

· 学科进展与展望 ·

中国科学计量学的发展:论文引文分析及中印比较*

梁立明

(河南师范大学科技与社会研究所, 新乡 453007)

[摘要] 通过对中国 *Scientometrics* 论文和引文做计量分析,用 WordSmith Tools 软件做主题词分析,并与科学计量学论文高产国印度做比较研究,我们对中国科学计量学的发展有了概略的了解:(1) 经过 1984—2000 年的漫长起步期,中国科学计量学在新千年迅速崛起,已经成为科学计量学论文高产国。(2) 在印度,国家研究机构是科学计量学的主要研究力量,而中国的主要研究力量是大学。(3) 和印度相比,中国科学计量学论文的高产作者不够高产。(4) 引文测度显示,新千年以来中国的科学计量学研究比印度的研究更受国际同行关注。(5) 引文研究和期刊研究占中国论文很大份额,研究中国问题的中国论文的比例远大于研究印度问题的印度论文的比例。(6) 最近 5 年中国论文与世界论文高频主题词基本一致,说明中国学者是在国际科学计量学热点领域从事研究的。论文最后给出发展中国科学计量学的几点建议及一个公开讨论的问题。

[关键词] 科学计量学,中国,印度,引文分析,主题词分析

1 引言

科学计量学 (Scientometrics) 是对科学自身进行定量研究的学科。20 世纪 60 年代初, D. Price 等人采用定量研究方法研究科学增长和科学论文的网络^[1,2] 等问题。与此同期, E. Garfield 创建了大型引文数据库 Science Citation Index, 倡导引文分析方法, 并与 I. Sher 一起构建了期刊影响因子指标^[3]。Price 和 Garfield 等人的工作开启了科学计量学研究范式的先河, 奠定了科学计量学的数据根基。此后 40 余年, 科学计量学逐步发展为一个成熟的学科:

(1) 以科学计量学命名的专业期刊 *Scientometrics* 1979 年创刊, 到 2009 年底出版了第 81 卷 (说明: 一年不止一卷);

(2) 2008 年以信息计量学 (该学科与科学计量学有很大的交集) 命名的专业期刊 *Journal of Informetrics* 创刊, 至此, 有包括 *Scientometrics*, *Journal of the American Society for Information Science and Technology* (JASIST) 在内的十几种国际期刊发表科学计量学论文;

(3) 两年一届的国际科学计量学系列会议始于 1987 年, 2009 年在里约热内卢召开了第 12 届大会;

(4) 科学计量学领域最高学术成就奖“普赖斯奖”1984 年首次颁发, 至今已颁发 15 次, Garfield 等 24 位科学家获此殊荣;

(5) 成立于 1993 年的国际科学计量学与信息计量学学会 (ISSI, International Society for Scientometrics and Informetrics) 凝聚各分支学科的研究力量, 系统而活跃地开展学术活动;

(6) 各国科学计量学工作者不仅用自己的研究成果积极为各自国家的科技评价、科技决策和科技管理效力, 而且共同为区域和世界科技发展服务。

多年来科学计量学与其姊妹学科文献计量学、信息计量学、网络计量学等相互依存共生共荣, 形成了生机勃勃的学科群。人们普遍认为, 现在是国际科学计量学发展的最好时期。

中国科学计量学研究亮相世界舞台不算太晚, 1984 年赵红州在 *Scientometrics* 上发表了中国的第一篇论文^[4]。其后 2 年, 赵红州、蒋国华又联名在 *Scientometrics* 上发表了中国的第 2 和第 3 篇论文。然而, 最初 10 年中国科学计量学研究的整体实力较弱,

* 国家自然科学基金资助项目。

本文于 2010 年 2 月 3 日收到。

1984—1993年中国总共发表了9篇 *Scientometrics* 论文,只占该刊10年发文总量的1.36%。接下来的10年中国学者发表了25篇 *Scientometrics* 论文,占该刊10年发文总量的2.57%。中国学者的科学计量学论文也散见于 *JASIST* 等期刊。而在国内,自20世纪80年代后期起,科学计量学研究已经开始为国家科技发展服务。最具规模的成果是中国科学技术信息研究所建构的中国科技论文与引文数据库(CSTPCD)及该所每年发布的中国科技论文统计与分析研究报告。中国科学院文献情报中心的CSCD数据库、清华大学的CNKI数据库、南京大学的中文社会科学引文索引(CSSCI)及其相关报告是另几项有代表性的工作。多年来,来自科学学、科技情报学、科技管理学等学科的专家学者在国内学术期刊上发表了一系列科学计量学理论研究与应用研究论文。科学技术哲学、高等教育管理等学科领域的学者也对科学计量学理论与方法表现出极大的兴趣。

