

# 主路径分析方法研究进展\*

刘懿 周丽英

(中国农业大学图书馆, 北京 100193)

**摘要:** 主路径分析方法是一种基于“连通性”降低知识网络复杂程度, 提取关键路径的定量和可视化方法, 在科学技术发展路径探测中具有重要的实践意义。本文从概念辨析、方法流变和应用实践3个角度对国内外主路径方法研究与应用的相关文献进行系统梳理, 旨在明晰主路径分析方法的发展脉络和研究重点, 发现已有研究的不足和空白点, 总结现有方法在路径探测中存在的问题, 提出主路径探测方法改进策略。现有主路径分析方法相关研究的关注重点主要在遍历权重的计算和路径搜索方法的选取等方面, 而关键节点的遴选、多元路径的探测及稀疏网络的密度增强等能够提高主路径分析效果的改进策略, 值得在未来进行深入研究。

**关键词:** 主路径分析; 引文网络; 方法改进; 研究进展

**中图分类号:** G350

**DOI:** 10.3772/j.issn.1673-2286.2019.10.002

主路径分析方法是一种基于“连通性”降低知识网络复杂程度, 提取关键路径的定量和可视化方法, 在科学技术发展路径探测中具有重要的实践意义。它基于时间流分析网络“连通性”<sup>[1]</sup>, 以“链接”为中心, 更加关注节点“连通性”, 是引文分析方法的突破性发展。引文网络是主路径分析方法的基础, 引文分析的思想由来已久, 但直到加菲尔德创办《科学引文索引》(Science Citation Index, SCI) 才使其进入实用阶段; 普赖斯在此基础上进行深入研究, 提出科学论文引用网络, 现代科学计量学由此滥觞<sup>[2-3]</sup>。早期引文网络的分析和研究多以节点为中心, 侧重中心性分析和聚类分析<sup>[4]</sup>; 随着非循环有向图研究的兴起, 以及复杂网络的引入, 有研究者开始关注引文网络中“链接结构”的重要性以及时间变量对研究结果的影响; 1989年 Hummon等<sup>[1]</sup>提出的主路径分析, 将引文网络的研究重点由“节点”转向“链路”, 开创了引文网络“连通性”分析的先河。

本文通过对主路径概念脉络的辨析, 描绘主路径方法的源起和发展; 通过对主路径识别算法的研究, 分析主路径方法的流变; 通过对现有应用和改进的梳理

及总结, 探讨主路径方法未来可能的发展方向。

## 1 主路径概念界定

### 1.1 算法描述性概念

为解决引文分析中以文献为中心、采用聚类或定标等方法研究节点相似性时, 无法展示时间变化对科学或技术影响的问题, Hummon等<sup>[1]</sup>提出一种“两步”提取引文网络中代表性路径的方法: ①通过为每个链接分配一个名为遍历权值的值属性来区分最初同样重要的引用链接; ②采用深度优先搜索算法的变体, 搜索具有最高遍历权值的链接以获得最重要的引文链。但是, 在研究中, 关于主路径的阐述仅停留在将提取出的关键性路径称为主路径, 并没有对其进行解释说明或给出规范定义。

随后, De Nooy等<sup>[5]</sup>尝试对主路径进行概念界定, 在其著作中将主路径定义为“在引文网络中, 主路径是从源点指向汇点的一条路径, 且该路径所含弧的遍历权值之和最高”, 并将主路径分析方法引入社会网络

\*本研究得到中央高校基本科研业务项目“面向双一流建设的数据支持及服务方法研究”(编号: 2019TC090)资助。

分析软件Pajek。从字面上看,这一定义涵盖了源点、汇点、最大遍历权重等主路径的必要结构元素,明确了主路径的结构,但是分析其内在逻辑不难发现,这一定义还是停留在对算法实现结果的简单表述,仍在用“主路径是怎样算出来的”来回答“主路径是什么”这一问题,并没有将“主路径”概念化。

## 1.2 路径解释性定义

利用主路径分析能够追踪引文网络中关键路径

的特性,Liu等<sup>[6]</sup>将其拓展至探测研究领域的发展轨迹,并提出了全局主路径(global main path)、本地后向主路径(backward local main path)、多元主路径(multiple main path)、关键路径主路径(key-route main paths) 4种变体,以解决传统主路径分析方法固有的弊端。表1列出了4种主路径变体的定义及其作用。

