

网络分析方法在引文分析中的整合研究

宋 歌

摘 要 随着社会网络分析在各学科领域的不断渗透,复杂网络理论的兴起和可视化技术的进步,引文分析面临着研究范式从个体主义到整体主义的转变,研究方法从统计分析到网络分析的融合。本文通过对引文网络形成过程和本质属性的探究,以及对网络分析方法理论前提和分析原理的考察,探索网络分析方法在引文分析中的整合途径。具体包括测度引文网络整体结构特性,探究网络结构的形成机制和演化过程,分析引文网络的细分结构、可视化引文分析过程与结果等方面的整合。图 1。表 2。参考文献 28。

关键词 网络分析方法 社会网络分析 复杂网络理论 引文分析 整合研究

分类号 G350

ABSTRACT With the penetrating of social network analysis into more and more disciplines, the rising of complex network theory and the development of visualization, citation analysis is undergoing tremendous changes, including research paradigm change from methodological individualism to holism, research method change from statistical analysis to network analysis. By detecting the forming mechanism and the essential attributes of citation network and by investigating the theoretical premises and principles of network analysis, the paper explores the ways to integrating the network analysis methods for citation analysis, such as measuring the whole structure properties of citation network, inquiring the formation mechanism and evolution process of network structure, analyzing the subdivision structure of citation network, visualizing the procedure and results of citation analysis and etc. 1 fig. 2 tabs. 28 refs.

KEY WORDS Network analysis methods. Social network analysis. Complex network theory. Citation analysis. Integration research.

CLASS NUMBER G350

1 引言

网络分析方法指主要以图论为数理基础的一种结构主义视角下的量化分析,包括社会网络分析(Social network analysis)和复杂网络理论(Complex network theory)。这两种理论与方法都源自社会学领域,然而因论证小世界理论而引爆物理学界复杂网络研究热潮的邓肯·瓦茨(Duncan J. Watts)却钟情于社会网络分析^[1],物理学界则另起炉灶,与社会网络研究分道扬镳。随着两种理论的日臻完善和对自然与社会现象的深入揭示,同根共源的两种理论势必在一些研究领域不期而遇。笔者发现,它们由于研究对象不同所形成的各具特色的研究方法,正好可以融合于对引文网络的研究中。持此观

点的还有大连理工大学的 WISE 实验室,他们认为:“把引文分析、复杂网络和社会网络三种理论与方法统一起来,将可能把科学知识图谱理论与方法提高到一个新的水平。”^[2]其中,对于社会网络分析,我国文献计量学界已经作出了积极响应。陈定权^[3]认为,我国同被引研究应该引入社会网络分析理论,与国外对比进行改进。岳洪江^[4]提到“引文分析只是为分析学科发展提供了一种基础数据,而社会网络分析可以利用这种引文数据揭示引文网络的结构特征,二者的结合可以成为一种有力的分析工具”。韩毅^[5]也认为“社会网络分析技术提供的结构划分方法为引文网络分析提供了在簇水平上研究网络结构的可能性”。

理念上的共识基于实践中的探索。从 1984 年加特雷尔^[6](Gattrell)以社会网络分析视角对

地理学期刊文献引文时序网络 and 同被引网络的阐释与分析,到1989年哈蒙^[7](Hummon)利用关键路径算法确定DNA理论的发展框架,再到2003年怀特^[8](White)利用网络分析方法对1998年利用多元统计方法得出的情报学作者同被引知识图谱的修正,以及近几年雷蒂斯托夫^[9-11](Leydesdorff)对于期刊及学科间引用关系的研究,显示出国外对网络分析方法在引文分析中的应用有一个循序渐进的研究过程,并且在2000年以后,随着可视化技术的突破形成了研究高潮。国内的相关研究始于2005年,虽然研究成果不多,但研究角度较广,主要涉及对某学科引文网络的小世界和无标度特性的验证,对某学科或主题的期刊、文献和核心作者的分析等^[12-16]。本文在前人研究的基础上,通过对引文网络形成过程和本质属性的探究,以及网络分析方法理论前提和分析原理的考察,对网络分析方法在引文分析中的整合进行初步探讨,以期对引文分析理论、方法与实践研究提供参考。

