

# 社会网络分析运用于学术人才挖掘的研究

段宇锋, 刘璇

(华东师范大学 商学院信息学系, 上海 200241)

**摘要:**以图书情报领域 1999-2008 年间的发文作者为主要研究对象建立起合著网络, 用传统的专家调查法确立 20 位学术权威, 发现这 20 位学术权威与合著网络中随机抽取的 1 000 位作者的四个指标值: 顶点值、点度中心度、中间中心度和接近中心度存在显著差异。以这 20 位学术权威的指标值为基准, 挖掘出符合条件的其他 17 位学术人才, 对这 17 位学术人才进行问卷调查判断他们的权威性。研究证明, 社会网络分析法运用于发现学术人才是有效而可行的。

**关键词:**社会网络分析法; 人才挖掘; 学术

**中图分类号:** C961      **文献标志码:** A      **文章编号:** 1008-5831(2011)05-0080-04

人才评价是科研管理与评价的重要理论和实践内容。目前, 人才评价主要采用文献统计和引文分析方法, 依据发表的成果数量、被引量以及 H 指数等指标。然而, 引用性质的差异、引文数据库的缺陷、“马太”效应等对评价结果的影响也一直为人们所诟病。因为研究合作是学者之间关系的直接反映, 也是体现其学术地位的重要形式。所以, 合作关系分析可以为人才评价提供客观依据。为此, 笔者的研究以科研合作为依据, 探索性地将社会网络分析方法运用于人才的评价和挖掘。

## 一、研究现状

社会网络分析萌芽于 20 世纪 30 年代, 形成于 20 世纪 60 年代。检索 EBSCO 数据库, 2000 年以前, 以社会网络分析为主题的文献仅 64 篇, 2008 年增长到 369 篇。社会网络分析已被运用到合著和引文网络研究、竞争情报、知识管理、网络信息资源评价以及图书馆资源配置等众多领域<sup>[1-10]</sup>。

社会网络分析主要集中在中心性、凝聚子群和核心-边缘结构三个方面, 本研究仅涉及中心性分析。Evelien Otte 等计算出信息科学领域的点度中心度指标用以判断网络的中心作者<sup>[11]</sup>; Zhu Weizhon 等特别探讨了 PageRank 和中心度指标之间的关系, 提出可以利用中间中心度作为参数对 PageRank 进行加权计算<sup>[12]</sup>; Jordi Palau 等介绍了如何利用点入度中心度和点出度中心度衡量人与团体的合作程度<sup>[13]</sup>。国内学者付允等指出中心性指标能够反映出作者在合作网络中的核心性及中心地位, 中心性越高的作者具有较高学术地位和较大影响力<sup>[14]</sup>; 梁潇研究发现信息加权的中心度可以用来寻找信息流动中的活跃分子以

收稿日期: 2011-07-22

**作者简介:** 段宇锋(1972-), 男, 山东定陶人, 华东师范大学商学院信息学系副教授, 博士, 主要从事网络信息资源管理研究; 刘璇(1983-), 女, 湖南邵阳人, 华东师范大学商学院信息学系硕士研究生, 主要从事信息资源管理研究。

及衡量各主体在整个网络中的地位<sup>[15]</sup>;邱均平将中心度指标作为判断图林博客核心的标准,认为中心度可以反映博客的重要程度,中心度高的博客可以看做博客圈中的核心且在圈子中有比较高的地位<sup>[9]</sup>;此外,孟微、庞景安将中心性的三个重要参数:度、介数和接近度(即点度中心度、中间中心度和接近中心度)应用于情报学科研合著网络的实证研究,发现作者三个指标值的排名存在一定的差异,同时具备三项高指标的作者很少<sup>[16]</sup>;苏娜根据社会网络分析的中心性指标对数字图书馆领域的研究趋势进行分析 and 预测,结果表明社会网络分析指标与其他方法相结合可以很好地识别学科领域的研究趋势<sup>[17]</sup>;魏瑞斌利用社会网络分析方法对关键词网络进行分析,发现领域的研究热点和研究内容的分布结构特征<sup>[18]</sup>;郑曦、孙建军构建了链接分析领域的作者合作网络,通过中心性分析可以找到在研究团队中起着至关重要信息枢纽作用的关键节点<sup>[19]</sup>。

