

引文分析存在的问题及其原因探究*

杨思洛

摘 要 目前引文分析已成为文献计量学的主要研究内容之一,但少有对其存在的问题进行分析。本文从引文分析理论(基础理论和引用动机)的不完善,引用过程中存在的不足,引文分析方法、工具和数据库的缺陷,引文分析应用与实践(科学评价和科学交流)的局限四个方面较系统地总结引文分析存在的问题及原因。目前应该以谨慎的态度看待引文分析,对其方法和结果作客观评价,在充分了解其优缺点的基础上进行完善与发展。图1。参考文献49。

关键词 引文分析 引文分析理论 引文分析方法 文献计量学

分类号 G353.21

ABSTRACT Citation analysis has become one of the main contents in bibliometrics, but few study has focused on its problems. The problems in citation analysis and their causes are summed up from four aspects in this paper: the weakness of citation theory, the deficiencies in the process of citation, the drawback of the citation analysis methods and databases, and the limitations in citation applications. At present, we should treat citation analysis with a cautious attitude, evaluate objectively the methods and the results obtained, and then improve it on the basis of fully understanding its merits and demerits. 1 fig. 49 refs.

KEY WORDS Citation analysis. Citation analysis theory. Citation analysis method. Bibliometrics.

CLASS NUMBER G353.21

自从16世纪后期论文引用制度形成以来,参考文献成为学术论文的第二特征,也是合理有效地进行科学交流的必要部分^[1]。20世纪中期,时任美国霍普金斯大学图书馆员的E. Garfield受1870年谢拔德引文的启发,开创了科学引文索引(SCI),掀开了引文分析的新篇章。随后“文献耦合”(Kessler, 1963),“科学引文网络”(Price, 1965),“同被引”(Small, 1973)、“引文可视化”(White, 1998)等相继提出。1998年,网络版数据库WoS(Web of Science)的问世进一步促进了引文分析的普及。在此过程中,对引文分析有效性、可靠性的争论与批判一直不断。从20世纪90年代开始,国内出现了关于引文分析反思性的文章,这是引文分析从介绍、引入,到应用发展的必然。进入21世纪,由于国

家科技经济的快速发展,科学评价及科技管理的需要,再加上信息可视化技术的突破、大规模引文数据库的出现,使得引文分析研究出现新的跨越式发展。但在人们热衷引文研究应用的同时,对引文理论和引文分析的不足相对研究较少。目前,从引文分析实践的惯势及其所支持的理论思维常态中走出来,对新环境下引文分析存在的问题作系统的探索,显得十分迫切和必要。

1 引文分析理论的不完善

1.1 基础理论

对引用行为的理论分析是进行引文研究的前提与基础。20世纪中期,一些学者意识到需

* 本文系教育部人文社科基金项目“网络引文的相关规律及其应用研究”(编号:09YJC870027)的研究成果之一。

要基础理论来解释作者的引用行为。1974年, Mulkey^[2]提出没有明显的证据表明引文反映了科学知识的影响过程。证据的缺乏使得对施引过程和原因知之甚少。此后,众多学者从不同方面对此问题进行了探索,特别是1998年科学计量学期刊 *Scientometrics* 以一期的篇幅,专题讨论引文分析理论。下面将对其中涉及的几个主要理论进行论述。

1.1.1 引文规范理论

1965年, Kaplan^[1]首先阐述了这一理论。他认为引用行为是通过教授对学生的口授和不同期刊的不断检验调整两方面而推广的;其主要功能是科学行为潜在准则的重申。引文规范理论基于以下假设:科学是一个由内部奖励和惩罚机制进行治理的规范机构;科学家为获得承认(以奖励或引文的形式)而交换信息(以出版物的形式)。引文是用于偿还智力债务的一种形式,“引文是学者付给同行的硬币工资”。后来许多学者企图从科学社会学的角度构建引文规范理论。但是该理论认为,当作者引用文献资料,则证明资料对他们有价值;而且,科学行为具有一致性,即他们引用的决定不会受功能不相关的特征影响,这些特征包括作者性别、种族、宗教和级别等;同时假定科学家们是无私的,不寻求通过奉承他人或自引来获得个人的利益;此外,该理论还认为科学家们对待他们自己的作品,持有和对待他人作品相同的怀疑态度^[3]。这些假设促进了引文分析的发展, Smith对假设作了进一步总结:文献被引用表示文献确实被作者利用;文献被引用是文献价值的反映(质量、意义与影响);最好的作品才会被引用;施引文献与被引文献之间是内容相关的;所有引文的价值相等^[4]。显然该理论是一种理想的状态,与现实情况存在着很大差距,也反映了这一理论的局限性。