2 引文分析中网络分析方法的选择与应用

引文网络可以演绎成时序网络、耦合网络 and 同引网络三种类型。根据网络节点的不同还可以分为期刊引文网络、文献引文网络和作者引文网络,它们分别是以选定研究样本中的期刊、文献和作者为网络节点的引文网络,节点之间的关系可以是引用、耦合或同引。从图论的角度说,文献时序网络是有向非循环网络,作者同引网络是无向赋值网络。以下按引文分析目的的不同,归纳可选的网络分析方法并说明适用情况。

2.1 节点核心性测度

节点核心性测度作者、文献、期刊甚至学科在其引文网络中居于怎样的中心地位和具备何种影响力,对当前的学术评价具有重要意义。

点度中心度(degree centrality)是节点的局部中心性指数,在评价节点核心性时有其局限

性。经常被用于评价的入度点度中心度指标,其含义是其他节点对某一节点关系数量的总和,实际上就是该节点的总被引量。

接近中心度(closeness centrality)虽然是全局中心性指数,但其考虑的是节点在多大程度上不受其他节点的控制,也不适合作为测度节点核心性的指标。如果分析节点信息传递的独立性或者有效性,可采用接近中心度。

特征向量中心度(eigenvector centrality)是在网络总体结构基础上找到最居于核心的行动者,而不关注局部的模式结构,适宜作为核心性指标。但是,特征向量中心度也有缺陷,当数据为有向数据时,利用该指标可能引起误导,因为有些位置的节点可能不被选择。因此,如果分析不对称数据,需要先进行对称化处理。

能够弥补这一缺陷的是核心-边缘结构分析中对于节点核心度(coreness)的测量。核心-边缘结构模型建立的基础就是关系的不对称性,因此对于有向网络来说,采用核心度作为测量节点核心性的指标更加适宜。

2.2 节点中介性测度

除了测度节点影响力的核心性指标,从知识流通角度对节点进行测度的结构性指标,能为引文评价工作另辟蹊径。在引文网络中处于不同位置的节点,对知识传播的作用是不同的,除了那些具有重大影响力的节点外,还有一些节点对知识的快速流动起到了重要作用。而且,对于在不同研究领域的知识流通中起到枢纽作用的节点也应该在评价体系中有所体现,因为新的知识生长点往往存在于研究领域或学科领域交汇之处。因此节点中介性测度能够弥补引文评价功能单一的缺陷。

如果一个节点处于许多其他节点连通的最短路径上,就说明该节点具有较高的中介中心度(betweenness centrality)。中介中心度是一种“控制能力”指数,即节点作为信息集散地的能力,可以很好地测度节点中介性。

结构洞^[17](structure holes)与中介中心度一样,也是围绕“局部依赖性”这个概念建立起来的。当两个点以距离2而不是距离1相连的时

候,说明这两点之间存在一个结构洞。结构洞的存在使得连接两点的第三者扮演经纪人或者中间人的角色。拥有越低的结构洞约束系数的节点,越具有获取多样化知识的能力,是潜在的创新节点。

因此,中介中心度与结构洞约束系数都可以作为测度节点中介性的指标。不同之处在于,前者只能处理二值矩阵,后者可以处理赋值矩阵,适用范围更广。总之,点度中心度、中介中心度和接近中心度的局限性在于它们没有考虑到节点之间信息交换的规模,因而仅适用于对二值网络的测量。

此外,媒介角色系数(brokerage roles)也可以作为赋值网络中评价节点中介性的指标。那些从自己所在的子群连接到别的子群的节点,往往在整个网络中发挥重要作用,这类节点被称为“边界跨越者”。它们通常是具有创造性的节点,因为它们能够从不同的群体中获得多方面的信息,综合不同的知识或思路形成新的创意。与结构洞约束系数的不同之处在于,它用于分析节点在子群内和子群间所扮演的不同类型的中介角色。因此,如果是基于子群的分析,宜采用媒介角色系数,否则应采用结构洞约束系数。

