

国内外引文分析研究进展综述

苑彬成^{1,2}, 方曙², 刘清³, 张晋辉^{1,3}

(1.中国科学院研究生院,北京 100190;2.中科院国家科学图书馆成都分馆,四川成都 610041;3.中科院国家科学图书馆武汉分馆,湖北武汉 430071)

摘要:对发表于2000-2008年间代表引文分析研究进展的大量情报学文献进行了评述,一方面勾勒出了引文分析发展的现状,另一方面也跟踪了引文分析里新出现的理论、方法,期望能为国内外同行了解引文分析最新进展、制定学科发展战略提供借鉴。

关键词:引文;引文分析;引文网络;自引;共引分析

中图分类号:G350 **文献标识码:**A **文章编号:**1007-7634(2010)01-0147-07

Overview on Progress in Citation Analysis at Home and Abroad

YUAN Bin-cheng^{1,2}, FANG Shu², LIU Qing³, ZHANG Jin-hui^{1,3}

(1. Graduate School of Chinese Academy of Sciences, Beijing 100190, China;

2. Chengdu Library of Chinese Academy of Sciences, Chengdu 610041, China;

3. Wuhan Library of Chinese Academy of Sciences, Wuhan 430071, China)

Abstract: In this article, we comment on a large number of documents published between 2000 and 2008 relevant to citation analysis in the area of Information Science. Not only we describe what it is like now, but also we trace new theories and methods. We hope that this article can provide a reference for peers at home and abroad to know better of its latest developments.

Keywords: citation; citation analysis; network citation; self-citation; co-citation analysis

引文分析是利用各种数学、统计学以及逻辑方法,对科技期刊、论文、著作等各种分析对象的引用或被引用现象进行分析研究,以便揭示其数量特征和内在规律,达到评价、预测科学发展趋势的目的^[1]。利用引文分析理论的指导来建立引文索引,已经有相当丰硕的成果,《科学引文索引》(SCI)即是全世界最具权威性的引文数据库,收录了世界各国超过6000种最重要科技期刊的引文状况。

1 引文分析研究的现状与进展

国内外对于引用文献的研究一直在蓬勃开展

收稿日期:2009-04-20

基金项目:中国科学院国家科学图书馆战略研究项目(NSL-001-<0503>)

作者简介:苑彬成(1983-),男,山东人,硕士研究生,主要从事情报计量分析研究;方曙(1957-),男,四川人,研究员,博士生导师,馆长,主要从事情报计量学理论与实践研究。

着,从一开始借引用文献探究论文与论文的相关性,到现在发展出很多延伸性的研究。近10年来,引文索引和引文分析得到越来越多的重视,2000-2008年国内外就有4000多篇相关文献公开发表,反映出有关引文分析的研究和应用仍是情报计量学中的研究热点。通过对这4000多篇文献进行主题统计,我们发现,近几年主要的关注研究对象或热点集中在“引文网络”、“自引与自引率”、“共引分析”、“影响因子”以及“引文分析和统计中的误差”等方面。

1.1 引文网络

在引文分析中,引文网络得到了学者们的高度

重视。国外方面,Davenport&Cronin(2000)^[2]将引文看作代表信赖的原型。他们认为在虚拟环境中,引文实际上代表了一种信赖。当A引用B的文章时,A认为B的观点是支持性的,而B引用的内容是被协商许可了的,引用内容的同时也就影响了B的思维。因此,引文网络系统可以被看成是信赖系统,引文索引则可以被看成是一个推荐系统——推荐被引次数多的文章。Fang&Rousseau(2001)^[3]将引文系统看成是网格,这些菱形网格对应的文献耦合与共引文献就是网格系统的基本结构。Small(2004)^[4]解释了引文理论与Merton关于科学人际关系网络理论的相互关系,他将引文看成是对知识债务的一种象征性偿还。Meister C&Meister M(2001)^[5],Chen&Hicks(2004)^[6],Brantle&Fallah(2007)^[7]等人认为研究专利引文网络,可以分析出其中知识流动和技术扩散的路径。Nerur(2005)^[8],Leydesdorff(2007)^[9]等人在期刊引文网络中分析了个别期刊在局部的相对影响力。除此之外,还有学者专门研究引文网络的结构,有的侧重于挖掘其中的核心文章,有的侧重于进行网络解构,还有的侧重于研究拓扑结构的动态变化及进化,如Shibata(2007)^[10],Valverde&Sole(2007)^[11]等。在这方面,国内研究相对较少,研究主题也比较分散,比较值得一提的是,柳泉波和许骏(2008)^[12],马楠和官建成(2008)^[13],以及黄亚明等人(2008)^[14]都引入了google的PageRank算法,分别从不同角度对引文网络进行了研究。