基于已有的研究成果,本研究以中国图书情报学合著网络为样本,归纳学术人才的网络中心性特征,进行人才挖掘的探索。

## 二、样本和工具

本研究利用维普期刊论文数据库收集样本,检索 1999 - 2008 年 G25 和 G35 类号下的所有文献。在删除一些非论文文献和重复文献后,得到 75 252 位作者,161 787 篇文献。发文数量大于 2 且与他人有合作关系的作者有 13 213 位。利用这 13 213 位作者的合著关系建立起合著网络,本研究不区分作者的次序,因而建立的合著网络为无向网络。

社会网络分析软件有多种,常用的有 UCINET、Pajek、NetMiner、Multinet、Stocnet、Structure 等。本研究使用 www. analytictech. com 网站提供的 UCINET 软件作为分析工具。

## 三、学术权威的特征研究

### (一) 学术权威的特征值计算

样本发文量的统计表明,1999 - 2008 年间,发文数量在 50 篇以上的作者有 105 位。运用专家调查法,确定了 20 位图书情报学领域的学术权威,即马费成、胡昌平、邱均平、吴慰慈、叶继元、王余光、陈传夫、张晓林、程焕文、吴建中、李纲、朱庆华、梁战平、武夷山、靖继鹏、叶鹰、王知津、范并思、柯平、于良芝。

社会网络分析理论指出,如果一个行动者与很多其他行动者有直接的关联,该行动者就居于中心地位,拥有较大的权力。衡量指标为点度中心度;如果一个行动者处于许多交往网络路径上,可以认为此人居于重要位置,因为他具有控制其他两个个体之间交

往的能力。衡量指标为中间中心度;一个点越是与其他点接近,该点在传递信息方面就更加容易,该点也越不依赖他者,因而居于网络的中心。衡量指标为接近中心度。利用 UCINET 计算出上述 20 位权威的点度中心度、中间中心度和接近中心度。为了排除奇异值对整体取值的影响,本研究采用中值。

发文量是一个非常重要的判断学术地位的指标,在合著网络中以顶点值(vertex value)表示。20 位权威的顶点值见表 1。

表 1 20 位学术权威的特征值

学术权威	顶点值	点度中心度	中间中心度	接近中心度
马费成	41	16	374 100	33 651 664
胡昌平	45	23	638 772	33 648 644
邱均平	154	55	2 171 129	33 645 640
李 纲	17	8	71 731	33 657 276
朱庆华	37	15	399 648	33 652 820
梁战平	42	15	306 637	33 652 384
武夷山	29	14	142 150	33 658 828
靖继鹏	17	14	516 378	33 648 232
叶 鹰	34	9	109 236	33 661 852
王知津	116	38	1 145 996	33 649 220
陈传夫	39	24	804 084	33 646 932
叶继元	37	14	309 070	33 652 164
范并思	66	7	89 339	33 652 700
柯 平	77	27	859 648	33 649 260
于良芝	29	14	168 132	33 665 344
张晓林	100	47	3 648 396	33 640 100
程焕文	51	8	158 184	33 668 772
吴建中	34	6	120 999	33 659 308
吴慰慈	50	13	309 304	33 548 624
王余光	25	7	54 569	33 660 796
平均值	52	18.7	619 875.1	33 648 528
中 值	38	14	307 853.5	33 651 914

### (二) 对照组的特征值计算

在整个合著网络中随机抽取 1 000 位作者,统计其顶点值,并用 UCINET 软件进行点度中心度、中间中心度、接近中心度分析,结果如表 2 所示。

表 2 对照组的特征值

作 者	顶点值	点度中心度	中间中心度	接近中心度
平均值	7.489	2.972	21 383.8	59 736 401.94

点度中心度即直接与该点相连的点的个数,也就是与该作者有过合著关系的作者的个数。一般来说,点度中心度都小于顶点值,直观上理解,就是作者与他人合著发文的数量总是小于该作者的总发文量。点的中间中心度测量该点在多大程度上控制他人之间的交往。中间点度中心度为该点的所有点的