从1984年中国首篇 *Scientometrics* 论文发表至今1/4世纪过去了,我们感知到中国科学计量学研究队伍的显著壮大和成果的日益增多,然而,中国的科学计量学研究在世界舞台上究竟表现怎样,和国际同行相比究竟处于何等地位,这却是仅凭经验和直观判断无法回答的问题。本文尝试用科学计量学方法给出初步评价。科学计量学是以基础研究和应用基础研究为主的学科,因此,相关评价主要基于学术论文和引文的分析。*Scientometrics* 是科学计量学领域最权威的期刊,也是发表科学计量学论文最集中且数量最多的期刊,该刊是公认的研究科学计量学领域发展状况的最佳样本^[5-7],我们也选定该刊论文和引文为样本。同时,我们选择2001年以前 *Scientometrics* 论文排名世界第5的印度为比较对象^[8],从论文、引文和研究选题3个方面作计量分析,解读中国科学计量学的发展。Thomson Reuters公司的Web of Science为本项研究的论文和引文数据来源,WordSmith Tools软件用于做主题词词频分析。

2 论文分析

2.1 论文的时间分布

1978—2009年 *Scientometrics* 共发表各类文献2653篇,含研究论文和综述论文2315篇,其余为书评、短讯等。我们将这些文献统称为 *Scientometrics* 论文。至少有1位作者地址栏含“Peoples R China”字段的论文被称为中国 *Scientometrics* 论文,简称中

国论文。至少有1位作者地址栏含“India”字段的论文则被称为印度论文。检索结果,中国论文115篇,占全部 *Scientometrics* 论文的4.3%;印度论文161篇,占6.0%。如果用 *Scientometrics* 论文数量作为一个显示度指标,印度在国际科学计量学领域的显示度要大于中国。图1展示出中印两国论文的时间分布。

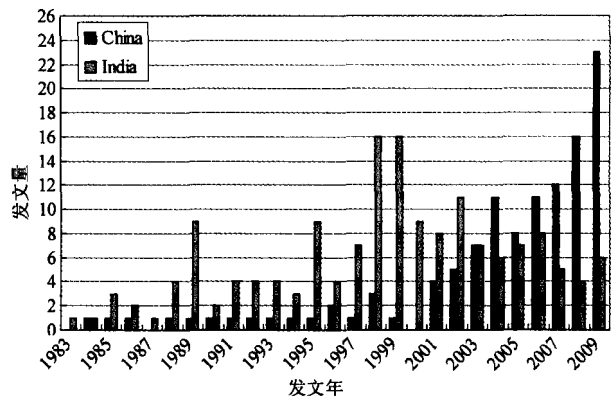


图1 中国和印度 *Scientometrics* 论文的时间分布

从图1可以看出,尽管印度论文总量远多于中国,但印度发表首篇 *Scientometrics* 论文的时间却只比中国早一年。1983年印度发表首篇论文后,其论文进入缓慢增长时期,第1个10年(1983—1992)年均发表论文3.1篇。第2个10年(1993—2002)是印度论文增长的黄金时期,10年发表了87篇论文。其中,1998和1999连续2年产出论文16篇,分别占 *Scientometrics* 全年发文量的18.0%和11.4%,标志着印度成为论文产出大国。2003年以后,印度维持一个较低的发文水平,年均发文6.1篇。

中国 *Scientometrics* 论文的时间分布和印度截然不同。1984—2000年的17年间中国总共发表18篇论文。也就是说,当印度科学计量学已在世界舞台拥有相当话语权的时候,中国的声音还十分微弱。图2是A. Schubert根据各国1978年(第1卷)至2001年初(第50卷)发表 *Scientometrics* 论文的数量绘制而成的,PRC和IND分别代表中国和印度^[8]。在Schubert的研究结果中,印度发文量排在美国、荷兰、匈牙利、德国之后,居世界第5位。中国排在第18位。

进入新千年后,中国论文数量迅速增长。2001—2009年中国在 *Scientometrics* 上发表了97篇论文,年均10.8篇。2009年更是达到创纪录的23篇,占是年该刊论文总量的12.0%,成为论文高产国家。

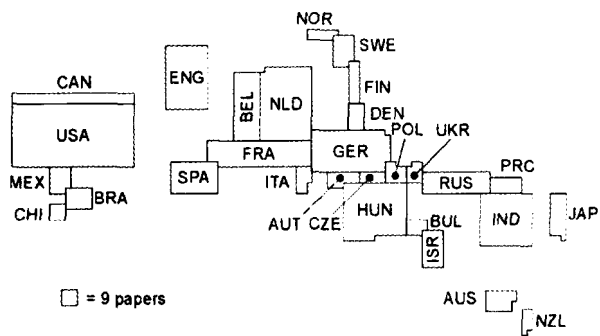


图 2 不同国家在 *Scientometrics* 上发表论文比例图(1978—2001 年初)

此图来源于文献[8]。原图标题：Proportional map of countries according to their publication activity in the journal *Scientometrics*。