图1是基于Liu等<sup>[6]</sup>提出的定义绘制的主路径新变体简单示意图,由于该图不包含子网,故未对多元主路径进行展示。其中,ACEH、ACEI、BCEH和BCEI为全局主路径,在图示网络所有的通路中,这4条路径

表1 主路径的4种变体

主路径变体	定义	作用
全局主路径	以节点在知识流动中的全局重要性为基础,提取网络中具有最大总遍历计数的路径	可以发现网络中全部的重要路径
本地后向主路径	所有从汇点指向源点的路径中,所含弧的遍历权值总计最高的一条路径	有助于探测当前活跃思想或观点的重要根源
关键路径主路径	从关键路径的两端而非源节点开始遍历搜索,以网络中拥有最高遍历计数的链路为关键路径提取主路径	使遍历计数最高的链接包含在路径中
多元主路径	基于一定准则,提取多条路径,如提取遍历计数排名前几的链接,或者提取遍历计数在容差阈值范围内的链接	有助于在下一层网络中检查重要路径

的总遍历计数最大,能反映出科学和技术的全局流动情况;ACEH和ACEI为本地后向主路径,这些路径由汇点H、I和J出发逆向进行遍历,依次选取遍历计数高的连通链路,直到抵达指定的源点A,本地后向主路径可以充分把握现有研究理论的根源和技术发端;ACEH、ACEI、BCEH和BCEI为关键路径主路径,图示网络中CE链路具有最高的遍历计数且不包含源点或汇点,是网络中的重要链接(关键路径),由CE出发分别向两端进行遍历搜索,这样提取到的路径可以确保最重要的链路在主路径上,不因强调连通性而遗漏重要链接。

除了文献6提出的主路径变体外, Lucio-Arias等<sup>[7]</sup>在综合考虑引文网络中节点属性的基础上,用节点的入度(indegree)和出度(outdegree)等中心性特征对主路径上包含的节点进行了限定,提出“主路径是基于链接在网络中的连通性,选取度数最高的节点构成的路径”。这一定义不仅考虑了路径的链接连通性,还包含了对路径节点的筛选。

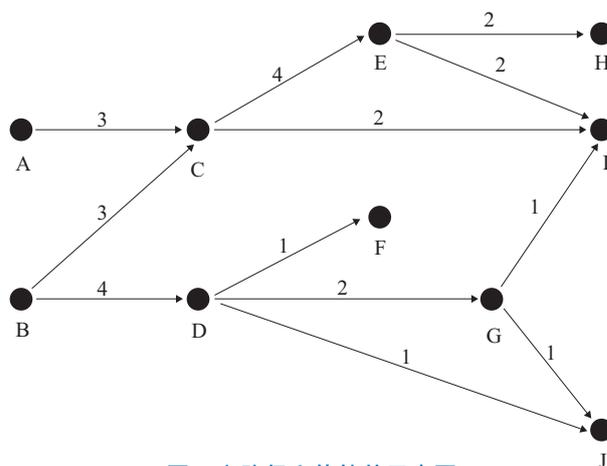


图1 主路径变体简单示意图

## 1.3 要素构成性定义

无论是为算法实现服务的算法描述性定义,还是面向既有路径解释的路径解释性定义,随着学术界对主路径研究的深入,都逐渐显露出无法满足研究需要的弊端,即基于这两种定义所提取出的主路径不能囊括预期的重要节点。

从根本上讲,主路径分析提取出的路径应该符合3

个基本要求：①主路径上的节点应该是重要的或关键的，能够体现学科领域发展的核心；②主路径在引文网络的所有路径中承载了最大信息量；③主路径体现的是节点之间紧密程度最高、最关键的关系。结合这些要素，马瑞敏等<sup>[8]</sup>将主路径定义为“反映学科领域发展脉络的重要节点之间关键关系的最大信息承载路径”。这一定义较全面地对构成主路径的实体和关系进行表述，反映了主路径的关键环节及核心要素，为后续研究奠定新的概念基础。

综上所述，主路径的概念随着研究的深入和拓展也在不断深化、发展，从简单描述到加入限制条件的理性思考，概念的变化反映出方法的变革和完善。对主路径概念的界定实质是对方法目标的一种声明，综合的主路径概念不仅揭示了符合要求的路径应具有的特征，还反映出主路径分析方法的适用场景和限制条件。