中间度的总和,中间中心度的值越大,说明对资源的控制能力越强,观察有一部分作者的中间中心度为0,表示该作者不能控制其他行动者,处于网络的边缘。接近中心度的计算有几种方法,我们选择的方法是 Geodesic paths only (Freeman Closeness),即只计算测地线(最短路径)的距离,得到的值为该点与所有其他点的最短距离之和。离中心点距离越远的行动者也在信息资源、权力、声望以及影响方面越

弱。接近中心度的值越大,说明该点越不是网络的核心点。

### (三)学术权威与对照组特征值的差异分析

运用 SPSS 软件进行差异分析。首先进行非参数假设检验,将 20 位学术权威作为样本组 A,对照组作为样本组 B,进行 Kolmogorov - Smirnov Test,检验两个样本所属总体的分布是否存在显著差异。结果表明,两组样本存在着显著差异(表 3)。

表 3 Kolmogorov - Smirnov Test

		顶点值	点度中心度	中间中心度	接近中心度
Most Extreme Differences	Absolute	0.933	0.875	0.877	0.701
	Positive	0.933	0.875	0.877	0.000
	Negative	0.000	0.000	0.000	-0.701
Kolmogorov-smirnov Z	4.131	3.875	3.883	3.104	
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	

## 四、学术人才挖掘

前一部分的研究表明,学术权威的顶点值、点度中心度、中间中心度和接近中心度与一般研究者具有显著差异。基于此,我们以顶点值 $\geq 38$ 、点度中心度 $\geq 14$ 、中间中心度 $\geq 307$ 、接近中心度 $\leq 33$ 为标准,对整个合著网络进行筛选,找到 17 个符合条件的学者,如表 4 所示。

表 4 17 位学者的特征值

学术人才	顶点值	点度中心度	中间中心度	接近中心度
侯汉清	96	41	1 562 760.750	33 647 252
初景利	79	32	2 793 939.250	33 640 240
马海群	78	21	705 956.938	33 649 116
李书宁	76	16	709 560.313	33 645 480
刘兹恒	56	19	1 165 278.875	33 646 976
盛小平	54	23	602 413.563	33 645 716
文庭孝	53	16	446 698.188	33 650 512
王子舟	53	45	1 372 393.375	33 645 488
肖希明	52	18	748 942.875	33 647 964
苏新宁	48	22	453 662.969	33 651 844
詹德优	46	20	779 653.875	33 651 228
王惠翔	45	17	456 319.438	33 648 912
焦玉英	43	19	420 144.844	33 654 756
张智雄	40	26	387 716.438	33 645 572
孙坦	39	29	1 038 466.313	33 543 516
赖茂生	38	20	582 402.688	33 649 036
毕强	38	19	383 697.688	33 650 312
平均值	55	23	859 412.3	33 641 476

为了检验挖掘结果的质量,本研究进行了小规模问卷调查,调查对象为图书情报领域的 10 位学者。调查问卷采用双盲法设计,将 20 位学术权威与

挖掘结果的顺序打乱,让参与者分别对这 37 位学者的权威度打分。分值为 1 - 10 之间的整数,对不熟悉的学者标记“一”。调查共发出 10 份问卷,回收问卷 9 份,有效问卷 8 份。

调查对象对挖掘出的 17 位学者基本上都比较熟悉,认为他们在图书情报学领域具有相当大的影响力。对调查问卷进行统计,20 位权威的平均分为 8.9,挖掘结果的平均分为 7.8。通过中国期刊网对挖掘结果的 H 指数进行统计分析,其平均 H 指数为 13.8,表明 17 位学者的论文数量和质量都比较高。

综上所述,本研究认为将社会网络分析方法运用于学术人才的评价具有一定的可行性。它克服了引文分析存在的问题,两者互为补充,有望显著提高学术评价的准确性。当然,本研究还只是社会网络分析在学术人才评价领域的初步探索,凝聚子群分析、核心 - 边缘结构分析、结构对等性分析等方面的引入将进一步推动这一领域的发展和完善。