需要提及中印两国科学计量学力量对比发生转折时期的重要事件。1991 年在印度召开了第 3 届国际科学计量学大会，随后，印度迎来其科学计量学发展的黄金时期。无独有偶，伴随中国筹办和承办 2003 年第 9 届国际科学计量学与信息计量学大会 (ISSI 大会)，中国科学计量学的春天也如期而至。国际会议的筹备和召开使科学计量学在会议承办国得到较大规模的普及和推广。依照会议主题撰写论文、国际同行评议专家对论文给出有价值的意见和建议、“近水楼台”地参加高水平的国际会议……这都有助于提升承办国科学计量学研究者的学术水平，吸引一批新成员投身该领域的研究工作。2003 年第 9 届 ISSI 大会在北京召开的次年，中国年发表 *Scientometrics* 论文数量首次超过印度。这种优势继而迅速扩大，到 2008 年中国发文量已达印度的 4 倍，2009 年为印度的 3.8 倍。

有趣的是，中国科学计量学发文量的飞速增长直接提升了中国在该领域的话语权。2003 年之前历届 ISSI 大会的亚洲主席都由印度学者担当。2005 年斯德哥尔摩第 10 届 ISSI 大会中国学者首次应邀担任中国-东亚区主席，印度学者则担当印度-南亚区主席。这种两分模式一直延续到第 11 届马德里和第 12 届里约热内卢 ISSI 大会。2005 年之后，有两名中国学者先后被聘为 *Scientometrics* 编委。2007 年两名中国学者同时被聘为 *Journal of Informetrics* 编委。

2.2 论文的机构分布

对中印两国 *Scientometrics* 论文作者的机构进行分析，我们发现两国发表论文的主力是截然不同的。在印度，研究机构是科学计量学研究的主力。印度的国家科学技术与发展研究所 (National Insti-

tute of Science, Technology and Development Studies) 发表了 80 篇论文，支撑印度科学计量学的半壁江山。发表论文最多的大学是 Karnatak University，仅发表 14 篇论文。在中国，大学是科学计量学的主要研究力量，发文量居前的 10 个机构中只有中国科学院和中国科学技术信息研究所两属于研究机构：Henan Normal Univ (15 篇)、Chinese Acad Sci (12 篇)、Beijing Univ Aeronaut & Astronaut (10 篇)、Harbin Inst Technol (10 篇)、Peking Univ (9 篇)、Inst Sci & Tech Informat China (8 篇)、Dalian Univ Technol (6 篇)、Tsing Hua Univ (6 篇)、Wuhan Univ (6 篇)、Fudan Univ (5 篇)。在世界机构发文量排名中 Henan Normal Univ 和 Chinese Acad Sci 分列第 30 位和第 44 位。

2.3 高产作者

高产作者在提升国家论文产出总量方面起着举足轻重的作用。中印两国 *Scientometrics* 论文高产作者分布存在很大差异。印度的高产作者数量多，发文量大，发文时间长，中国作者望尘莫及。表 1 列出两国的高产作者。中国发表 4 篇以上论文的作者共计 13 位，印度发文 7 篇以上的作者 12 人。居印度发文量首位的 Gupta BM 在 1990—2009 的 20 年间发文 27 篇，发文量居世界个人排名第 10 位。紧随其后的是 Garg KC，在 1985—2006 年的 22 年中发文 23 篇，世界排名第 12 位。中国高产作者前两位(梁立明、官建成)的发文量分列第 31 和 48 位。和印度相比，中国科学计量学的高产作者不够高产。

表 1 中国和印度 *Scientometrics* 论文高产作者

中国		印度	
姓名	发文量	姓名	发文量
Liang LM(Liang L, Liming L)	13	Gupta BM(Gupta B)	27
Guan JC(Guan J)	10	Garg KC	23
Yu G	9	Nagpaul PS	12
Ho YS	6	Kumar S	10
Li YJ	5	Arunachalam S	9
Zhou P	5	Bhattacharya S	9
He TW	4	Karisiddappa CR	9
Jin BH(Jin B)	4	Rao IKR	9
Jiang GH	4	Sengupta IN	8
Ma N	4	Kalyane VL	7
Qiu JP	4	Kumar V	7
Yu DR	4	Kundra R	7
Zhao H	4		

3 引文分析

我们从引文分析的角度审视和比较中印两国

Scientometrics 论文的学术影响。引文分析所用数据检索自 Web of Science, 检索时间为 2010 年 1 月 1 日至 3 日。

3.1 一般引文指标

中国的 115 篇 *Scientometrics* 论文有 76 篇被引用, 被引率 66.1%。共被引用 465 次, 篇均被引 4.04 次。465 次引用被包含在 349 篇施引论文中, 其中, 不含作者自引的施引论文(即他引论文)为 305 篇, 占 87.4%。不排除自引, 中国论文的 h 指数为 11, 即有 11 篇论文被引 11 次或 11 次以上, 其余论文被引频次均低于 11。

印度 161 篇论文有 132 篇被引用, 被引率 82.0%。总被引频次 811, 篇均被引频次 5.04。811 次引用被包含在 556 篇施引论文中, 其中, 他引论文 482 篇, 占 86.7%。不排除自引, 印度论文的 h 指数为 13。

比较中印两国各项引文指标, 可做出如下评价:

(1) 印度论文被引率和篇均被引频次均高于中国。

此处, 在做出评价之前需先界定一个概念——论文年龄。2009 年的论文年龄 = 2009 - 论文发表年份 + 1。例如, 在 2009 年, 2000 年发表的论文年龄为 2009 - 2000 + 1 = 10 年。可以计算出印度 161 篇论文的平均年龄为 11.6 年, 中国 115 篇论文的平均年龄为 5.9 年。印度论文的平均年龄是中国论文平均年龄的 2 倍, 平均可被引年限也是中国论文平均可被引年限的 2 倍, 理论上讲, 印度论文的被引率和篇均被引频次都应该高于中国。

(2) 中印两国他引论文占全部施引论文的百分比几乎相同。

换句话说, 中印两国论文的自引率相差无几。这一结果显示出两国作者自引行为的相似性。

(3) 中国的 h 指数略低于印度的 h 指数。

注意两个事实: 中国发表 115 篇论文, 比印度少发表 47 篇论文; 中国论文远比印度论文年轻。在这样的条件下, 中国论文的 h 指数仅比印度低 2, 说明中国论文被引用情况是优于印度的。

3.2 同龄论文的篇均引文比较

中印两国论文的年龄分布不同, 一般引文指标只能大致反映两国的引文情况, 只有比较两国同年发表论文的篇均引文才能客观反映两国论文的学术影响。2000 年以前中国发表论文太少, 不便比较, 表 2 列出的是 2001—2009 年中印两国同年发表论

文的篇均被引频次。这里没有排除自引, 因为“一般引文指标”部分显示两国论文的自引率非常接近。9 年中有 2 年印度的篇均被引频次高于中国, 其余年份中国的篇均被引频次高于印度, 有些年份相差悬殊。结论是, 最近 9 年中国论文在国际上的影响大于印度论文在国际上的影响。

表 2 中国和印度: 同龄论文的篇均引文比较(2001—2009 年)

发文年	中国			印度		
	论文数	总被引频次	篇均被引频次	论文数	总被引频次	篇均被引频次
2001	4	24	6.0	8	33	4.1
2002	5	55	11.0	11	50	4.5
2003	7	23	3.3	7	51	7.3
2004	11	86	7.8	6	27	4.5
2005	8	80	10.0	7	28	4.0
2006	11	49	4.5	8	15	1.9
2007	12	31	2.6	5	5	1.0
2008	16	13	0.8	4	4	1.0
2009	23	3	0.1	6	0	0.0

3.3 他引论文作者的国家(地区)分布

分析他引论文作者的国家(地区)分布可以使我们了解中印两国论文学术影响的辐射情况, 得到的主要结论是:

(1) 中国作者引中国论文的比例(15.1%)高于印度作者引印度论文的比例(12.5%)。(2) 中国作者引印度论文的比例(8.3%)远高于印度作者引中国论文的比例(1.9%), 表现出一种极端的不对称。部分原因是由于中国论文普遍年轻, “被引用窗口”较小; 而印度近年来发文比较少, 所以引用别人论文的机会也较少。还有没有其他原因? 值得探究。(3) 除中印两国的本国作者外, 美国作者对中印两国论文引用最多, 分别达到 10.5% 和 12.3%。从篇次百分比看, 英国、比利时及荷兰作者对中国论文的关注要多于对印度论文的关注, 日本和法国则对印度论文的引用更多一些。(4) 和印度比较, 引用中国论文作者的国家(地区)分布更集中一些, 排序前 6 位国家的作者引用篇次之和占到总篇次的 59.8%。印度该比例为 52.7%。

3.4 高频被引论文

高频被引论文是反映一个国家论文影响的重要指标。表 3 列出中印两国各 5 篇最高频被引论文。

表3 中国和印度: *Scientometrics* 5篇最高频被引论文

国家	作者	论文题目	年卷期	总被引 频次	他引 频次
中国	Liang LM 梁立明	<i>h</i> -index sequence and <i>h</i> -index matrix: Constructions and applications	2006,69(1)	32	30
	Leydesdorff L Zhou P 周萍	Are the contributions of China and Korea upsetting the world system of science?	2005,63(3)	30	23
	Jin BH 金碧辉 Wang B	Chinese science citation database: Its construction and application	1999,45(2)	29	20
	Ren SL 任胜利 Rousseau R	International visibility of Chinese scientific journals	2002,53(3)	28	24
	Guan JC 官建成 Ma N	A comparative study of research performance in computer science	2004,61(3)	18	9
印度	Arunachalam S Srinivasan R Raman V	International Collaboration in Science—Participation by the Asian Giants	1994,30(1)	29	26
	Egghe L Rao lkr	Classification of Growth-Models Based on Growth-Rates and Its Applications	1992,25(1)	29	22
	Arunachalam S Garg Kc	A Small Country in A World of Big Science—A Preliminary Bibliometric Study of Science in Singapore	1985,8(5—6)	24	18
	Nagpaul PS	Contribution of Indian Universities to the Mainstream Scientific Literature—A Bibliometric Assessment	1995,32(1)	21	21
	Bhattacharya S Basu PK	Mapping a research area at the micro level using co-word analysis	1998,43(3)	19	16