## 2 主路径分析算法研究

由于主路径概念的内涵包括节点和链接两个核心构成要素，因此本文将依据这两个核心要素将主路径分析算法划分为关键节点遴选和主路径识别两个方面。其中关键节点遴选算法主要用于从网络中筛选重要节点，对网络进行降维；而主路径识别算法是基于链接的遍历权重以及路径的搜索策略，从网络中识别核心路径。

### 2.1 关键节点遴选

关键节点的识别一直是图论的重要研究议题之一。复杂网络的兴起使得关键节点识别研究跳出传统特征识别方法和调查量表方法的局限，引入网络环境下的社会网络分析法、系统科学分析法、PageRank算法和HITS算法等识别方法。

引文网络中对于关键节点的遴选，目前主要有2种方式。①利用传统文献计量视角下的被引频次来测度节点的重要性，如Lucio-Arias等<sup>[7]</sup>曾以高被引论文为研究对象来研究学科领域演化轨迹。②利用信息搜索领域的算法来测度节点的影响力，如马瑞敏等<sup>[9]</sup>就综合利用改进PageRank和遍历权重来测度节点重要性。

主路径分析对于关键节点的遴选既要考虑节点的影响力和控制力，也要确保主路径的汇点能够体现出最新的研究成果<sup>[10]</sup>。高影响力和强控制能力从不同角

度体现了节点的重要性，保障了节点遴选的科学性和全面性，而汇点的新颖程度保障了主路径的前进性。

### 2.2 主路径识别

主路径识别涉及遍历计数、路径搜索两个关键步骤与算法。

#### 2.2.1 遍历计数算法

知识流动的关键链接或者知识演化的重要部分可以用相互连通的路径来表示。遍历计数是表征知识流动或知识演化中节点对路径连通重要程度的指标。Hummon等<sup>[1]</sup>提出搜索路径链接计数（Search Path Link Count, SPLC）、搜索路径节点对（Search Path Node Pair, SPNP）和节点对投影计数（Node Pair Projection Counts, NPPC）3种用于计算引文网络中节点显著性的遍历计数方法。Batagelj<sup>[11]</sup>认为上述3种遍历计数的原始权值分配方法计算过程复杂，可能会遗漏重要节点，因此对其进行了改进，并提出另一种简单、高效的有效权值计算方法——搜索路径计数（Search Path Count, SPC）。Verspagen<sup>[12]</sup>把主路径分析扩展应用于专利引文网络，并基于SPC遍历计数提出一种顶部路径演变网络（Network of the Evolution of Top Path, NETP）的新算法，该算法通过连接不同时间间隔的顶部路径来构造主路径，以揭示最优路径在一定时间间隔之后发生的变化。表2比较了SPC、SPLC、SPNP、NPPC等遍历计数算法，从中可以看出，NPPC不适用于大型网络，复杂网络研究兴起后该算法使用较少。根据遍历的复杂程度可以推断出遍历权值 $SPNP \geq SPLC \geq SPC$ ，SPC算法因其流入值与流出值相同的优良特性得到了广泛应用；但是结合知识流通性来看，SPLC算法更适用于科学技术发展过程中的知识扩散场景。Batagelj等<sup>[10]</sup>将路径搜索算法集成到了社会网络分析软件Pajek中，提高了主路径分析的使用效率。

#### 2.2.2 路径搜索算法

路径搜索算法是主路径分析的技术核心，也是正确提取主路径的要点之一。Hummon等<sup>[1]</sup>采用按分支进行搜索的深度优先搜索算法，从源点出发对网络中的分支进行逐一遍历，简化网络，揭示重要知识流的

表2 发现系统分面展现控制策略比较

算法	原理	知识流通特性	适用性
SPLC	在考虑搜索路径差异的前提下, 计算通过链路的节点对数量, 其路径的终点必须是汇点之一	中间节点不仅传递知识, 而且是新知识产生的起源, 这种情况会导致早期节点或链接的遍历权值较低	适用于知识和技术发展中的知识扩散场景
SPNP	在考虑搜索路径差异的前提下, 计算通过链路的节点对数量, 对路径的起点和终点没有特殊要求	中间节点既可以传递知识, 又能生产新知识, 还能储存已有知识, 因此中间节点和链接获得较高的遍历权值	适用于强调中介重要性的网络
NPPC	在不考虑搜索路径差异的情况下, 计算通过链路的节点对数量	中间节点既可以传递知识, 又能生产新知识, 还能储存已有知识, 因此中间节点和链接获得较高的遍历权值	只考虑通过路径的2个不同节点的可达性, 对遍历路径的选择并不敏感, 不适合于大型网络
SPC	在考虑搜索路径差异的前提下, 计算通过链路的节点对的数量, 其中起点必须是源点, 并且终点必须是汇点	中间节点仅起传递知识的连通作用, 既不产生新知识也不储存已有知识	适用于符合基尔霍夫节点定律的场景, 即流入遍历权值之和等于流出遍历权值之和, 适用于大型引文网络