1.2 自引与自引率

关于自引和自引率的研究是引文分析的内容之一。目前在国外研究最多的是关于期刊的自引和自引率。一般认为,期刊高自引有助于提升其影响因子;也有研究人员研究去除期刊自引数对影响因子的影响,认为未去除影响因子的数据可以采用;还有研究显示期刊自引与自引率与影响因子的负相关性。以情报和图书馆学以及遗传学方面的杂志为研究对象,Nisonger(2000)^[15]检测了除去自引后对这些杂志影响因子的影响,其结论是可以使用没有去除自引杂志的影响因子(因为这种情况下影响因子会大些)。Fassoulaki等(2000)^[16]对1995-1996年出版在《英国麻醉学杂志》上的六种麻醉方面的期刊进行了调查,得出了自引对期刊本身的影响因子具有正影响效应的结论。在随后开展的一个研究中,Fassoulaki等(2002)^[17]又论证了影响因子对自引和自引率的负影响效应。Aksnes(2003)^[18]对挪威作者在

1981-1996年间发表的文章开展了大规模的作者自引研究,结果发现:在文章发表后的最初一段时间内自引率较高;文章的作者数量对自引率也有影响;学科不同,自引用率也不相同。Gl?nzl&Schlemmer(2004)^[19]对1992年被WoS编入索引的所有文章进行了研究,其研究结果支持前面Aksnes的结论:自引率随时间的推移而下降。他们还发现自引与非自引(或称为外引)并不是相互独立的,自引是引用程序的一个有机组成部分。在随后的一个研究中,Gl?nzl(2004)^[20]指出从宏观上计算文献计量学指标时,没有必要除去自引。

目前在国内关注较多的则是关于论文的自引和自引率。目前,国内科技论文的过度自引现象比较严重,潘文涛和武夷山(2007)^[21]曾对过度自引现象的弊端加以痛陈。但是从理论上说,自引本身并无错误,因为它能够反映科学研究的连续性、继承性、相关性,反映著者学术研究的发展道路、研究方向及研究过程。在肯定自引具有必要性的同时,近年来国内学者探讨比较多的是适度自引问题。以《编辑学报》为阵地,先后有熊春茹(2002)^[22]、朱大明(2004)^[23]等多位研究人员对自引的原则、不当自引的控制等问题进行了探讨。许多科研人员还在实证方面做了大量关于自引对期刊影响的研究,例如蒋颖和金壁辉(2000)^[24],李建辉(2007)^[25]等。另外,每年不仅会有关于某本期刊自引情况分析的文章,而且还有针对某领域内各种核心期刊自引的对比分析的文章。例如《<情报杂志>自引分析》一文就是对《情报杂志》1995~2004年的自引情况进行统计,对它的语种自引、著者自引、机构自引、学科自引以及期刊自引分别进行了研究^[26]。

1.3 共引分析

自1973年被提出以来,共引分析已成为引文分析里面一种潜在多产的分析方法,它不仅可以用来揭示科学结构的发展现状乃至变化情况,还可以用来进行前沿分析、领域分析、科研评价等,进而为宏观科技决策提供先行支持,为科技规划与评估提供基础。

自从McCain(1990)^[27]对共引分析的步骤详加阐述后,大家一直以来都按照这种约定俗成的方法来进行共引分析。但近几年来,以《JASIST》和《Scientometrics》为主要阵地,开展了大量关于优化共引分析的讨论。在2005年以前,该讨论主要集中在相似性系数的选取方面,例如White(2003b)^[28]和Bens-