分析中印两国最高频被引论文,我们得到这样一些认识:

(1) 中国5篇最高频被引论文都是最近10年发表的,最晚的一篇发表于2006年。印度的5篇都发表于10年以前,最早的一篇发表于1985年。显然,最近10年中国的研究比印度的研究更受关注。

(2) 中印两国高频被引论文被引频次属于同一数量级。

(3) 印度每篇高频被引论文的作者中至少有一位是印度的高产作者(见表1);并非所有中国高频被引论文的作者中都有中国的高产作者。

(4) 中印两国高频被引论文作者中均有国际作者。印度的国际作者L. Egghe是比利时学者,中国的国际作者L. Leydesdorff和R. Rousseau分别来自荷兰和比利时。特别值得提出的是,现任国际科学计量学与信息计量学学会主席Rousseau教授1992—2009年期间先后与16位中国学者合作发表了14篇*Scientometrics*论文,对促进中国科学计量学发展起到十分重要的作用。

4 主题词及主题词组分析

主题词分析是对论文研究内容及研究方法进行考察的有效途径^[9]。利用WordSmith Tools软件的Wordlist和Concordance功能可以得到中国和印度*Scientometrics*论文的主题词表和主题词组表,在此基础上,我们定位两国的研究兴趣,并与世界科学计

量学大背景做比较。

4.1 高频主题词

利用WordSmith Tools软件的Wordlist功能分别对115篇中国*Scientometrics*论文和161篇印度*Scientometrics*论文的题目作词频分析,得到中国和印度两个词表。去掉虚词,并将同根及同义词归类,得到实词主题词表(略)。

分析高频主题词可以总结出两国研究选题的一些特点。

China(Chinese, China's)为中国论文的最高频主题词,115篇中国论文中有50篇研究的问题与中国相关,占全部论文的43.5%。印度作者也关注本国问题,但论文题目中含主题词India(Indian, India's)的印度论文的比例为32.1%,远小于中国相应比例。

中印两国研究兴趣各有侧重。例如,主题词citation(s)和cited表征的引文研究占中国论文的17.4%,但只占印度论文的8.6%。主题词journal(s)提示的期刊研究也占中国论文的17.4%,但该词并非印度前10位的高频主题词。表示科学合作的collaboration(collaborative, collaborator)和表示技术研究的technology(technological)分别占印度论文的12.3%和9.3%,这两组词均不在中国前十位高频词之列。

中印两国均有论文专门研究计量指标(index, indices, indicator(s)),但计量指标研究并非两国最

热门选题。

中印两国学术用语有一定差异。例如,中国学者多用“bibliometric”(15 频次),少用“scientometric”和“scientometrics”(共 6 频次)。印度学者多用“scientometric”和“scientometrics”(共 28 频次),少用“bibliometric”和“bibliometrics”(共 14 频次)。

2003 年北京 ISSI 大会后中国科学计量学飞速发展,将中国论文主题词与世界论文主题词作比较

可以从研究选题方面揭示这种飞速发展的原因。最近 5 年(2005—2009 年) *Scientometrics* 共计发表 726 篇论文,其中,至少有 1 位中国作者的论文 70 篇,至少有 1 位印度作者的论文 30 篇。用 WordSmith Tools 软件的 Wordlist 功能分别对世界、中国和印度论文的题目作词频分析,得到 3 个主题词表。表 4 列出 3 个词表中频次最高的 10 个主题词。

表 4 世界、中国和印度:词频排序前 10 位的主题词(2005—2009 年)

世界 726 篇论文	频次	中国 70 篇论文	频次	印度 30 篇论文	频次
science(s), scientific	236	China, Chinese, China's	27	research, researchers	17
research, researcher's, researcher(s)	213	science(s), scientific	26	India, Indian, India's	10
citation(s), citable, cited, citing	139	research, researcher	19	analysis	8
analysis(-es), analytical, analyze(-ing)	134	analysis(-ses)	18	index, indices, indicator(s)	7
index, indices, indicator(s)	115	journal(s)	15	scientometric, scientometrics	6
journal(s), journal's	105	citation, cited	12	science	5
bibliometric(al), bibliometrics	79	index, indices, indicator(s)	11	study	5
impact(s)	74	bibliometric	10	citation, cited	5
study(-ies)	72	publication(s)	10	country(-ies)	4
publication(s)	70	study	9	publication(s)	4

我们注意到,表 4 展示的中国词频前 10 位的主题词和世界前 10 位的主题词基本一致,说明中国学者是在国际科学计量学热点领域从事研究的,例如引文研究、指标研究、期刊研究等。除了聚焦国际研究热点外,中国论文的中国特色依然十分明显。China(Chinese, China's)依然是排序第一的主题词,70 篇论文中有 27 篇研究中国问题,占到 38.6%。当然,和 1984—2009 年期间研究中国的论