2.3 引文网络嵌套结构测度

网络的嵌套结构与网络的等级结构不同,它揭示的不是权力层次结构,而是基于节点之间联系的紧密程度的层级关系,因此该方法主要用于勘测科学或学科结构而不宜用于评价。嵌套结构的测度主要利用k-核分析。

k-核即成分(最大关联子图)中的点都至少与该成分中k个其他点邻接。一个简单的成分就是一个“1k-核”,如果忽略所有度数为1的点,就得到一个“2k-核”,依此类推,进而考察剩余各个点之间的关联结构。因此,一个k-核便是在整个网络中的一个凝聚力相对较高的区域。利用k-核的一个明显好处在于对成分结构的研究可以运用度标准区分高低凝聚力的领域。这样,在进行引文分析时就能方便探

测出文献或作者的嵌套结构,过程中不仅可以分层展示,还可以根据需要进行不同k-核进行分析。

此外,n-派系、n-宗派、k-丛也可以用于嵌套结构分析。其中,k-丛与k-核的机理一样,是基于子群内部成员之间关系频次的(点的度数),而n-派系和n-宗派是基于子群成员之间可达性和直径基础上的。需要注意的是,n-派系和n-宗派在应用方面有很大局限性。首先n-派系的直径有可能大于n,其次n-派系可能是一个不关联图。一个n-派系中的两点可能通过一个长度不超过n的测地线连接在一起,这条测地线可能包含外部的点,并且这两个点之间不存在一条仅仅包含n-派系的成员的途径。n-宗派虽然限定了派系本身的直径不超过n,但是与n-派系一样,当n大于2的时候,缺乏解释力。因此,一般来说,可运用k-核分析来测度网络的嵌套结构。

2.4 引文网络等级结构测度

网络的等级结构分析一定是基于网络整体结构的,并且所有点和关系都应该被同时考虑,而不是把关注点仅限定在连接某些点的特定路径上。核心—边缘结构分析、位置分析和角色分析符合这个要求。

(1) 核心—边缘结构分析

核心—边缘模型(core-periphery model)的目的是对现实社会中表现出来的核心—边缘结构进行量化处理。在引文分析中可用该算法分析期刊引文网络和作者引文网络,判断期刊或作者在“核心—半边缘—边缘”的层次结构中所处的位置,或者计算出每个节点的“核心度”,再根据数据分布特征进行分层。其中各个节点在各层次中是可以流动的,但就整体而言,层次结构是相对稳定的。运用该模型时需要注意,只有单一核心网络才适用此方法。因此提出以下建议:可以先利用派系分析、聚类等方法考察现实网络是否存在多个核心。有的网络可能拥有较多核心,而利用核心—边缘模型分析的时候,可能找不到这些核心。

(2)位置分析与角色分析

社会网络分析中的“位置”(positions)和“角色”(roles)概念是比较抽象的,它们不同于中心度、派系等具有现实基础的概念,但也正因为如此,位置分析和角色分析能够对引用行为和网络结构进行一般化分析,得到具有推广意义的结论。更重要的是,我们可以在引文分析中运用此方法,根据类似的引用行为将各知识节点归类,并且解释是什么因素使这一类节点不同于其他类别的节点。

“网络位置”指的是一系列在网络关系或者互动中相似的节点。由于位置概念基于节点子集之间的关系相似性,因此,这个概念与凝聚子群的概念截然不同,处于相同位置的节点之间不必然有直接或者间接的关系。根据分析网络位置的结构对等性原理和对图书情报学期刊的实证研究结果^[18],在引文分析中,网络位置分析从两个维度对引文网络中的节点进行了区分。第一个维度是学科相似性,它把研究领域相近的文献、作者,学科领域相近的期刊分在了一组;第二个维度是节点权威性,将与其他节点引用关系相似的节点分为了一组。因此,如果对节点的学科分布情况比较熟悉,可运用位置分析同时察看各个节点学科分群和权威性两个方面的情况。

角色分析基于“规则对等性”原理,要比“结构对等性”更抽象一些,它不像结构对等性那样要求节点跟其他节点具有相同的关系,只是要求占有相同位置的节点以相同的方式相关联,即关注的是关系之间的联系。在引文分析中就是对位置分析中第二个维度的专门测度。因此,引文网络中的各种角色就可以根据一系列引用与被引用关系之间的联系来定义。角色分析可以不受节点学科属性的干扰而仅仅考察节点之间的权威结构。