man(2004)^[29]等人建议采用皮尔森相关系数 r 进行共引分析,但 Ahlgren 等人(2003)^[30]则反对采用这种度量,建议用 cosine,jaccard 等代替。针对自己采用皮尔森相关系数 r 的批评,White(2003)^[31]采纳一种新技术——网络寻址定位(PFNETs),它不需要使用相似性系数,直接使用共引矩阵的原始数据。Leydesdorff(2005)^[32]则认为,若想采用皮尔森相关系数 r 进行分析,应该对共引矩阵中原始数据进行对数转换,使之符合正态分布。在随后的研究中,Leydesdorff 和 Bensman(2006)^[33]对应用对数转换处理引文数据的方法进行了进一步的探讨。但 2006 年以后,该讨论扩展到了相似系数矩阵的生成方式方面:是该利用共引矩阵还是引文矩阵。例如,Leydesdorff 和 Vaughan(2006)^[34]提出应将相似性系数用于引文矩阵进而生成相似系数矩阵,而共引矩阵本身就是临近矩阵,应直接导入 SPSS 进行多元分析,无须由一个临近矩阵转化成另外一个临近矩阵(相似系数矩阵),并就该观点给出了自己的例证。Schneider&Borlund(2007)^[35]也对相似性矩阵的生成方式进行了探讨。但随后,Waltman&Van Eck(2007)^[36]指出 Leydesdorff 和 Vaughan(2006)的观点是基于 SPSS 程序的错误得出的结论,并给出了自己的反驳意见。Leydesdorff(2007)^[37]对于反驳的意见并不十分赞同,并基于社会网络分析软件 pajek 的分析结果再次给出了坚持自己观点的论据。我们相信,这些文章无疑会给共引分析带来更大范围的讨论。

国内方面,近年来以综述性的文章居多,例如王建芳和冷伏海(2006)^[38],耿海英和肖仙桃(2006)^[39]等人概括地介绍了共引分析的起源、分类、发展及国外的应用诸方面。实证方面,大多数学者也是沿着约定俗成的方法进行共引分析,但也有少数学者加入到优化共引分析方法的大讨论中,例如邱均平、马瑞敏等(2008)^[40-41]针对目前学术界对共引分析方法中共引矩阵对角线该如何设置的激烈讨论,给出了自己的解决方案。

整体来看,近几年国内外关于共引分析的研究进展体现在以下方面^[39]:①综合运用多种分析方法。例如将共引分析结果和文献耦合、共词聚类、词频统计等方法的分析结果加以比较分析;②不断融入新的技术。由最初借用多维尺度技术进行降维,到现在用 PFNETs 替代 Pearson 相关系数,甚至引入自组织映射(SOM)技术、潜在语义索引(LSI)技术等;③扩展至网络结构研究。用网页共链分析(WCA)反映网络本身的结构和网络中知识的结构;④不断探究共引

分析中的一些细节问题,包括相似性计算方法的优化,如何对合著者进行所有作者的共引分析等。

1.4 影响因子

作为引文分析的延伸性研究,毋庸置疑,关于影响因子的争论是情报计量学近年发展历程中一个非常重要的现象。其中 2002 年由《自然》杂志刊登的一篇报道引发的一场讨论很有代表性。Adam(2002)^[42]的这篇文章报道了关于引文计数和影响因子的争论,并讨论了在计算影响因子时“可引用”论文的定义。Walter 等人(2003)^[43]强调期刊的影响因子和引用计数作为衡量科学质量的手段被误解和误用了,建议以“最佳文章”的同行排名取而代之。Lundberg(2003)^[44]认为影响因子是衡量期刊、期刊文章以及作者价值的糟糕手段。Colquhoun(2003)^[45]指出影响因子只是一个平均值:在一份影响因子相对较低的期刊中可能存在高引用数的论文,或者说即便发表在一份影响因子很高的期刊上这篇论文也可能被较少引用。《美国放射学杂志》的主编 Rogers(2001)^[46]也抱怨影响因子是杂志编辑们操纵的数字游戏。

尽管有如此之多的负面观点,但似乎人们并不考虑完全放弃这一指标。正如 Lankhorst & Franchignoni(2001)^[47]所说:“尽管对于影响因子的使用被认为是不合理的,但无论如何这种情况仍然继续,部分原因可能在于目前还没有提出更好的公认的评价手段。”学者们在关注影响因子的缺点的同时,还提出了一些改进的建议,如 Van Leeuwen&Moed(2002)^[48],Egghe&Rousseau(2002)^[49]等。结合其他的手段和方法,采纳影响因子作为评价期刊的手段暂时来说是无法替代的。