1.1.2 社会结构理论

从社会结构角度关注科学交流和科学研究,社会结构领域学者认为科学的形成是协商过程的结果,在此过程中一方靠劝说的手段使其他人信服。从此角度,当作者引用文献时,他们通过组编以前的文献,说服读者相信他们观

点的合理性。大多情况下,作者常常仅仅是为了自身的需要而选择引用文献,就像推销和广告商品一样。White从两方面对此理论进行了总结:作者引用是为了达到自己目的而特意歪曲文献的本意;作者为获得认同不成比例地引用已有的权威作品^[5]。MacRoberts也认为引用的主要动机是说服而不是引文规范理论所宣称的“致谢”或“还债”^[6]。尽管此理论得到了许多学者的赞同,但也有众多学者从实证的角度加以反对。例如 White通过对文献被引频次的研究表明,引文呈正态分布,即引用名人和引用不知名人的著作都较少,大多数引文来自中等声誉作者的文献。这与结构主义者所宣称的作者为了说服而喜欢引用“大牌”作者的文献的说法矛盾。Moed和Garfield在系列论证基础上,也认为“说服不是主要的引用动机”^[7]。

1.1.3 心理过程理论

众多学者认为引用行为可当作一种心理认知的过程,相应地,引文理论的构建应该基于对个体引用者的研究,具体可通过访谈技术、放声思考法(thinking aloud)、行为模式记录等方法实现^[3]。Harter认为“引文活动是一个动态、复杂的认知过程”^[8],他提出在科学研究中,通过检索系统或其他途径获取文献是一个相关性判断的过程,引起了研究者认知的变化;研究过程中,参考的文献对作者造成影响(包括对研究的概念构架、问题的提出、研究方法及结果的解释);最后,这些参考过的文献与研究是特别相关的,它们被整合到呈现研究结果的论文参考文献列表中。参考文献列表以及特定的引文,告知读者引文与研究的历史相关性。虽然Harter的观点得到了一些后续研究者的支持,但也出现了很多反对者,批评他无视社会文化的影响,如认知能力、教育背景和所处物理环境的差异等。正如 Nicolaisen所评论:“Harter没有研究为什么最初文献被认为是相关的;当作者在后来的论文中参考这些文献,他们使用了什么标准。”因为缺乏广泛深入的研究,减少了Harter引用理论的说服力。

1.1.4 自反引文理论(the Reflexive Citation Theory)

一些学者关注于引文的表征特征,探究引文如何反映和代表科学,而不是为什么作者引用文献。Wouters 在其博士论文及系列著作中对这一观点进行了深入分析,他认为从作者引用行为角度来寻求解释引文理论将进入死胡同^[9],因此必须放弃这种追求,而是集中精力于引文的象征特点和指示能力。类似于信息语义学理论,他的理论基于对“参考”和“引文”是两种不同标志的解释。不同于“参考”,“引文”是无量纲的也是无意义的。仅仅在引文分析家的手中引文才得到意义,因此对作者引用行为的研究,有利于解释参考模式,而不是引用模式。但是,自反引文理论带有和语义信息学理论同样无法解决的问题,因为它无法处理误报(false positives),例如,在文献计量学领域,一般认为引文是衡量论文质量的指标,这种方法假定被引量 and 文献质量齐头并进,因而是线性相关的。引文分析的合法性研究也因此寻求科研质量和被引用的频次间表现一种线性关系,但许多研究表明,两者并非如此理想的关系^[3]。

1.1.5 引文进化理论

基于生物进化角度,1970 年以色列生物学家扎哈维发表“不利条件原理”(handicap principle),含义是动物和人类不是在做出最冒险、最过分的行为之余侥幸能兴旺,而正是因为有这类行为而兴旺。引文的引用行为也可从这个角度进行理论的探究,Nicolaisen 在其博士论文中对将此原理引入引文理论中进行了深入研究,但也受到了一些学者的质疑^[10]。