递归过程。Liu等<sup>[6]</sup>基于贪心策略从源点向汇点进行前向搜索, 提取具有重要贡献的后代节点, 尤其是那些高被引节点, 以生成“本地主路径”, 但是这样的搜索策略方法不能很好地反应知识或技术的来源与演变。Liu等<sup>[6]</sup>还提出从汇点向源点进行反向搜索的后向搜索策略, 这种追根溯源的搜索方式可以将更多高引用的节点纳入主路径, 为主路径分析提供了独特的视角。但是, 前向搜索和后向搜索的区别仅在本地主路径中有所体现, 搜索方向的差异在全局主路径中没有意义。Batagelj<sup>[11]</sup>尝试将关键路径算法引入主路径分析, 以克服本地主路径只考虑局部最优解而遗漏网络中其他重要链接的问题。

关键路径算法的搜索目标是从源点到汇点具有最大路径长度的路径, 利用该算法可以识别出包含最大总遍历计数链接的路径。基于此, Verspagen<sup>[12]</sup>通过穷举法提取出全局主路径, 为分析主路径的重要性提供了一个全新的视角。由于单一的主路径无法完整地表征知识流序或技术轨迹演进的方向, 并且无论是本地主路径还是全局主路径都可能遗漏权重最高的链接。为了将更多关键节点和重要链接纳入主路径中, Verspagen<sup>[12]</sup>提出在全局主路径的基础上多次抽取权重较大的路径, 形成多条主路径, 即多元主路径法。Liu等<sup>[6]</sup>提出关键路径搜索策略, 来提升主路径对重要链接的涵盖程度。关键路径搜索不再从某一固定的节点开始进行前向或后向搜索, 而是从关键路径(通常是具有最高遍历权值的链接)开始, 同时向路径的两端进行

搜索, 直到抵达源点和汇点, 最后将提取到的路径进行拟合, 形成引文网络的主路径。

Yeo等<sup>[13]</sup>提出利用二阶马尔可夫链代替遍历权重进行路径搜索的聚合方法; 马瑞敏等<sup>[8]</sup>将Pathfinder算法与主路径分析所使用的搜索算法进行对比分析, 提出改进主路径搜索算法的可能性; Park等<sup>[14]</sup>提出利用遗传知识持久性测量(GKPM)方法来搜索主路径; 程洁琼等<sup>[15]</sup>提出基于边链接影响力流的主路径搜索算法。

### 2.3 主路径方法改进

尽管主路径分析方法得到了广泛的应用, 但是传统主路径分析仍存在如重点文献缺失、重要路径覆盖不全等局限和不足<sup>[6,13]</sup>, 在实际应用场景中受到许多限制。研究者在应用和研究的过程中提出了以下改进办法。①全局主路径搜索、关键路径主路径搜索、多元主路径搜索等多主路径搜索方法的提出从局部最优路径扩展到全局最优路径, 改进了本地主路径单一路径涵盖不全的问题。②利用二阶马尔可夫链、边链接影响力流等其他权重指标代替传统的遍历计数, 使主路径回溯能力更强, 覆盖的知识领域和技术范围更加全面, 动态演化过程更加明显。③许多研究者在网络构建方面进行了一些尝试, 基于关联规则或时序关系所构建的引文网络使节点之间不再局限于直接(间接)引用关系, 降低了因缺乏引用关系而遗漏关键文献的风险。④结合关联规则或边链接影响力流为初始链接赋权, 构造

了新的遍历权重指标,为因初始遍历权重无差别而产生的问题提出了解决方案。

### 3 应用实践

主路径分析方法的研究可以分为方法和应用两大类。方法类的研究从方法论层面出发,聚焦主路径方法与算法的改进;应用类的研究以实用为主,研究主路径方法在不同领域的实践与应用。