作为一个研究热点,影响因子的未来走向虽然未必会跌宕起伏,但一定要引起我们的重视,至少我们应该关注其可能的变化和修正。如果有非常好的修正方案,未必不能在我们的实际工作中进行探索性应用,特别是在面向科研评价战略情报研究活动中进行应用。这方面国内学者也开展了一些工作,例如马廷灿、黄可等人(2007)^[50]在分析期刊影响因子局限性的基础上,提出了对影响因子的一种修正方法,提出了期刊综合影响力指数,具有很好的可操作性。

1.5 引文分析和统计中的误差

引文分析以其独特的科学评价功能而备受推崇,然而随着应用的深化,越来越多的人开始对引文

分析的有效性与可靠性提出质疑。因为无论是引文分析的对象,还是引文分析的方法本身,都存在一些虚假和错误的成份。虽然不是很热门,但针对引文分析和统计误差的研究一直伴随着SCI的成长历程。尽管专论相对而言不是很多,但在不少采用引文数据进行研究工作的文献中多少都存在对引文分析中缺陷和错误的分析和研究,很多文献也以善意的提醒或者警示来告知读者要谨慎地使用引文数据。

2000年,Fenton等人研究了医学文献中参考文献的准确性,发现存在高达37.5%的错误^[51]。2002年,《自然》发表了一篇被广泛转载的短文,在这篇文章里,作者提醒要注意ISI引文统计数据中的错误^[52]。2004年,有研究人员研究错误引用的情况,追踪错误引用的源头^[53]。2006年,Wetterer则用真实的案例告诉读者错误的引用如何导致谬误的结果^[54]。

目前,国内对于引文虚假和误差讨论最多的当属伪引和漏引。有些引文概念,如转引、崇引、媚引等,它们的共同特点是作者在文后列出了一些与论文内容无关的参考文献,有的学者将它们统称为伪引。罗时嘉(2001)^[55],胡锦涛(2003)^[56]以及郭小兰(2004)^[57]等人剖析了伪引的机理,指出了它的危害,并就如何消除伪引提出了建议。漏引方面,据统计,国外期刊90%以上的论文都附有引文篇目,国内科技期刊论文附有引文篇目的也达到了88%,而在国内社会科学期刊和论文里附有引文篇目的仅占54.32%,漏引现象较为突出。对此,戴勇和张四新(2004;2005)^[58-59]全面分析了漏引的成因,指出了它可能造成的严重后果,并提出了防止漏引的意见和建议。除了进行理论上的探讨,也有大量学者针对引文分析和统计中的这些不规范行为进行了实证分析,例如梁立明和钟镇(2007)^[60],邓迎和侯汉清(2008)^[61]等。

2 引文分析新出现的理论与方法

关于引文行为和引文分布的研究,长期以来是情报计量学的重要领域。从理论上讲,可以说这是一个充满宝藏的领域:底层的数据明晰而易懂,但如果深入研究,又蕴含了非常多的内容,我们可以看到有人从社会学(引文行为)角度来分析这些数据,也有人从科学学(引文分布)角度来分析这些数据,不一而足。随着信息技术的发展,对引文数据的深度分析具备了可操作性(虽然这种可操作性有限。在这一领域,商业元素起决定性作用,信息的不对称性还比较

突出),近几年诞生了一些重要的理论和方法,产生了一些新的亮点。

2.1 科学的睡美人现象,第n次被引和未被引

所谓“科学的睡美人”(Sleeping Beauties in Science)是指那些长期不被引用(10-20年休眠)然后突然成为高被引论文的论文。该理论是由荷兰雷顿大学的著名学者Raan提出的。他在2004年发表了名为“科学的睡美人”^[62]的短篇通讯,按照Google Scholar的数据,该文被引用34次(截止2009年2月25日)。其实这一现象加菲尔德1980年就曾关注过,即文献的重要性或价值在文献出版后很久才得到认可。Raan提出了三个变量:①睡眠深度。包括两种形态:文献年均被引次数最多1次为深度睡眠,1-2次则为次深度睡眠。②睡眠时长。指睡眠状态持续的时间。③唤醒强度。指睡眠结束后四年的年均被引频次。Raan设定不同的睡眠时长(5-10年)来研究不同情况下的唤醒强度,并分析了一篇关于弦理论的文献(1988年发表,唤醒强度64)。他认为在论文发表时,相关的理论物理研究还处于“睡美人状态”,到1995年人们开始意识到该文献的价值,于是纷纷引用。这样,睡美人被唤醒了。Burrell(2005)^[63]在一篇文献中认为现有的引文过程模型能够解释这一现象。Simkin(2007)^[64]也提出一种“引文的数学理论”,认为他们的理论能够解释“科学的睡美人”现象。