### 参考文献:

- [1] LIU XIAOMING, BOLLEN J, NELSON M L, et al. Co-authorship networks in the digital library research community [J]. Information Processing & Management; 2005(41): 1462 - 1480.
- [2] Newman M E J. Scientific collaboration networks. I. Network construction and fundamental results [J]. Physical Review E, 2001, 64:016131
- [3] KREUZMAN H. A co-citation analysis of representative authors in philosophy: Examining the relationship between epistemologists and philosophers of science [J], Scientometrics, 2001(5): 525 - 539.
- [4] HOWARD D, WELLMAN W B, NAZER N. Does citation re-

- flect social structure longitudinal evidence from the “Globe-net” interdisciplinary research group[J]. Journal of the American Society for Information Science and Technology, 2004,55(2):111 - 126.
- [5] 包昌火,谢新洲,申宁. 人际网络分析[J]. 情报学报, 2003,22(3):365 - 374.
- [6] 李亮,朱庆华. 社会网络分析方法在合著分析中的实证研究[J]. 情报科学,2008,26(4):549 - 555.
- [7] 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. 2008,31(2):184 - 188.
- [8] 王晓光,社会网络范式下的知识管理研究述评[J]. 图书情报知识,2008(4):87 - 91.
- [9] 邱均平,马瑞敏. 图林博客的社会网络分析[J]. 图书情报工作. 2008,52(11):6 - 9.
- [10] 高凡,徐引麓. 图书馆联盟的社会网络资源配置[J]. 中国图书馆学报,2006(3):14 - 16.
- [11] OTTE E, ROUSSEAU R, Social network analysis; a powerful strategy, also for the information sciences[J] Journal of Information Science, 2002,28:441 - 453.
- [12] ZHU WEIZHONG, CHEN CHAOMEI, ALLEN R B. Analyzing the propagation of influence and concept evolution in enterprise social networks through centrality and latent semantic analysis[M] Washio T, et al. PAKDD 2008, LNAI 5012, 2008:1090 - 1098.
- [13] PALAU J, MONTANER M, BEATRIZ L' OPEZ, JOSEPLIU' S DE LA ROSA. collaboration analysis in recommender systems using social networks[M] KLUSCH M, et al. CIA 2004, LNCS 3191, 2004.
- [14] 付允,牛文元,汪云林,李丁. 科学学领域作者合作网络分析——以《科研管理》(2004 - 2008)为例[J]. 科研管理, 2009,30(3):41 - 46.
- [15] 梁潇. 三螺旋创新模式中信息流主体研究——以信息论加权的网络分析法[J]. 情报科学,2008,26(1):115 - 119.
- [16] 孟微,庞景安. 我国情报学科研合著网络研究及其特征参数分析[J]. 情报理论与实践,2009,32(8):12 - 15.
- [17] 苏娜,张志强. 社会网络分析在学科研究趋势分析中的实证研究——以数字图书馆领域为例[J]. 情报理论与实践,2009,32(9):79 - 83.
- [18] 魏瑞斌. 社会网络分析在关键词网络分析中的实证研究[J]. 情报杂志,2009,28(9):46 - 49.
- [19] 郑曦,孙建军. 链接分析领域的作者合作网络及其分析[J]. 图书情报工作. 2009,53(4):29 - 32,105.

## An Exploratory Study of Social Network Analysis Method Applied to Establish the Academic Talents

DUAN Yu-feng, LIU Xuan

(Information Department of Business School, East China Normal University, Shanghai 200241, P. R. China)

**Abstract:** This article aims at studying co-author network of library and information science according to all the authors in the field of library and information science during 1999 - 2008. Using traditional expert survey method to identify 20 authorities, it is found that the 20 authorities have significant difference in the four indices: vertex value, point centrality, betweenness centrality, closeness centrality, compared to the random 1000 authors'. Take the 20 authorities' index value as a benchmark to seek the other 17 eligible authors, we conducted a questionnaire survey to weight their authority. It is concluded that social network analysis method is feasible and effective to determine academic talents.

**Key words:** social network; analysis method; talents mining

(责任编辑 彭建国)