早期主路径分析方法以面向应用的研究为主,多用于梳理领域知识流的发展脉络。Hummon等<sup>[1]</sup>最早使用主路径分析方法揭示DNA理论发展史的主流知识脉络;Hummon等<sup>[16]</sup>随后将其应用于社交网络分析领域;Carley等<sup>[17]</sup>梳理了冲突解决领域关键知识流的发展;Moore等<sup>[18]</sup>研究了社会资本概念在公共卫生领域的演化发展,生成了公共卫生领域社会资本的概念谱系,确定了公共卫生中社会资本概念化的知识来源、有影响力的文献和进一步的发展方向;Calero-Medina等<sup>[19]</sup>将文献计量图与主路径分析结合起来,通过分析科学出版物研究知识的创造和转移过程;祝清松等<sup>[20]</sup>利用引文主路径分析方法进行了碳纳米管纤维领域演化路径识别。

随着专利计量的发展,主路径方法的应用范围扩展至专利引文分析领域,研究者开始利用这一方法绘制技术发展轨迹。Verspagen<sup>[12]</sup>研究了燃料电池领域的技术发展轨迹;Fontana等<sup>[21]</sup>重建了局域网领域的主要技术轨迹,对以太网数据传输标准领域技术变化的动力学机制进行了研究,并识别出局域网技术中的技术瓶颈;杨中楷等<sup>[22]</sup>识别和研究了太阳能光伏电池板领域的技术轨迹;Mina等<sup>[23]</sup>同时在冠状动脉病症研究和治疗领域的论文及专利中采用主路径方法,突出了知识和技术共同演化的模式,以寻求解决医学创新中出现的问题;Lu等<sup>[24]</sup>探索了纳米技术发展引发的广泛社会问题的主要研究前沿和层次结构;张丰等<sup>[25]</sup>研究了燃料电池领域的技术发展路径并简述了其对中国汽车领域发展的影响。

探测科学和技术的结构与演化也是主路径分析方法的重要应用领域。Martinelli<sup>[26]</sup>应用主路径方法识别通信交换机在7个不同时代的技术进步中发生的变化;Epicoco<sup>[27]</sup>研究了半导体小型化轨迹的长期演变;Ramlogan等<sup>[28]</sup>对冠状动脉成形术文献进行了深入分

析,采用主路径方法检测到了对抗严重冠状动脉疾病衰弱效应的新方法的引入,为领域内的变化提供了有力证据;Lucio-Arias等<sup>[7]</sup>力图超越科学知识的历史重构,运用主路径分析,凸显科学领域发展的结构性主干。

利用主路径分析方法追踪历史发展,阐明科学或技术领域的演变,提纲挈领地对某个领域的发展演变进行精练描述,逐渐成为文献综述的得力方法之一。Bhupatiraju等<sup>[29]</sup>利用主路径分析总结了企业家精神、创新研究和科学技术研究等领域的主要研究轨迹;Colicchia等<sup>[30]</sup>利用主路径分析,综述了从供应链风险管理背景下的动态角度研究知识创造、转移和发展过程的一系列重点文献;Harris等<sup>[31]</sup>利用主路径分析,对公共卫生服务和系统研究领域的文献进行了梳理和综述;Liu等<sup>[32]</sup>利用主路径分析,综述了为数据包络分析带来研究动力的活动以及该领域的4个研究前沿。

随着研究的不断深入,主路径分析方法逐步由应用研究转向改进探索。为弥补传统主路径分析方法的不足,解决其在实际应用中遇到的问题,研究者不断探索改进,提出增强使用效能的新思路。Lu等<sup>[24]</sup>综合运用聚类分析方法和文本挖掘技术,基于对一组纳米技术伦理发展历史文献的分析,划分了其研究领域和各领域的层级结构;Calero-Medina等<sup>[19]</sup>将文献计量制图引入路径分析,研究科学出版物创造和传递知识的过程,探索研究领域的“吸收能力”;Tu等<sup>[33]</sup>通过合并主路径中的相似主题论文,提出概念路径;Kim等<sup>[34]</sup>将主路径分析与技术关联分析相结合,以确定主路径上某些技术接合点产生的衍生路径;Lucio-Arias等<sup>[7]</sup>通过主路径方法和路径依赖方法对HisCite绘制的时间线图进行了改进,利用定量测量进一步扩展了算法史学;祝清松<sup>[35]</sup>提出利用语义方法抽取引文内容中的主题,计算引文内外的关联强度,用该指标替代遍历权重来构建加权引文网络,结合时序关系进行多主路径提取;陈亮等<sup>[36]</sup>将文本挖掘技术与引文网络分析方法相结合,提出通过语义相似度总和进行路径搜索,同时抽取多个技术主题路径的技术演化分析方法;万小萍等<sup>[37]</sup>基于主路径分析方法,利用关联分析改进了关键专利节点的选取和专利关系的描述,提出从大规模专利文献中快速提取技术演进路径的新方法;韩毅等<sup>[38]</sup>提出一种全新的基于连通性的领域演化结构识别方法,他们以主路径为种子文献提取与主路径关联的最大核心弱组分及凝聚子群,以此展示丰富的领域演化结构。