目前引文分析基础理论远未完善,影响最大的是引文规范理论和社会结构理论,但这两种理论恰好存在对立性,各有支持者,都有合理的一面,同时又有许多不足。形成此现象的原因主要有:一是由于理论提出的目的不同,角度不同,方法不同,对本质认识的不同,从而产生不同的看法;二是引文的形式及所处的外部环境在不断发展、变化,在不同的发展阶段,人们对于它的认识也是发展变化着的;三是由于引文现象的复杂性,研究者因各人所涉及的学科领域、所使用的杂志和论文类型、出版物的可获得性及技术因素的不同,都会影响到研究者的

认知与结论。

1.2 引用动机

引文理论往往建立在引用动机的基础上。引用动机是相当复杂的,它深深地影响作者的引用行为及引用结果。但引用行为动机的差异,使得引文分析时无法有效地控制各种因素,可行性因而大打折扣。目前更严重的问题是对引用动机也没有统一的归纳。20 世纪下半期开始,众多学者关注于此领域,Weinstock 系统归纳文献被引证的 15 种原因,这些引证动机对于科学发展、文献交流来讲,属于正常动机。但是,也存在一些与上述目的相反的动机。Thorne 归纳了 19 种不遵守学术规范和学术道德的行为:为阿谀某人而引证;为互相吹捧而带有偏见的引证;以自诩为目的的引证;为支持某一学术研究派别利益的不正当引证;迫于权威压力的引证等^[11]。Brooks 根据前人的研究,将引文动机分为 7 类^[12]:①新颖性,只引用最新的资料,以彰显自己搜集资料的新颖程度。②负面证据,引用反面的证据批评、否定、反驳他人的作品。③操作型信息,沿用其他研究者的研究方法、结果、参考工具等。④说服,引用文献说服他人。⑤正面评价,对于同领域的论文或相近的研究成果给予正面评价。⑥提醒,提醒读者注意背景数据、原始数据和最新资料。⑦社会认同,为了得到他人在学术上的认同而引用。Liu 在 1993 年绘出了引用行为的内在动机和外因因素在引用过程中的关系图,充分表现了引用过程中复杂的影响因素^[13]。1994 年, Baird 及 Oppenheim 提出作者引文的 17 种动机^[14]。

对引文动机研究缺乏一致性的分类主要原因是由于对作者引用动机的分析方法受到众多非议。目前,引用动机的分析方法主要分为两种:①背景与内容分析法;②问卷调查与访谈法。早期研究常用背景与内容分析法,主要通过研究者对特定论文集的参考文献进行分类,分析施引论文的内容达到对引文特征分析的目的。例如在 1975 年, Moravcsik 和 Murugesan 使用自建分类表,对发表在 1968—1972 年的《物理评论》上 30 篇论文的 706 个参考文献进行了分

类。而在网络环境下该方法也频繁使用,例如 Kousha 和 Thelwall^[15],还有 Vaughan 和 Shaw^[16] 对网络引文进行的系列分类分析。但此方法遭到质疑,一方面并没有统一的分类标准,另外研究归类时,主观性太强,施引者动机的多样性对分类也是一种挑战。从问卷调查与访谈法这一分析方法的角度来看,有学者提出从现象学的视角分析,即通过直接调查了解作者引用文献的动因。例如 Brooks 归纳了引用理论模型,总结出 7 种引用动机,对 26 位学者在近期发表论文的参考文献动机进行了问卷调查^[17]。但是以上分析都存在一个根本性问题:不能解释为什么被引文献在最初被作者发现是相关的。一个人往往无意识地或有意识地不承认引用特定文献和不引用其他文献的原因。因此,调查一个作者引用或不引用文献的动机,不能揭示作者已经施引的实际引用原因。由于引用行为因人而异,没有一定的标准,没法表明引用文献与被引用文献是否真有相关性。由以上两种分析方法得出的结论可信性都存在问题:因为这些研究的设计方案间的变化与差异很大,其结果很难具有可重复性;研究方法普遍存在不足;而且有证据表明,施引者存在引用动机的多样性及动机归纳的随意性^[18]。