一般而言,“科学的睡美人”现象源于“超前研究”ⁱ。对科学家而言,“超前研究”既具有吸引力,同时也让人担忧。担忧的原因是其文献或研究成果从此杳无音讯,无人问津。

第n次被引与“科学的睡美人”现象在本质上相同,研究的都是引文的时间分布问题。不过前者揭示的是普遍性的问题,或者说基于统计学的宏观引文行为,主要探讨引文的时间分布特征;后者则存在一个“唤醒”的问题(唤醒条件)。与第n次被引相对应的,还有一个概念“未被引率”,近几年被提的比较多。在进行评价应用时,人们也常常考虑这一指标。在进行宏观引文分析时,无论所取得时间间隔有多长,始终有部分文献落在了无人问津的境地,在发表后若干年内都没有被引的记录。这类文献中的大部分将继续沉寂,一部分可能会收到引用记录,当然还有一些可能就是“科学的睡美人”,其受关注的程度突然会大增。

虽然引文的时间分布很久以来就得到关注,但过去受技术和数据的限制,相关的研究不太深入,往

往是小数据量的分析和推测,Raan 所开展的研究主要得益于 CWTSii 的数据积累。把“科学的睡美人”当作一种新的理论来重点介绍,主要是考虑到这一种现象为我们提供了一个比较独特的、富有挑战的研究领域。按照 Raan 的理论,如何界定三大变量,是否能超前发现“超前研究”等都很具有研究意义。

2.2 h 指数

Hirsch(2005)^[65]提出 h 指数来评价研究人员学术成就的时候,恐怕并未料想到所能引发的轰动效应。从时间上讲,作为一种方法,面世只有 3 年多的时间,比较新颖,但是从其短短的发展历程来看,我们不得不承认其发展和膨胀的速度是如此之快,以至于在不知不觉中一些商业数据库推出了基于 h 指数的指标,显得有些“成熟”了。把 H 指数作为一种新方法来推荐,主要原因一是考虑到其相对于传统引文分析方法的创新性贡献,二是考虑到其影响非常深远。

其创新性贡献在于,把人们垢弊多年的论文数量、被引频次用一种独特的方式展现,既保留了它们的本质特征,又在这些特征的基础上提出了不一样的体系,即所谓“h 篇论文分别被引用了至少 h 次”。这一种体系是否能够比较准确地反映一个人的学术成就并没有定论,但是从对顶尖科学家的分析结果来看,确实具有一定的价值。当然,也存在一定的弊端,比如依然没有解决跨学科比较的问题,与学者从事研究的时间长短存在关联。

之所以说 h 指数影响深远,主要还在于其评估小数据集时的潜力。其他传统的文献计量方法和指标在面临小数据集时往往无能为力,或者力不从心。其次,人们通过后续的大量研究(对 h 指数的修正、补充和完善,如衍生出来的 g 指数,a 指数等),大大拓展了 h 指数的适用范围。第三,一些商业数据库开始把 h 指数作为一项服务内容,如 Scopus 就推出了作者的个人 h 指数分析。

2.3 自我中心的引文研究

所谓自我中心(ego-centered)的引文研究始于某个独立的特定对象。围绕这个特定对象,有一批不同类型的作者或作者群:特定作者的合作者,特定作者引用的作者,引用特定作者的作者,与特定作者共被引的作者。这种作者和作者群的复杂关系构成了研究自我中心的引文分析的基础。White(2000)^[66]首先采用这种方法,先是研究了 Garfield 的引文网络,