## 4 结论与对策

主路径方法是引文分析方法在网络环境下的拓展应用,在图论的基础上结合社会网络分析方法和复杂网络分析方法,其在科学技术发展关键路径提取中的应用具有广阔的发展前景。主路径分析方法的相关研究不仅有从算法描述、路径解释、要素构成等不同侧面对其概念进行界定的理论研究,也有基于“节点”和“链接”两个核心构成要素,从关键节点遴选、主路径识别两个方面对其算法进行改进的方法研究,还有将其用于不同领域知识演化脉络梳理及技术发展轨迹绘制的实践研究。而且,随着主路径分析方法在学术界的广泛使用,其理论、方法及应用还在不断充实和发展。

在日益丰富的研究中,传统主路径分析方法获得了极大的发展,这期间也涌现了许多行之有效的改进方法,但是当前主路径分析方法仍存在一些突出问题。①如何验证方法改进的有效性?目前主要采取实证领域专家验证和方差分析两种方法,但是这两种方法都不能强有力地说明或验证方法改进的合理性和有效性,提出一种切实可行的验证方法是改进主路径分析方法不可忽视的重要环节。②传统主路径分析方法强调节点之间的连通性,而对节点本身并未做深入研究,缺乏对关键节点进行分类和遴选,后续的改进方法也仅是尝试性地提出过2种节点识别和遴选方法,未见更深入的研究。③为增强网络的连通性,已有学者尝试用其他关系构建引文网络,但是主路径分析方法对网络结构十分敏感,其他关系或混杂关系所构建的引用网络是否合理,是否会影响主路径分析的准确性都需要更深入的研究。④主路径分析方法是基于图论和社会网络分析的一种实用方法,现阶段已有研究大多围绕边的遍历权重和搜索路径的算法选择,对构成主路径分析方法的网络拓扑次序研究十分有限。

应对当前主路径分析方法存在的问题,应该从节点、网络、链接3个方面展开研究。首先,需要借助复杂网络研究中关键节点遴选理论,对节点进行分类遴选,选择合适的指标,确定恰当的截断阈值,筛选出网络中的关键节点;其次,根据文献性质选取合适的关系构造引文网络,既要保证网络的连通性也要兼顾对关键节点的覆盖程度,同时还应增强对稀疏网络处理方法的研究;再次,基于关系类型或内容分析为初始链接赋予不同的权重,并且根据不同的初始值和遍历计数构造

新的遍历权重指标,改进遍历算法;最后,结合计算机技术对引文网络的算法核心进行优化对提高主路径分析的效果具有积极作用。

### 参考文献

- [1] HUMMON N P, DEREIAN P. Connectivity in a citation network: the development of DNA theory [J]. *Social Networks*, 1989, 11 (1): 39-63.
- [2] PRICE D J D. Networks of scientific papers [J]. *Science*, 1965, 149 (3683): 510-515.
- [3] GARFIELD E. Citation indexes for science: a new dimension in documentation through association of ideas [J]. *International Journal of Epidemiology*, 2006, 35 (5): 1123-1127.
- [4] SHIBATA N, KAJIKAWA Y, TAKEDA Y, et al. Comparative study on methods of detecting research fronts using different types of citation [J]. *Journal of the American Society for Information Science and Technology*, 2009, 60 (3): 571-580.
- [5] DE NOOY W, MRVAR A, BATAGELJ V. Exploratory Social Network Analysis with Pajek: Revised and expanded edition for updated software [M]. Cambridge: Cambridge University Press, 2018.
- [6] LIU J S, LU L Y Y. An integrated approach for main path analysis: development of the hirsch index as an example [J]. *Journal of the American Society for Information Science and Technology*, 2012, 63 (3): 528-542.
- [7] LUCIO-ARIAS D, LEYDESDORFF L. Main-path analysis and path-dependent transitions in HistCite (TM)-based historiograms [J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (12): 1948-1962.
- [8] 马瑞敏,张欣.基于Pathfinder算法的领域知识交流主路径发现研究[J].*情报学报*,2016,35(8):856-863.
- [9] 马瑞敏,杨雨华.基于节点重要性的领域主路径发现新探索[J].*情报杂志*,2018,37(3):71-78,93.
- [10] BATAGELJ V, MRVAR A. Pajek-program for large network analysis [J]. *Connections*, 1998, 21 (2): 47-57.
- [11] BATAGELJ V. Efficient Algorithms for Citation Network Analysis [EB/OL]. (2003-09-14) [2019-09-08]. <https://arxiv.org/pdf/cs/0309023.pdf>.
- [12] VERSPAGEN B. Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research [J].