文占全部中国 *Scientometrics* 论文的比例 43.5% 相比,该比例已经缩小了。印度也有三分之一论文研究本国问题,这一比例小于中国的相应比例。

4.2 高频主题词组

利用 WordSmith Tools 软件的 Concordance 功能可以得到词组表。表 5 列出世界 726 篇论文和中国 70 篇论文题目中频次最高的 4 个词组。30 篇印度论文的主题词组非常分散,最高频次为 2,此处不再列出。

表 5 世界和中国:词频排序前 4 位的主题词组(2005—2009 年)

世界 726 篇论文	频次	中国 70 篇论文	频次
bibliometric analysis(-es)(bibliometric study)	38	bibliometric analysis(-es)(bibliometric study)	8
<i>h</i> -index (<i>h</i> -indices, Hirsch index, Hirsch-index, Hirsch-type indices)	36	publication delay(s)	5
impact factor(s)	28	<i>h</i> -index (<i>h</i> -indices, Hirsch type indices)	4
self citations(self citation(s))	15	impact factor(s)	4

从表 5 中看出,中国最高频主题词组和世界最高频主题词组是一致的,都是文献计量分析与研究。中国第 3 高频词组和世界第 2 高频词组是一致的,说明中国融入了始自 2005 年末的 *h* 指数研究热潮。影响因子研究是科学计量学领域一个传统而又常新的选题,impact factor(s)是世界第 3 高频主题词组,5 年频次达到 28。中国论文中该主题词频次为 4,显示出中国学者对影响因子研究的兴趣。世界排序第 4 的高频主题词组为 self citation(s),在中国论文中,该词组词频为 1。特别值得注意的是中国论文

的第 2 高频主题词组 publication delay(s)。世界论文该词组频次为 6,其中 5 次出现在中国论文的题目中。中国的这几篇论文主要研究期刊出版时滞对期刊各项指标的影响。中国与世界主题词组的比较研究展示出中国与世界整体研究趋势的一致性,同时凸现中国研究选题的特殊性。

5 发展中国科学计量学的几点建议

本文作者 10 年前曾写过一篇文章“科学计量学与信息计量学:从世界看中国”^[10],文中谈到发展中

国科学计量学的几点建议。那是作为中国惟一代表刚刚参加完在墨西哥举行的第7届国际科学计量学与信息计量学大会,感慨于国际科学计量学与信息计量学的飞速发展和中国在该领域的沉寂,觉得有义务为中国科学计量学的发展尽绵薄之力。如果说10年前建言献策是企盼中国科学计量学尽快崛起,现在提建议则是为了中国科学计量学的持续发展和新的阶跃。

一些常规的建议,例如,切近国际学术研究前沿,以原始创新研究立足国际舞台,建设好研究队伍,加强国际国内学术交流与合作,发挥学术团体的作用等等,尽管很重要,但已经众所周知,就不再重复了。以下是些很具体的建议。

5.1 了解科学计量学的学科结构,扩展选题视角

盘点 *Scientometrics*, *Journal of the American Society for Information Science and Technology* 和 *Journal of Informetrics* 等国际期刊上中国学者发表的科学计量学论文,我们有这样一个整体感觉:选题范围比较窄,以应用研究为主,真正开展理论探讨和方法研究的论文不多。这种研究取向是一种瓶颈,会限制中国科学计量学向广度和深度的进一步发展,应该引起我们的高度重视。

2008年以色列学者 Judit Bar-Ilan 在 *Journal of Informetrics* 上发表了一篇气势磅礴的综述论文,题目是“21世纪之初的信息计量学述评”^[11]。在这篇论文中,科学计量学与信息计量学研究被分为13个方面:(1) General works on informetrics; (2) Theory; (3) Methods and techniques; (4) Citation analysis; (5) Indicators; (6) Webometrics; (7) Journals; (8) Open access and electronic publications; (9) Productivity and publications; (10) Collaborations; (11) Research policy; (12) Patent analysis; (13) Databases。每个方面又包含若干类选题,例如,“Methods and techniques”包括 Mapping and visualization, Text and data mining and other linguistic techniques, Network analysis, Classification 和 Other methods。建议我国科学计量学与信息计量学研究者耐心地读一读这篇长达50页、引用了600多篇参考文献的信息量极大的论文,这对我们了解科学计量学与信息计量学的学科结构,把握学科研究态势,从理论和应用两方面扩展选题视角大有裨益。

5.2 结合学科优势,发展特色方向

科学计量学是一个年轻的学科,不像数学和物

理学那样有着深厚的研究根基和稳定的研究队伍。中国科学计量学的研究者来自不同学科,有着不同的知识结构和研究背景。如何结合自己的学科优势发展科学计量学的特色研究方向,我们提供3个例子。

大连理工大学的科学计量学研究力量主要源于科学学与科技哲学。在刘则渊教授领导下,大连理工大学创办了网络-信息-科学-经济计量实验室(WISE LAB),实现了科学学范式下科学计量学的可视化转向。通过引进信息可视化国际顶尖学者、旅美华人陈超美教授及培养自己的力量,形成了以知识图谱技术为核心的研究特色,完成了多个国家级项目,取得了一系列有影响的成果^[12]。