2.5 引文网络分群测度

网络分群方法有两大类:一类是基于关系属性的凝聚子群分析(cohesive subgroup analysis),即群落结构测度;另一类是基于结构对等

性的位置分析。前者比较具体,包含众多算法,后者比较抽象,但是如果研究者对某领域的期刊、作者等知识节点比较熟悉的话,位置分析可以从整体网络的角度为学科分类提供依据,而且在有向赋值网络中比凝聚子群分析更具优势。位置分析前面已有介绍,以下具体谈凝聚子群分析。

凝聚子群分析对于期刊、作者、文献的分群,学科结构的研究都是一种精细的工具,它描述如何根据一定的模式把网络中的节点分派到各个子群之中。但是由于分析凝聚子群的算法很多,又是基于对关系的不同限定,彼此之间既有联系又有区别,因此有必要根据引文网络的特点和引文分析的目的厘清各种凝聚子群算法的适用情况,总结出具有可操作性的凝聚子群分析的一般方法和步骤。

“凝聚子群”在网络研究中并不是一个具有明确含义的概念。大体上说,“凝聚子群是这样一些节点的子集,子集中的节点之间具有相对较强的、直接的、紧密的、经常的或者积极的联系”^[19]。成分、派系、n-派系、n-宗派、派别、k-丛、k-核等都属于“凝聚子群”的范畴,都可以看成是“凝聚子群分析”的各个类型。它们分别是基于关系的互惠性、可达性、度数或子群内外关系来判定的。对于不同的算法在分析子群时有不同的处理方式,其中两个常规方法如下:

(1)对于派系严格的子群来说,很可能在网络中出现派系重叠的情况,即子群间存在一个或多个共同的节点,因此一般将一个包含重叠关系的派系群当作一个凝聚子群(如社会圈),而不是将每个单独的派系分别列为子群。

(2)k-核分析将形成一个嵌套(nested)结构,高一阶的k-核总是包含低一阶的k-核,所以一个节点可能同时属于几个k-核。不同成分中的k-核节点之间也不相连。因此,为了得到凝聚子群,研究者需要删除k-核中的低阶节点,直到网络分为几个相对紧凑的成分。