稍后又研究了其他 8 位情报学家的引文特征。2001 年他还详细阐述了生成个人引文地图(包括上面提到的四种对象)恰当的数据收集方法^[67]。Bar-Ilan(2006)^[68]从 WoS,Citeseer 和 Google Scholar 等数据库收集信息,得出了关于计算机科学家 Michael Rabin 的一系列引文特征。还有其他一些学者也采用这种方法分析不同作者的引文网络和绘制引文图谱,如 Cronin 和 Shaw(2002)^[69]等。还有研究人员尝试利用特定作者的引文和合著关系网络,来分析该作者的知识结构,揭示该作者对各个主题域的贡献,如 Huang TH 和 Huang ML(2007)^[70]。这里有必要指出,以特定作者为中心,还存在一些其他网络,例如人际关系网络,是属于社会网络分析(SNA)中个体中心网的研究范畴,在这里不再详述。但将它里面的诸多概念引入引文网络分析,像网络密度,中心性,核心-边缘结构等,势必会给引文分析带来新的生机和活力。这种新方法的运用在国内也初见端倪,例如,郭永正(2008)^[71]就尝试对中、印两国国际科学合作的个体中心网结构进行了比较,发现两国国际科学合作个体中心网的网络密度、中心势以及核心合作伙伴基本相同,但有明显不同的链密度。

3 结 语

本文对发表于 2000-2008 年间代表引文分析进展的大量情报学文献进行了评述,一方面勾勒出了引文分析发展的现状,另一方面也跟踪了新出现的理论、方法,期望能为国内外同行了解引文分析最新进展、制定学科发展战略提供借鉴。

参考文献

- 1 庞景安. 科学计量研究方法[M].北京:科学技术文献出版社,2002:216-217.
- 2 Davenport, E., Cronin, B. The Citation network as a Prototype for Representing Trust In Virtual Environments [A]. In: Cronin, B and Atkins, H.B. (eds.), The web of knowledge: a Festschrift in honor of Eugene Garfield [C]. Metford: Learned Information Inc, 2000:517-534.
- 3 Fang, Y., Rousseau, R. Lattices in citation networks—An investigation into the structure of citation graph [J]. Scientometrics, 2001, 50(2): 273 - 287.
- 4 Small, H. On the shoulders of Robert Merton—Towards a normative theory of citation [J]. Scientometrics, 2004, 60(1): 71 - 79.
- 5 Meister C, Meister M. Trends and trajectories in MEMS-related technologies: an analysis on the basis of patent appli-