上海交通大学高等教育研究院刘念才研究团队的研究优势是高等教育研究。从2003年开始该团队连续发布“世界大学学术排名(ARWU)”。ARWU引起国际一流科学计量学专家的关注,由此引发的讨论深化了关于大学评价的认识,改进了大学排序的工作。从2005年开始该团队发起并连续主办三届“世界一流大学国际研讨会”,A van Raan等4位国际科学计量学普赖斯奖获得者与会和访问交流。

武汉大学的情报学研究在国内有举足轻重的地位。该专业的邱均平教授整合信息管理学院、学校图书馆、计算机中心与图书情报研究所等单位的相关力量,联合组建武汉大学中国科学评价研究中心。该中心将科学计量学的理论与方法广泛应用于评价领域,定期出版系列研究成果,编辑出版《评价与管理》杂志,并面向社会开展评价服务。

中国科学计量学需要多发展一些这样的特色研究,特色研究的累积会形成中国的特色和优势。

5.3 国家研究机构承担应用研究和基础研究双重任务,要努力做到双重精彩

在中国 *Scientometrics* 论文数量的机构排序中,中国科学院和中国科学技术信息研究所是进入前10名的两家研究机构。作为国家研究机构,他们承担着重要的国家任务,其中不少任务用到科学计量学方法。

例如,中国科学院国家科学图书馆科学前沿分析中心金碧辉研究团队基于各种大型数据库的数据,采用科学计量学方法、数据挖掘和可视化技术,分析、监测和映射世界科学前沿与科学热点,监测世界科学前沿的发展态势,完成了多项国家项目,出版了诸如《世界科学中的中国》等系列研究报告。国家科学图书馆还构建了科技论文与引文分析数据库,

一年一度发布科技论文与引文分析报告。这些工作为科技决策和管理人员提供了决策和管理依据。

中国科学技术信息研究所承担了国家科技部中国科技论文统计与分析项目,选择 SCI 等国际数据库,利用科学计量学的理论与方法对中国的科研产出状况、在世界所处的位置影响以及世界研究态势的变化等进行了全面的分析研究。该研究所构建了中国科技论文与引文数据库,每年定期向社会发布《中国科技论文统计与分析》年度研究报告和《中国科技期刊引证报告》,为国家的科技决策提供了有力的支撑。作为该所的总工程师,武夷山研究员在主持完成国家任务的同时在国际科学计量学舞台上也十分活跃。2008年12月,中国科学技术信息研究所与汤森路透集团签署了战略合作协议,成立了科学计量学联合实验室,设立开放基金,开展与科学计量学相关的基础理论、技术与应用研究。

我们的国家研究机构远不止这两家。国家研究机构不仅承担着国家指派的应用研究任务,还扮演着基础研究国家队的角色。在科学计量学领域能否做到,以及如何做到双重任务,双重精彩,这是需要认真研究的问题。印度的国家科学技术与发展研究所发表了80篇 *Scientometrics* 论文,发文量居世界机构排名第二。他们的经验是不是值得我们研究和借鉴呢?

5.4 及时学习和充分利用各种数据平台的技术支持,提高研究的效率和技术含量

目前可用于科学计量学研究的数据平台多种多样。在国际上,应用最广泛的是汤森路透(Thomson Reuters)公司的数据库 Web of Science。近年来,Scopus 和 Google Scholar 的适用性也日益提升。国内可用于科学计量学研究的数据平台发展非常快,中国科学技术信息研究所的 CSTPCD 数据库、中国科学院国家科学图书馆的 CSCD 数据库、清华大学的 CNKI 数据平台等都为国内科技论文和引文分析做出了重要贡献。国内外的数据库技术、检索技术和数据分析技术日新月异,科学计量学研究者要与时俱进地学习和了解这些新进展,提高研究的效率和研究的技术含量。

以汤森路透公司为例。该公司基于 Web of Science 数据库推出的 InCites 工具在机构科研绩效评估方面有特殊的意义。该分析工具能够提供包括研究成果数量、学术影响力、学科特色和科研合作等方面在内的多指标分析体系和国际对比基准(Benchmark)。InCites 中不但采用了多种先进的计

量指数,还能够将多角度分析结果以可视化的方式集中呈现在一个分析面板中,便于分析人员很快地了解和掌握机构研究现状。

5.5 全面盘点国家自然科学基金资助的科学计量学项目,提高科学基金资助效益

国家自然科学基金为发展中国的科学计量学事业立下汗马功劳,主要表现在两个方面。第一,国家自然科学基金委员会管理科学部是2003年北京国际科学计量学与信息计量学大会的主要赞助者。前文已经说明,正是那次大会引来了中国科学计量学的春天。第二,多数科学计量学的高产作者、高频引用论文的作者都是在国家自然科学基金项目的支持下开展研究工作的。国家对发展科学计量学十分重视,投入了可观的资金。建议国家自然科学基金委员会管理科学部对资助的科学计量学项目进行一次全面盘点,分析选题,考察业绩,了解队伍,发现问题,做到心中有数,为改善资助政策和项目管理,从而进一步提高资助功效提供依据。这种盘点也可以为我们提供国家自然科学基金支持一个学科从起步到国际化的范例。