将Wouter de nooy等人^[20]制定的凝聚子群分析决策树稍作修订,绘制凝聚子群分析流程图(见图1)。

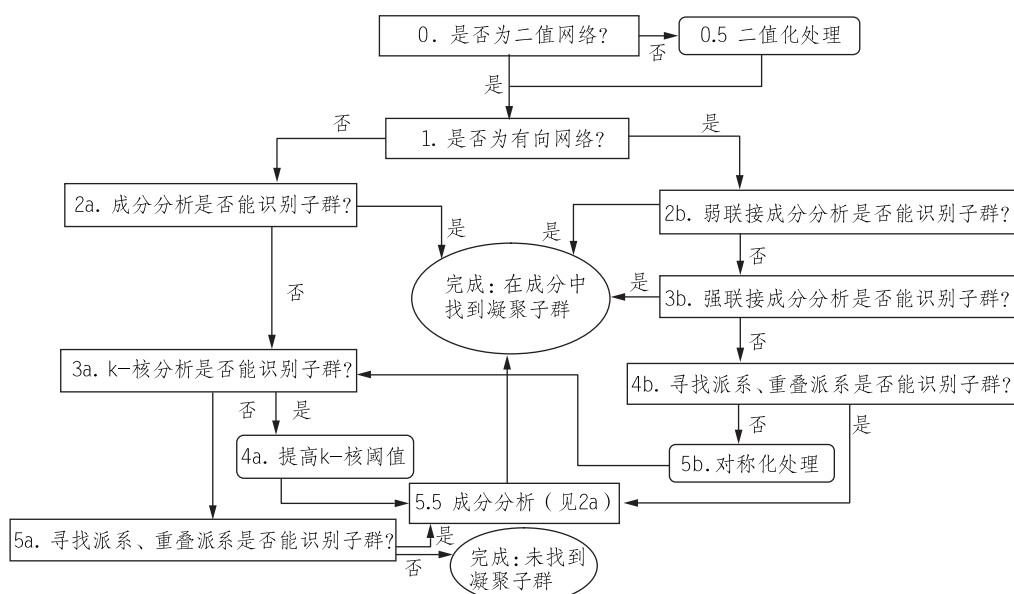


图1 凝聚子群分析流程

至于选择哪种技术来分析凝聚子群,主要依赖于网络的密度。在密集网络中,重叠派系可以很好地揭示网络的基干构架;而成分和k-核分析更适合拆分疏松的网络。建议在最初试探性的研究中,先进行成分分析再运用k-核分析,如果k-核过大,有必要进一步细分,再查找其中的派系、重叠派系或完备三方组。也有的研究者^[21]认为,在分析凝聚子群时,应该先分析定义比较严格的子群,然后分析界定比较松散的子群。例如,可以先分析“派系”,如果不存在派系,再进一步分析n-派系等。但本文认为,采用由松散到严格的分析路径能更好地把握引文网络的总体结构,同时不会遗漏网络的细部结构,并且容易判断采用的算法和设定的阈值是否合适。对应图1,将凝聚子群分析的一般过程总结如下:

(1)如果数据是二值的,直接看第一步;如果数据是赋值的,需要进行二值化处理。一般可采用平均同引强度、平均被引率等均值或网络密度作为阈值。以相似性数据(数字越大表示关系越近)为例,要确保大于某个指定值的数字重新编码为“1”,否则为“0”。也可以设置不同的阈值进行分析,以便检验数据的稳健性,找

出隐含在数据中的比较完备的图形结构。

(2)如果数据是无向的,进行a组分析;如果数据是有向的,进行b组分析。对于无向网络,分析其成分;对于有向网络,先分析其弱联接成分,如果在弱联接成分中没有找到子群,再分析其强联接成分。成分分析是凝聚子群分析的最简单形式,并且有时候成分可以为我们提供用来回答问题的充分信息。如果情况确实如此,分析就到此结束。

(3)如果成分分析的结果没有为我们提供充分的信息,则对无向网络进行k-核分析。如果能够获得比较清晰的嵌套结构,则提高k-核阈值,去除低阶k-核中的节点,将简化后网络的成分作为凝聚子群,结束分析流程。如果经过k-核分析还是没有找到子群,则在网络中寻找派系或重叠派系。如果寻找失败,则网络中不存在凝聚子群,结束分析流程。如果寻找成功,则进行成分分析,获得凝聚子群,结束分析流程。

(4)对于有向网络,如果强联接成分分析未能得到凝聚子群,则寻找派系或重叠派系。如果寻找成功,则相继步骤与5a相同。如果未找到则需要对称化网络,将所有的单向连接转换成双向或无向连接,之后利用得到的对称矩阵

进行 k-核分析,相继步骤与 3a 相同。

(5)如果在 4b 和 5a 中没有或仅找到几个派系,可以尝试作如下调整:如果派系的规模最小值为 4 或者更多,可以降低规模,但是不能降低到 3 以下;如果数据经过了二值化处理,则对于相似性数据来说,需要降低阈值,对于相异性数据来说,需要提高阈值。如果分析的结果是找到了太多的重叠派系,有必要的話可以把上述两个步骤颠倒过来进行,即增加派系的最低规模,最终就是改变阈值。

另外,在流程图中没有提及分派分析 (factions),它与派系分析不同的是可以找出指定数目的派别。只要知道可以分为几个区,针对网络中的任何一个成分都可以进行分派分析。比如在 2a、2b、3b、4b、5b 的分析中如果能推演出网络中大体存在多少个分组,就可以直接进行分派分析;又如在已经分析出的凝聚子群中,如果知道或者需要分为多少个小派别,也可以进行分派分析,找出子群内部的子-子群,这就需要结合具体的关系网络进行勘测。

2.6 引文网络演化测度

复杂性研究与社会网络研究的结合,为知识传播网络的演化问题开启了一扇窗。引文网络的生成和演化可以利用已有的网络拓扑基本模型来描述,而引文时序网络的演化路径可以通过关键路径分析得以呈现。