2 引用过程中的不足

引文是作者引用过程的结果,但目前引用过程中存在许多不足。梁立明等在利用 ISI 数据库检索 *Nature* 的被引情况时发现,有一篇 1970 年发表的论文总被引频次超过 19 万次。但在这些引用该论文的记录中,出现了大量的错引,仅仅卷码和页码标注就有 518 种错引类型,4,133 条错引记录,错引率 2.13%。他们推测是由于科学家的引文失范行为,也就是引用者在没有阅读原始文献的情况下,复制了其他论文的参考文献,即间接引用或转引^[19]。Simkin 估计仅仅 20% 的引用者读了原文,并认为一篇高频引用论文的错引记录按照错引类型的被引频次从高到低排列,呈齐普夫分布^[20]。Prabha 对 19 个研究者的问卷调查表明,转引占

所有引用的 4%^[21]。Moed 和 Vriens 检验了来自 5 种学术期刊的 4,500 篇论文,被 25,000 篇论文所引用,在引文数据库中 10% 的引文在标题、作者姓名或页码上存在误差。

漏引在学术论文中普遍存在。在实际中,没有一位著者会列出全部参考过的书目,即使引用过的文献,对于著者的影响程度及发挥的作用明显不同。Macroberts 通过多个学科的实证分析表明漏引问题普遍存在,最严谨的学者也仅列出其实际参考文献的 64%。例如在 15 篇遗传学发展史方面的论文中,应该列出 719 篇参考文献才能完全包括其中所含已有出处的信息内容,但实际引文只列出 216 篇,仅占应引文献的 30%^[22]。在国内这一现象更为严重,2003 年我国科技论文中的 1,576 种统计源期刊,其中没有参考文献的论文就有 25,283 篇 (9.21%)^[23]。这么多科技学术论文居然可以没有一篇参考文献,这一现象曾被外国人认为是不可思议的。金标的调查也发现,在 12 种优秀图书馆学期刊中,有约 25% 的文章省略或没有文后参考文献;只有 17.5% 的引文作了标注,而且采用多种标注方式;参考文献用语使用相当混乱^[24]。另外,在国内这一问题还存在其他更为严重的现象:一些学术期刊公开声明,为了节省篇幅,刊物将来稿中的参考文献删去,或者不经作者同意擅自增删;有的刊物提出所有刊登论文必须引用该刊 1-2 篇论文;还有的期刊,为了进入核心期刊阵列,故意提高被引率。这些引用过程中的失范行为,对引文分析无疑是巨大的打击。

引用过程中产生的问题是多方面原因造成的。漏引问题产生的原因有可能是由于作者写作过程时间较长,无法再找到原文的出处,最终只好不引;有些作者外文水平较差,知识面窄,遗漏与论文相关的重要文献;某些作者不负责任,为图省力,简单列几篇,未一一列出全部引文;还有作者对引文不重视,不管相关程度如何,随便列出其中一部分;也有作者主观故意造成的因素,这种故意漏引则是作者学术道德的深层问题^[25]。引用过程的不足不只与作者的引用态度和行为习惯相关,文献的产生和出版过

程中各环节都可能对其造成影响。编辑对引文的处理方式过于草率、引文规格多样化、引文规范的过时等都是常见的现象。例如《GB/T7714-2005 文后参考文献著录规则》是推荐性国家标准,但是在各期刊中执行的普遍度不高;另外相关规范不明致使著录混乱;电子文献著录规范也过于笼统且自相矛盾^[26]。

3 引文分析方法与工具的缺陷

引文分析是文献计量学的重要组成部分,目前形成了整套方法、工具和测度指标体系,但很多方面一直受到学者的质疑。

3.1 引文分析方法

引文分析常见的方法可分为两大类:将引文款目作为独立计量单位的分析和将款目间的联系作为计量单位的分析^[27]。其中前者因为相对简单,较为成熟,但也存在分析方法的改进问题,如对影响因子的计算年限、最大引文年限、引文半衰期的数据统计方法等;而后者,因为可以得出更深层次的有意义结论,是目前引文分析的重点,但也存在众多缺陷和问题。例如,共引和耦合是引文分析法中的重要概念,共引分析法与耦合分析法是对学科内部结构与研究群体可视化的知识发现和信息聚类的重要方法,然而共被引与耦合究竟在多大程度上体现着两主体的相关性还是一个涉及较少的论题。所以共引或耦合分析究竟能够多大程度准确定量地描述主体间的结构与关系仍是未知。