- Advances in Complex Systems, 2007, 10 (1) : 93-115.
- [13] YEO W, KIM S, LEE J M, et al. Aggregative and stochastic model of main path identification: a case study on graphene [J]. Scientometrics, 2014, 98 (1) : 633-655.
- [14] PARK H, MAGEE C L. Tracing technological development trajectories: a genetic knowledge persistence-based main path approach [J]. Plos One, 2017, 12 (1) : 18.
- [15] 程洁琼, 万小萍, 刘向. 技术主路径分析: 基于边链接影响力流的路径搜索 [J]. 现代情报, 2019, 39 (5) : 24-29, 37.
- [16] HUMMON N P, CARLEY K. Social networks as normal science [J]. Social Networks, 1993, 15 (1) : 71-106.
- [17] CARLEY K M, HUMMON N P, HARTY M. Scientific influence: an analysis of the main path structure in the journal of conflict-resolution [J]. Knowledge-Creation Diffusion Utilization, 1993, 14 (4) : 417-447.
- [18] MOORE S, SHIELL A, HAWE P, et al. The privileging of communitarian ideas: Citation practices and the translation of social capital into public health research [J]. American Journal of Public Health, 2005, 95 (8) : 1330-1337.
- [19] CALERO-MEDINA C, NOYONS E C M. Combining mapping and citation network analysis for a better understanding of the scientific development: The case of the absorptive capacity field [J]. Journal of Informetrics, 2008, 2 (4) : 272-279.
- [20] 祝青松, 冷伏海. 基于引文主路径文献共被引的主题演化分析 [J]. 情报学报, 2014, 33 (5) : 498-506.
- [21] FONTANA R, NUVOLARI A, VERSPAGEN B. Mapping technological trajectories as patent citation networks: an application to data communication standards [J]. Economics of Innovation and New Technology, 2009, 18 (4) : 311-336.
- [22] 杨中楷, 刘佳. 基于专利引文网络的技术轨道识别研究——以太阳能电池板领域为例 [J]. 科学学研究, 2011, 29 (9) : 1311-1317.
- [23] MINA A, RAMLOGAN R, TAMPUBOLON G, et al. Mapping evolutionary trajectories: Applications to the growth and transformation of medical knowledge [J]. Research Policy, 2007, 36 (5) : 789-806.
- [24] LU L Y Y, LIN B J Y, LIU J S, et al. Ethics in nanotechnology: What's being done? What's missing? [J]. Journal of Business Ethics, 2012, 109 (4) : 583-598.
- [25] 张丰, 刘啦, 缪小明. 燃料电池研发的热点领域与关键技术: 基于专利数据的实证研究 [J]. 情报杂志, 2017, 36 (1) : 54-58.
- [26] MARTINELLI A. An emerging paradigm or just another trajectory? Understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry [J]. Research Policy, 2012, 41 (2) : 414-429.
- [27] EPICOCO M. Knowledge patterns and sources of leadership: Mapping the semiconductor miniaturization trajectory [J]. Research Policy, 2013, 42 (1) : 180-195.
- [28] RAMLOGAN R, MINA A, TAMPUBOLON G, et al. Networks of knowledge: the distributed nature of medical innovation [J]. Scientometrics, 2007, 70 (2) : 459-489.
- [29] BHUPATIRAJU S, NOMALER O, TRIULZI G, et al. Knowledge flows-Analyzing the core literature of innovation, entrepreneurship and science and technology studies [J]. Research Policy, 2012, 41 (7) : 1205-1218.
- [30] COLICCHIA C, STROZZI F. Supply chain risk management: a new methodology for a systematic literature review [J]. Supply Chain Management-an International Journal, 2012, 17 (4) : 403-418.
- [31] HARRIS J K, BEATTY K E, LECY J D, et al. Mapping the multidisciplinary field of public health services and systems research [J]. American Journal of Preventive Medicine, 2011, 41 (1) : 105-111.
- [32] LIU J S, LU L Y Y, LU W M. Research fronts in data envelopment analysis [J]. Omega-International Journal of Management Science, 2016, 58: 33-45.
- [33] TU Y N, HSU S L. Constructing conceptual trajectory maps to trace the development of research fields [J]. Journal of the Association for Information Science and Technology, 2016, 67 (8) : 2016-2031.
- [34] KIM J, SHIN J. Mapping extended technological trajectories: integration of main path, derivative paths, and technology junctures [J]. Scientometrics, 2018, 116 (3) : 1439-1459.
- [35] 祝青松. 语义增强的引文分析方法与应用实验研究 [D]. 北京: 中国科学院大学, 2014.
- [36] 陈亮, 杨冠灿, 张静, 等. 面向技术演化分析的多主路径方法研究 [J]. 图书情报工作, 2015, 59 (10) : 124-130, 115.
- [37] 万小萍, 刘向, 闫肖婷, 等. 基于关联分析的技术演进路径发现 [J]. 情报学报, 2018, 37 (11) : 1087-1094.
- [38] 韩毅, 童迎, 夏慧. 领域演化结构识别的主路径方法与高被引论文方法对比研究 [J]. 图书情报工作, 2013, 57 (3) : 11-16.

