-15.
- [45] 郭春侠, 叶继元. 基于共词分析的国外图书情报学研究热点[J]. 图书情报工作, 2011, 55(20):19-22.
- [46] 王兰敬. 2004-2009年我国图书馆、情报与档案管理学科的研究热点与重点领域——基于CSSCI来源文献关键词的分析[J]. 图书情报工作, 2011, 55(16):68-71.
- [47] 马费成, 望俊成, 陈金霞, 胡超. 我国数字信息资源研究的热点领域:共词分析透视[J]. 情报理论与实践, 2007, 30(4):438-444.
- [48] 安新颖. 基于改进信息熵的干细胞研究领域共词分析[J]. 图书情报工作, 2011, 55(2):37-40.
- [49] 沈建通, 姚乐野. 多元统计与社会网络分析法在知识图谱应用的实证研究[J]. 情报杂志, 2009, 28(8):34-36.
- [50] 刘蓓, 袁毅, BOUTINE. 社会网络分析法在论文合作网中的应用研究[J]. 情报学报, 2008, 27(3):407-417.
- [51] 宋歌. 社会网络分析在引文评价中的应用研究[J]. 图书情报工作, 2010, 54(14):16-19.
- [52] 程慧平, 余凡. 基于语义web的社会网络分析应用模型研究[J]. 情报科学, 2011, 29(4):615-618.
- [53] 惠淑敏. 基于影响强度的社会网络搜索算法研究[J]. 图书情报工作, 2012, 56(2):111-115.
- [54] 姜鑫. 基于“结构洞”视角的组织社会网络内隐性知识共享研究[J]. 情报资料工作, 2012(1):32-36.
- [55] 周洋, 郑彦宇, 刘志辉. 国内外竞争情报研究的比较研究[J]. 情报杂志, 2011, 30(10):10-15.
- [56] 邓蓓. 基于社会网络分析的我国竞争情报论文统计分析[J]. 情报科学, 2011, 29(6):949-951.
- [57] 高继平, 丁堃. 专利研究文献的可视化分析[J]. 情报杂志, 2009, 28(7):13-16.
- [58] 王贤文, 张曦, 丁堃. 基于专利共被引的世界500强企业技术竞争的专利地图分析[EB/OL]. [2012-6-28]. <http://www.paper.edu.cn/index.php/default/releasepaper/content/201011-166>.
- [59] 赵蓉英, 许丽敏. 知识管理研究现状的可视化分析[J]. 情报理论与实践, 2010, 33(12):48-52.
- [60] 赵蓉英, 王菊. 图书馆学知识图谱分析[J]. 中国图书馆学报, 2011(37):41-50.
- [61] 邱均平, 温芳芳. 近五年来图书情报学研究热点与前沿的可视化分析——基于13种高影响力外文源刊的计量研究[J]. 中国图书馆学报, 2011(192):51-60.
- [62] 马丽娜, 赵蓉英, 王敏. 基于知识图谱的国内外企业知识共享比较研究[J]. 图书情报工作, 2011, 55(10):20-24.
- [63] 邱均平, 胡文君, 罗力. 基于知识图谱的国际网络搜索引擎研究现状与前沿分析[J]. 图书情报工作, 2010, 54(24):89-94.
- [64] 朱光, 董尹, 张华坤. 基于知识图谱的mashup研究可视化分析[J]. 现代情报, 2012, 32(2):151-155.
- [65] 奉国和, 吴敬学. 国内机构知识库研究文献的可视化分析[J]. 图书情报工作, 2011, 55(22):95-100.
- [66] 肖明, 陈嘉勇, 李国俊. 基于citespace研究科学知识图谱的可视化分析[J]. 图书情报工作, 2011, 55(6):91-95.
- 梁晓婷 女, 1987-, 华南师范大学经济管理学院信息管理系, 在读研究生。 广州 510006
- 奉国和 1971-, 博士, 华南师范大学经济管理学院信息管理系, 副教授。 广州 510006
- (收稿日期:2013-01-04 编发:王宗义)