对于引文分析方法存在缺陷的原因,一方面引文分析方法出现的时间较短,许多方面还有待于进一步深入探索,需要进一步完善,例如在引文分析过程中引入权重进行差别性计算。另外引文分析方法的机理存在不足,其计算方法随着引文的发展也需要进一步改进。以现在应用得最为广泛的共引分析方法为例,因为引用文献的原因多种多样,两篇论文可能出于完全不同的原因或从不同的角度引用同一篇早期文献,一篇可能是引用其方法,另一篇可能是引用其结果,那么这两篇文献在内容上的联系就

有可能是虚假的。同样,一篇论文也可能出于完全不同的目的同时引用两篇已有的文献。引文有些是发生在前言和篇名中;有些在正文中;有些发生在结论或讨论中。引文过程中有四种类型:点对点、点对面、面对点和面对面。在这些情况下,作者对原文的引用内容和程度是不同的。Edge D 以《为什么我不是一个共引研究者》为题,从6个方面对共引分析进行了强烈的反对与批评^[28]。

3.2 引文数据库

基于引文分析研究与评价的质量、合理性和可靠性很大程度上依赖于所用引文数据的精确性和全面性,不准确或不全面的数据往往造成不精确甚至错误的结果,而数据库的检索功能与响应速度对引文分析的效率也影响重大。在50多年前,加菲尔德建立SCI,此后的几十年里,该系列数据库作为唯一的综合性引文分析工具而被广泛使用。但一直以来,对SCI功能的质疑和缺点的批评不断。首先,收录的期刊比例很少,《科学引文索引扩展版》只收录了约6,000种重要期刊,约占各国现有科学期刊总数的3%;另外,所收录的期刊在国别地区、学科及语种等方面十分不均衡,例如对我国现行数千种期刊只收10余种,许多国家、许多学科、许多语种的较高水平的刊物被排斥在外;还有,选刊标准受到彼此相差很大的“影响因子”干扰,马太效应迭现;由于文献之间的引用与被引用关系极其复杂,仅仅依据SCI所作出的评价和结论往往有一定偏差^[29];而且在人文社会科学,学者引用的文献类型更多的是图书或报告等,通过论文数据库进行引文分析的结果将扭曲实际情况^[30]。另外,也存在数据著录格式(如人名和地名的不统一)和遗漏的问题。

在国内,引文索引库的研究虽然起步较晚,但成功地借鉴了SCI及其他数据库建设的经验,近年来取得了令人瞩目的成绩。出现了有一定特色的系列引文数据库,较常用的有中国学术期刊全文数据库、中国科学引文数据库、中文科技期刊引文数据库、中国社会科学引文索引。但与国外相比,国内引文库还有一定差距,引文

数据库间缺少必要的沟通与合作,尚存在不少问题。如数据产品质量不高、生产周期过长、不支持引文链接、引文产品影响小等^[31];在期刊选取、用户支持、系统功能、系统性能、系统规范性等方面存在诸多问题^[32]。

随着科学技术的进展,网络环境下出现了多种引文分析工具,如学科专业性引文数据库 CiteSeer(计算机科学)、SmealSearch(工商管理)和 RePEc(经济学);综合性有 Scopus 和 Google Scholar 等。尽管这些新出现的数据库比传统的 SCI 在某些方面有优势,如内容更全面、索引更深入或功能更强,在引用关系的基础上建立了全面的链接关系。但存在的问题也很普遍,不如传统引文数据库规范、系统、权威,不够稳定、适用。而且对同一目标,使用不同引文数据库分析,得出的分析结果有很大不同,甚至截然相反。目前缺乏对这些数据库的系统比较,不知道哪些数据库适应哪些领域或者适合什么特定应用的需要。正如 ISI 引文索引的诞生极大地促进了引文分析理论和方法的发展,网络引文库中新的数据和工具也会进一步促进信息计量学理论和方法的扩展、提高和创新。现在要研究的问题包括:如何将基于第一被引著者的引文分析研究扩展到所有被引著者、如何使用类似于 Google 的 PageRank 的更复杂而巧妙的算法进行基于引文分析的科研评估、如何利用大规模的数字信息研究引文网和知识流的结构和特点^[33]。