## 作者简介

刘懿，女，1995年生，硕士研究生，研究方向：信息计量与科学评价。

周丽英，女，1979年生，博士，副研究馆员，通信作者，研究方向：信息计量与科学评价，E-mail: zhouly@cau.edu.cn。

### Literature Review about Main Path Analysis Method

LIU Yi ZHOU LiYing

( China Agricultural University Library, Beijing 100193, China )

**Abstract:** The main path analysis method is a quantitative and visual method based on “connectivity” to reduce the complexity of knowledge network and extract the key path. It is practical significant in the scientific and technological development path detection. This paper systematically combs the relevant literatures for the study and application of the main path method at home and abroad from the three angles of concept analysis, method evolution and application practice. The aim is to clarify the development and research focus of the main path analysis method, to find the shortcomings and blank points of the existing research, to summarize the existing methods in the path detection problems, and to propose the main path detection method to improve the strategy. The focus of the research on the existing main path analysis method is mainly in the calculation of traversal weight and the selection of path search method, and the selection of key nodes, the detection of multiple paths and the density enhancement of sparse networks have a positive effect on improving the analysis effect of the main path, which is worth further studying in the future.

**Keywords:** Main Path Analysis; Citation Network; Method Improvement; Research Progress

(收稿日期: 2019-09-18)

## ■ 书 讯 ■

# 《汉语主题词表》

《汉语主题词表》自1980年问世以后，经1991年进行自然科学版修订，在我国图书情报界发挥了应有作用，曾经获得国家科学技术进步二等奖。为适应网络环境下知识组织与数据处理的需要，由中国科学技术信息研究所主持，并联合全国图书情报界相关机构，自2009年开始进行重新编制工作，拟分为工程技术卷、自然科学卷、生命科学卷、社会科学卷四大部分逐步完成。目前工程技术卷和自然科学卷已出版。

《汉语主题词表（工程技术卷）》共收录优选词19.6万条，非优选词16.4万条，等同率0.84，在体系结构、词汇术语、词间关系等方面进行了改进创新。《汉语主题词表（自然科学卷）》共收录专业术语12.4万条，包含数学、物理学、化学、天文学、测绘学、地球物理学、大气科学、地质学、海洋学、自然地理学等学科领域，收词系统、完整，语义关系丰富、严谨，每条词汇都有相应的学科分类号表现其专业属性，并与同义英文术语对应。同时，建立《汉语主题词表》网络服务系统，提供术语查询、文本主题分析、知识树辅助构建等服务。《汉语主题词表》可用于汉语文本分词、主题标引、语义关联、学科分类、知识导航和数据挖掘，是文本信息处理及检索系统开发人员不可或缺的工具。

《汉语主题词表（工程技术卷）》已于2014年由科学技术文献出版社出版，分为13个分册，总定价3 880元。

《汉语主题词表（自然科学卷）》已于2018年5月由科学技术文献出版社出版，分为5个分册，总定价1 247元。两卷均可分册购买。