(1)BA 无标度模型与适应度模型

无标度是复杂网络的一个重要特性,文献

计量学的三大定律即洛特卡定律、齐普夫定律和布拉德福定律在数学形式上都呈幂律分布,具有标度变换下的不变性。国内外的大量研究也已经证明引文网络同样具有无标度特性,是无标度网络。与小世界网络是要建立一个具有明确拓扑特性的图不同,无标度网络的建模目标把重点放在掌握模型的动态特性上。其中 BA 无标度网络模型和适应度模型很适合用来对引文网络进行动态分析。

BA 模型考虑了网络的增长特性和优先连接特性,这两点符合引文网络的形成过程。引文网络的规模是在不断扩大的,每个月都会有大量新的科研论文发表,而且新发表的文章更倾向于引用一些已被广泛引用的重要文献。适应度模型与 BA 模型的区别在于,在适应度模型中的优先连接概率与节点的度和适应度之积成正比,而不是仅与节点的度成正比^[22]。这样,如果一个年轻的节点具有较高的适应度(与节点的内在性质相关),那么该节点就有可能在随后的网络演化过程中获取更多的连接。这一模型算法符合许多实际网络,在引文网络中表现为一些高质量的科研论文在较短时间内就可以获得大量引用的情况。

无论是 BA 无标度网络模型还是适应度模型,既能增减节点,又能重布连线,适合描述引文网络的生成和演化。表 1 总结了引文网络无标度特性的形成机制。正是节点的增长性和连接的择优性使得引文网络呈现无标度幂律分布。

表 1 引文网络无标度特性的形成机制

无标度特性	引文网络形成过程
节点数量 N 不确定	新的文献不断发表
节点连接概率不同	①经常被引用的文献会有更大的可能性被再次引用
	②一些高质量的文献在较短时间内就可能获得大量引用

(2)关键路径搜索算法

在生物范畴,遗传基因蕴含着生命体的延续与演进关系,而在科学研究领域,参考文献标识了学术传承与发展的轨迹。引文时序网络最详细地记载了科学研究的这种传承关系。加菲尔德的引文编年可视化系统 HistCite^[23]就是展

现引文时序网络的一种方式,但是该系统缺少更为细致的定量模型,“对于文献之间引用关系的密切程度无法反映”^[24],不能对关键文献及科研发展路径作出精确描述。而由哈蒙^[25]及其同事提出的关键路径搜索算法(main path analysis)弥补了这一缺陷,其独特之处正是在于关注

了引文时序网络的形成过程。

把引文网络看作一个传输知识信息的通道系统,其中如果有一篇高被引论文不但承继了很多以往的研究成果,并且还增添了实质性的新知识,那么它或多或少会降低先前一些论文的参考价值。因此,这篇文献就成为了承载着大量知识流的通道枢纽。可想而知,有很多引用路径都需要经过的文献节点一定比只有很少引用路径经过的文献节点更加重要。那些最重要的参考文献及其引用关系就构成了一条或多条关键路径,勾勒出科学发展的框架图。关键

路径搜索算法计算的就是每一个特定引用路径或文献在连通其他文献的引用路径时被需要的程度。这种测度指标被称作一篇文献或一条引用路径的“传输权重”。提取出拥有高传输权重的路径作为关键路径(main path)或者关键路径成分(main path component),它们即是传递文献知识流的“主干道”,可以通过分析其演进过程了解学术研究融合与分裂的模式,以及学术共同体的形成。

为了便于查看,表2总结了对应于不同的分析目的,网络分析方法的选择及适用情况。

表2 网络分析方法的选择与适用情况

研究目的	分析方法	适用情况
节点核心性	特征向量中心度	无向网络、双向网络
	核心度	有向网络
节点中介性	中介中心度	二值网络
	结构洞约束系数	赋值网络
	媒介角色系数	包含子群的赋值网络
网络嵌套结构	k-核分析	二值网络
网络等级结构	核心—边缘结构分析	单一核心网络
	角色分析	有向赋值网络
	位置分析	有向赋值网络
网络分群	凝聚子群分析	二值网络
	位置分析	有向赋值网络
网络整体特性	规模、密度/簇系数、凝聚力、平均距离、可达性、幂指数	各种网络
网络演化	以上各方法的时序分析	对应的各种网络
	关键路径分析	非循环网络
	BA 无标度网络模型、适应度模型	动态网络