6 需要进一步讨论的问题

1984—2009年期间研究中国问题的论文占全部中国 *Scientometrics* 论文的比例高达43.5%,远大于研究印度问题的印度论文的比例(32.1%)。这使我们想起多年前曾经研究过的国际著名科技政策研究机构英国苏塞克斯大学的 SPRU,词频分析显示,SPRU 论文中与英国相关的研究占到34.4%^[9]。SPRU 的论文包括在国际上发表的论文,也包括在本土发表的论文。一般说来,本土期刊发表的论文研究本国问题的比例更大一些。应该如何看待中国学者在国际期刊 *Scientometrics* 上发表高比例的研究中国问题的论文? 对此的回答一定是仁者见仁智者见智。一种可能的意见是,国际期刊的读者遍布世界,多在国际期刊上发表研究中国问题的论文可以让世界更好地了解中国。尤其是,中国国力的稳步增强使世界了解中国的渴望与愿望与日俱增,发表研究中国问题的论文刚好有助于满足国际上了解中国的愿望。相悖的意见是,研究中国问题的比例过大意味着中国学者投入其他国际性研究选题的力量减少,从科学研究国际化的角度看,本国研究的比例占到40%似乎高了些。投稿国际期刊,如何在研究中国问题与研究其他国际性选题之间保持必要的张力? 能否寻求一种动态的平衡? 不仅是科学计量

学研究者,其他管理科学、社会科学、人文学科以及交叉科学研究者也会有同样的困惑。我们将此作为一个公开讨论的问题提出来,希冀引发学界同仁的思考。

参 考 文 献

- [1] Price D. Little Science, Big Science. New York: Columbia University Press, 1963.
- [2] Price D. Networks of scientific papers. *Science*, 1965, 149 (3683): 510—515.
- [3] Garfield E, Sher I. New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, 1963, 14: 195—201.
- [4] Zhao H. An intelligence constant of scientific work. *Scientometrics*, 1984, 6(1): 9—17.
- [5] Schoepflin U, Glänzel W. Two decades of “scientometrics”—An interdisciplinary field represented by its leading journal. *Scientometrics*, 2001, 50(2): 301—312.
- [6] Dutt B, Garg K C, Bali A. Scientometrics of the international journal *Scientometrics*. 2003, 56(1): 81—93.
- [7] Hou H, Liu Z et al. Visualization of international networks of Chinese scientometrics. Berlin: Proceedings of WIS 2008, Berlin, 1—11.
- [8] Schubert A. The Web of *Scientometrics*—A statistical overview of the first 50 volumes of the journal. *Scientometrics*, 2002, 53 (1), 3—20.
- [9] 梁立明,李小宁. SPRU 科研选题的词频分析与计量研究. *科研管理*, 2003, 3: 97—108.
- [10] 梁立明. 科学计量学与信息计量学: 从世界看中国. *科研管理*, 2000, 3: 95—101.
- [11] Judit Bar-Ilan. Informetrics at the beginning of the 21st century—a review. *Journal of Informetrics*. 2008, 2: 1—52.
- [12] 刘则渊. 基于知识图谱的科学学与科技管理学科建设——大连理工大学科学学博士点学科建设情况汇报. 全国第四届科技政策与管理学术年会, 南京, 2008.

THE DEVELOPMENT OF CHINA'S SCIENTOMETRICS: PUBLICATION-CITATION ANALYSES AND COMPARISON BETWEEN CHINA AND INDIA

Liang Liming

(Institute for Science Technology and Society, Henan Normal University, Xinxiang 453007)

Abstract The development of China's scientometrics could be understood from three aspects: analysis of publications and citations of China's *Scientometrics* papers; topic words analysis using WordSmith Tools; the comparison between China and India—the productive country in scientometrics. Some findings are as follows. 1) After a longer launching period during 1984—2000, China's scientometrics has grown up in the new millennium. China has become a productive country in scientometrics. 2) In India the national institutions do the majority of their scientometric research, while in China universities' research performance is better. 3) China's productive authors are not as powerful as India's. 4) It is revealed by the citation measures that in the new millennium China's scientometric studies attract more attention from the world than India's studies do. 5) Studies on citations and journals share a large proportion of Chinese papers. The ratio of Chinese papers focusing on China's problems is higher than that of Indian papers focusing on India's problems. 6) In the last five years the topic words of high frequency in Chinese papers accord with those in the global papers. This fact illustrates that Chinese scholars are working at the hot topics in scientometrics. Some suggestions regarding how to develop China's scientometrics are proposed, and one open problem is offered for the further discussion.

Key words scientometrics, China, India, citation analysis, topic words analysis