- cation data [EB/OL]. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1558743&isnumber=33123>, 2009-12-16.
- 6 Chen CM, Hicks D. Tracing knowledge diffusion [J]. *SCIENTOMETRICS*, 2004, 59(2): 199-211.
 - 7 Brantle TF, Fallah MH. Complex innovation networks, patent citations and power laws [A]. In *PICMET '07, Vol.1-6, Proceedings - management of converging technologies [C]*. Portland: Oregon, 2007: 540-549.
 - 8 Nerur S, Sikora R, Mangalaraj G, et al. Assessing the relative influence of journals in a citation network [J]. *Communications of the ACM*, 2005, 48(11): 71-74.
 - 9 Leydesdorff L. Visualization of the citation impact environments of scientific journals: An online mapping exercise [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(1): 25-38.
 - 10 Shibata N, Kajikawa Y, Matsushima K. Topological analysis of citation networks to discover the future core articles [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(6): 872-882.
 - 11 Valverde S, Sole RV, Bedau MA, et al. Topology and evolution of technology innovation networks [J]. *Physical Review*, 2007, 76(5).
 - 12 柳泉波, 许骏. 基于链接分析的文献个性化排序算法 [J]. *中山大学学报(自然科学版)*, 2008, 47(6): 87-92.
 - 13 马楠, 官建成. 基于网络结构挖掘算法的引文网络研究 [J]. *情报学报*, 2008, 27(4): 584-590.
 - 14 黄亚明, 王琳辉, 金碧辉. 期刊引文网络影响测度研究 [J]. *情报学报*, 2008, 27(2): 265-270.
 - 15 Nisonger, T.E. Use of the Journal Citation Reports for serials management in research libraries: an investigation of the effect of self-citation on journal rankings in library and information science and genetics [J]. *College and Research Libraries*, 2000, (61): 263-275.
 - 16 Fassoulaki, A., Paraskeva, A., Papilas, K [et al.]. Self-citations in six anaesthesia journals and their significance in determining the impact factor [J]. *British Journal of Anaesthesia*, 2000, 84(2): 266-269.
 - 17 Fassoulaki, A., Papilas, K., Paraskeva, A. [et al.]. Impact factor bias and proposed adjustments for its determination [J]. *Acta Anaesthesiologica Scandinavica*, 2002, 46(7): 902-905.
 - 18 Aksnes, D. W. A. Macro study of self-citation [J]. *Scientometrics*, 2003, 56(2): 235-246.
 - 19 Glanzel, W., Thijs, B., Schlemmer, B. A bibliometric approach to the role of author self-citations in scientific communication [J]. *Scientometrics*, 2004, 59(1): 63-77.
 - 20 Glanzel, W., Thijs, B. The influence of author self-citations on bibliometric macro indicators [J]. *Scientometrics*, 2004, 59(3): 281-310.
 - 21 潘文涛, 武夷山. 自引、他引: 说不尽的故事 [J]. *科技导报*, 2007, 25(24): 85.
 - 22 熊春茹. 关于科技论文参考文献自引问题的商榷 [J]. *编辑学报*, 2002, 14(6): 456-456.
 - 23 朱大明. 参考文献合理自引与不当自引的区分标准 [J]. *编辑学报*, 2004, 16(1): 76-76.
 - 24 蒋颖, 金碧辉, 刘筱敏. 期刊论文的作者合作度与合作作者的自引分析 [J]. *图书情报工作*, 2000, (12): 23-28.
 - 25 李建辉等. 自引对科技期刊影响因子作用的量化研究 [J]. *编辑学报*, 2007, 19(2): 154-157.
 - 26 骆柳宁. 《情报杂志》自引分析 [J]. *情报杂志*, 2005, 24(11): 125-127, 130.
 - 27 MacCain. Mapping authors in intellectual space: a technical overview [J]. *Journal of the American Society for Information Science and Technology*, 1990, 41(6): 433-443.
 - 28 White, H. D. Author cocitation analysis and Pearson's r [J]. *Journal of the American Society for Information Science and Technology*, 2003, 54(13): 1250-1259.
 - 29 Bensman, S. J. Pearson's r and author cocitation analysis—A commentary on the controversy [J]. *Journal of the American Society for Information Science and Technology*, 2004, 55(10): 935.
 - 30 Ahlgren, P., Jarneving, B., Rousseau, R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient [J]. *Journal of the American Society for Information Science and Technology*, 2003, 54(6): 550-560.
 - 31 White, H. D. Pathfinder Networks and Author Cocitation Analysis: A Remapping of Paradigmatic Information Scientists [J]. *Journal of the American Society for Information Science and Technology*, 2003, 54(5): 423-434.
 - 32 Leydesdorff, L. Similarity measures, author cocitation analysis, and information theory [J]. *Journal of the American Society for Information Science and Technology*, 2005, 56(7): 769-772.
 - 33 Leydesdorff, L., Bensman, S. Classification and powerlaws—The logarithmic transformation [J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(1): 1170-1186.
 - 34 Leydesdorff, L., Vaughan, L. Co-occurrence matrices and their applications in information science—Extending ACA to the Web environment [J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(12): 1616-1628.
 - 35 Schneider, J.W., P. Borlund. Matrix Comparison, Part 1: Motivation and Important Issues for Measuring the Resemblance Between Proximity Measures or Ordination Results [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(11): 1586-1595.