3 引文分析中网络分析方法与统计分析方法的结合

从方法论角度来说,网络分析方法的整体主义(holism)方法论补充了一般统计分析的个体主义方法论(methodological individualism)。所谓个体主义方法论是把个体视为独立单位,按照个体的内在属性和规范特征来解释个体行

动的一种研究范式。而网络分析者从关系而非范畴的角度界定总体和样本,从相互联系而非孤立的视角描述和分析资料。体现在具体方法上,就是“网络分析者特别是整体网络分析者较多地运用行列式的数学方法,而较少地运用个体主义的统计技术”^[26]。国外学者对于网络分析在方法论上的优越性已经进行了很多讨论。例如 Emirbayer^[27]认为,网络理论把解释建立在关系模型之上,从关系视角进行的解释优于从

个体属性视角作出的解释。Wellman^[28]提出,网络分析方法直接针对社会结构的模式化的关系本质,从而可以补充甚至超越主流的统计方法。笔者认为,研究范式的转变是大势所趋,引文分析应该基于引文网络结构而不仅仅是引文条目,但是基于实体论的个体主义方法也是同样重要的,不能一概抛弃。比如网络分析方法和多元统计方法在分析结果上就可以互为印证和补充。有时候仅仅从网络关系的角度,或仅仅从个体属性的角度给出的解释都是不充分的,将两种不同视角的分析方法结合起来进行引文分析,能够更好地理解知识传播现象。具体来说,在引文分析中,网络分析方法与统计分析方法至少可在以下方面结合使用。

(1) 鉴于因子分析的精确性和网络分析的解釋力及可视化优势,可将因子分析与位置分析、角色分析等网络分析方法相结合,对网络分群或网络结构作出全面、精确的判定,之后再运用网络分析软件将结果可视化。

(2) 对于小样本的研究,或者经过缩减数据不会导致信息大量流失的样本,可以直接采用多元统计分析方法,因为以“地图”形式显示的知识节点更符合人们的认知习惯。对于大型引文网络,也可以先运用网络分析简化网络,然后再运用多元统计方法进行分析和显示。

(3) 在网络分析中,可先运用一般统计方法了解引文网络的基本静态特征,或运用多元统计分析检测研究对象是否符合某种网络分析方法的适用前提。例如,可以用聚类分析检验引文网络是否为单一核心以及初步了解核心、亚核心的分布情况。

4 结语

总的来说,网络分析方法在引文分析中的整合,可以利用复杂网络理论研究引文网络的整体结构特性,探讨网络结构的形成机制和演化过程;利用社会网络分析法分析引文网络的细分结构以及结构与个体间的相互作用;同时运用各种网络分析软件,如 Pajek、KNOT、UCI-NET 等展示基于节点间关系的引文网络结构,

加强引文分析过程和结果的可视化。

实际上,网络分析方法是图论结合代数方法、概率统计等发展起来的,它从来不排除其他数学方法的运用。尤其是通过近年来的迅速发展,网络分析方法已经涉及了一些较为复杂的数理分析和高级统计方法。本文论述中之所以形成网络分析方法与统计分析方法对立的局面,是基于网络理论和引文分析各自发展历史的,也是为了达到论述清晰的目的。笔者认为,在引文分析中要考虑的不是分析方法的替换,而是在方法论的层面挖掘网络分析的优势,将其更好地融入到当前的引文分析研究中。另外需要强调的是,网络分析不仅是对关系或结构加以分析的一套技术,它还是一种理论方法,是一种结构分析观点。我们不仅要了解网络分析的具体方法,还要理解其基本原理,从而更好地加以应用。

参考文献:

- [1] Watts D J, Dodds P S, Newman M E J. Identity and search in social networks[J]. Science, 2002, 296(5571):1302-1305.
- [2] 刘则渊,陈悦,侯海燕,等. 科学知识图谱:方法与应用[M]. 北京:人民出版社,2008:20-21.
- [3] 陆伟,李纲,谢阳群,等. 面向创新型国家建设的知识组织与服务——中国科协第 150 次青年科学家论坛会议综述[J]. 图书情报知识,2008(3):101-104.
- [4] 岳洪江. 管理科学知识扩散网络的结构研究[J]. 科学学研究,2008(4):779-786.
- [5] 韩毅,张克菊,金碧辉. 引文网络分析的方法整合研究进展[J]. 中国图书馆学报,2010(4):83-89.
- [6] Gattrell A C. Describing the structure of a research literature: Spatial diffusion modeling in geography[J]. Environment and Planning,1984B:11.
- [7] Hummon N P, Doreian P. Connectivity in a citation network: The development of DNA theory[J]. Social Networks,1989(11):39-63.
- [8] White H D. Pathfinder networks and author co-citation analysis: A re-mapping of paradigmatic information scientists[J]. Journal of the American Society for

- Information Science and Technology, 2003, 54 (5): 423-434.
- [9] Leydesdorff L. Clusters and maps of science journals based on bi-connected graphs in the journal citation reports [J]. Journal of Documentation, 2004, 60(4): 371-427.
- [10] Leydesdorff L. The delineation of nanoscience and nanotechnology in terms of journals and patents: A most recent update [J]. Scientometrics, 2008, 76(1): 159-167.
- [11] Leydesdorff L, Ismael Rafols. 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.
- [12] 金碧辉, Loet Leydesdorff, 孙海荣, 等. 中国科技期刊引文网络: 国际影响和国内影响分析 [J]. 中国科技期刊研究, 2005(2): 141-146.
- [13] 仲艳秋, 张晗, 冷荣新, 等. 利用社会网络分析法和聚类法研究心血管疾病知识结构的比较 [J]. 中华医学图书情报杂志, 2007(6): 77-80.
- [14] 徐媛媛, 朱庆华. 社会网络分析法在引文分析中的实证研究 [J]. 情报理论与实践, 2008(2): 184-188.
- [15] 岳洪江, 刘思峰. 管理科学期刊同被引网络结构分析 [J]. 情报学报, 2008(3): 400-406.
- [16] 陶乃航, 赵星, 彭晓东, 等. 中心度指标在期刊引文网络分析中的运用及改进 [J]. 图书情报工作, 2009(14): 143-146.
- [17] Burt R S. Structural holes: The social structure of competition [M]. Cambridge, MA: Harvard University Press, 1992.
- [18] 宋歌, 叶继元. 基于 SNA 的图书情报学期刊互引网络结构分析 [J]. 中国图书馆学报, 2009(3): 33-41.
- [19] Wasserman S, Faust K. Social network analysis: Methods and applications [M]. Cambridge: Cambridge University Press, 1994: 249.
- [20] De Nooy W, Mrvar A, Batagelj V. Exploratory network analysis with Pajek [M]. New York: Cambridge University Press, 2005: 78.
- [21] 刘军. 社会网络分析导论 [M]. 北京: 社会科学文献出版社, 2004: 172-173.
- [22] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用 [M]. 北京: 清华大学出版社, 2006: 34.
- [23] Garfield E, Pudovkin A I, Istomin V S. Algorithmic citation-linked historiography-mapping the literature of science [C]. Proceedings of the American Society for Information Science and Technology, Annual Meeting, November 2002, Philadelphia, 2002: 14-24.
- [24] 李运景. 基于引文分析可视化的知识图谱构建研究 [M]. 南京: 东南大学出版社, 2009: 43.
- [25] Hummon N P, Doreian P. Connectivity in a citation network: The development of DNA theory [J]. Social Networks, 1989(11): 39-63.
- [26] 张文宏. 社会网络分析的范式特征——兼论网络结构观与地位结构观的联系和区别 [J]. 江海学刊, 2007(5): 100-106.
- [27] Emirbayer M, Goodwin J. Network analysis, culture, and problem of agency [J]. American Journal of Sociology, 1994(99): 1411-1454.
- [28] Wellman B, Berkowitz S D. Social structures: A network approach [M]. Cambridge University Press, 1988: 19-61.

宋歌 东南大学情报科技研究所讲师。通讯地址: 南京市四牌楼 2 号。邮编: 210096。

(收稿日期: 2010-07-21)