- 36 Waltman, L., Van Eck, N. J. Some comments on the question whether co-occurrence data should be normalized [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(11):1701-1703.
- 37 Leydesdorff, L. Should Co-Occurrence Data be Normalized: A Rejoinder [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(14):2411-2413.
- 38 王建芳, 冷伏海. 共引分析理论与实践进展[J]. *中国图书馆学报*, 2006, 32(1):85-88.
- 39 耿海英, 肖仙桃. 国外共引分析研究进展及发展趋势[J]. *情报杂志*, 2006, 25(12):68-69, 72.
- 40 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.
- 41 邱均平, 马瑞敏. 网络环境下 ACA 方法的应用研[J]. *图书情报工作*, 2008, 52(2):85-87
- 42 Adam, D. The counting house [J]. *Nature*, 2002, (415): 725 - 729.
- 43 Walter, G., Bloch, S., Hunt, G., Fisher, K. Counting on citations—A flawed way to measure quality [EB/OL]. http://www.mja.com.au/public/issues/178_06_170303/wal10537_fm.html 2009-12-16.
- 44 Lundberg, J. D. The “omnipotent” Science Citation Index impact factor [J]. *Medical Journal of Australia*, 2003, 178(6): 253 - 254.
- 45 Colquhoun, D. Challenging the tyranny of impact factors [J]. *Nature*, 2003, (423): 479.
- 46 Rogers, L. F. Impact factor—the numbers game [J]. *American Journal of Roentgenology*, 2001, 178(3): 541 - 542.
- 47 Lankhorst, G. J., Franchignoni, F. The ‘impact factor’—An explanation and its application to rehabilitation journals [J]. *Clinical Rehabilitation*, 2001, 15(2): 115 - 118.
- 48 Van Leeuwen, T. N., Moed, H.. Development and application of journal impact measures in the Dutch science system [J]. *Scientometrics*, 2002, 53(2):249 - 266.
- 49 Egghe, L., Rousseau, R. A general framework for relative impact indicators [J]. *Canadian Journal of Information and Library Science*, 2002, 27(1):29 - 48.
- 50 马廷灿, 黄可, 陈永波. 浅析影响因子的局限性及其一种修正[J]. *情报杂志*, 2007, 26 (4): 151- 152, 155.
- 51 Fenton, J. E., Brazier, H. [et al.]. The accuracy of citation and quotation in otolaryngology/head and neck surgery journals [J]. *Clinical Otolaryngology and Allied Sciences*, 2000, 25(1):40 - 44.
- 52 Anon. Errors in citation statistics [J]. *Nature*, 2002, 415 (6868):101.
- 53 Ho, Y.-S. Citation review of Lagergren kinetic rate equation on adsorption reactions [J]. *Scientometrics*, 2004, 59(1):171 - 177.
- 54 Wetterer, J. K. Quotation error, citation copying, and ant extinctions in Madeira [J]. *Scientometrics*, 2006, 67(3): 351 - 372.
- 55 罗时嘉, 洪蓉. 论社科著作的“硬伤”: 引文错误. *中国矿业大学学报*[J], 2001, 3(4):134-140.
- 56 胡锦涛. “伪引”的存在与消除[J]. *佳木斯大学社会科学学报*, 2003, 21(1):131-132.
- 57 郭小兰. 引文及其自引与伪引 [J]. *现代情报*, 2004, 24(4): 150-151.
- 58 戴勇, 张四新. 文献的不正当引文现象剖析[J]. *现代情报*, 2004, 24(10):202-203, 205.
- 59 戴勇, 张四新. 学术论文故意漏引现象分析[J]. *现代情报*, 2005, 25(4):48-50.
- 60 梁立明, 钟镇. 错引现象折射出的科学家群体引文失范行为—以 Nature 上一篇 19 万次高频引用论文的错引记录为例[J]. *自然辩证法研究*, 2007, 23(6):62-65.
- 61 邓迎, 侯汉清. 同一期刊的引文数据存在差值原因的分析—兼论期刊引文数据的规范化 [J]. *情报理论与实践*, 2008, 31(2):218-222.
- 62 Anthony Van Raan. Sleeping beauties in science [J]. *Scientometrics*, 2004, 59(3):461-466
- 63 Burrell, Q. L. Are “sleeping beauties” to be expected? [J]. *Scientometrics*, 2005, 65(3): 381 - 389.
- 64 M.V. Simkin, V.P. Roychowdhury. A mathematical theory of citing [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(11):1661-1673.
- 65 Hirsch, J. E. An index to quantify an individual’s scientific research output [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102 (46): 16569 - 16572.
- 66 White, H. D. Toward ego-centered citation analysis [A]. In B. Cronin & Atkins H. B. (Eds.), *The Web of knowledge*[C]. Medford, NJ: Information Today, Inc., 2000: 475-496.
- 67 White, H. D. Author -centered bibliometrics through CAMEOs—Characterizations automatically made and edited online [J]. *Scientometrics*, 2001, 51(3):607 - 637.
- 68 Bar-Ilan, J. An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes [J]. *Information Processing and Management*, 2006, 42 (6):1553 - 1566.
- 69 Cronin, B., Shaw, D. Identity-creators and image-makers—Using citation analysis and thick description to put authors in their place [J]. *Scientometrics*, 2002, 54(1): 34 - 49.
- 70 TH Huang, ML Huang. Visualization of Individual’s Knowledge by Analyzing the Citation Networks [J]. *Computer Graphics, Imaging and Visualisation*, 2007, 465-470.
- 71 郭永正. 国际科学合作个体中心网的中印比较[J]. *科学与科学技术管理*, 2008, 29(8):18-23.

(责任编辑:刘凤勤)