3.3 分析工具

引文分析往往涉及大规模数据,为提高分析效率出现了许多引文分析工具,专门用于信息计量分析的工具具有 Bibexcel、Citespace、Histcite 等,通用处理工具有 SPSS、SAS、Pajek、Ucinet 等。现在引文可视化是研究重点,但不同的可视化工具、不同的方法与过程都会产生不同的结果,而且不同软件产生的可视化图谱的侧重点也不同。不管是因为具体的软件,还是这些软件中的具体算法,或者说是因为不同形式及不同的角度,使得可视化结果不同,其效率和效果都存在。表面上看来较为客观的引文计量分析,实际上是十分主观的不太可信的结果。这

些软件工具对数据往往具有特定要求,很多只适合处理 WoS 题录,特别是缺少中文引文分析工具。而在引文分析工具中一般会用到数学算法。例如,在相似性测度时,传统共引分析中,文献共引分析代表人物 Small 主要采用 Salon 余弦测度或 Jaccard 系数测度文献间的相似性,而作者共引分析的代表人物 White 主要用 Pearson 相关系数进行作者间的相似性计算。目前学术界对此仍然没有达成共识。它们的数学机理不同会给计算结果造成显著影响,例如皮尔逊相关系数自身存在的问题明显地影响到引文分析结果的真实展示^[34]。

目前中文引文分析工具的缺乏,一方面与国内的引文数据库相关,很多中文引文库数据很不规范,在对其进行引文自动分析时存在困难,虽然有少数人尝试开发了软件工具,但没有像国外一样免费拿出供他人共享;另一方面,这也与国内对通用引文分析工具开发的不重视相关。

4 引文分析应用的局限

引文分析主要有以下应用领域:对科学家、出版物和科学机构进行定性和定量评估;模拟科学技术的历史发展;信息资料的查找与检索;预测学科结构与发展趋势^[35]。

4.1 科学评价方面

引文分析模式被认为能反映、评估论文或作者表现的影响力,但实际上存在许多问题。早在 20 世纪 60 年代,加菲尔德就告诫人们,仅用引文分析作为评价指标必须慎之又慎,并指出引文分析在评价中可能出现的混杂与疏漏^[36]，“盲目地作出被引频次最多的著者就该得诺贝尔奖金的结论是荒谬的……谁用引文统计在不同层次,尤其是在个人的层次上评价研究工作的绩效,谁就必须了解引文统计的微妙性和局限性”。2009 年 9 月 12 日,84 岁的加菲尔德博士访华期间表示:“到今天 SCI 的主要作用仍然是一个文献检索工具,而评估科技研究成果只是 SCI 的衍生功能之一。”“引用期刊影响因子来评估个人科研能力是

不合理的。”国内学者叶继元也表示：“应清醒地看到，引文是科学对话的一种方法，是作者认为对自己的研究有用的资料，并不是专门对研究质量的评价。”^[37]

另外，基于引文的科学评价有系列指标，长期使用但争议最多的是期刊影响因子。例如，2009年度《期刊引用报告》(*Journal Citation Reports*, 简称 JCR)中，《结晶学报 A 辑》的影响因子猛然从 2008 年的 2.051 升至 49.926，在 JCR 全部期刊中排第 2 位。引起《结晶学报 A 辑》急剧变化的直接原因，是该刊 2008 年所发表的《SHELX 简史》一文获得了很高的被引次数。《结晶学报 A 辑》影响因子即年指标突变，完全是一篇论文之功，这个典型事例也充分说明：受少数高被引论文左右，期刊影响因子高低不能准确反映期刊影响力大小^[38]。影响因子的主要问题有：易被人操纵、统计错误、不能跨学科比较、选源标准问题以及对非英文期刊不公平等；另外，影响因子建立在“所有引文重要性等价”为假设的基础上，实际上这是存在问题的，如在总被引频次相等的情况下，被 *Nature* 等引用的论文影响力显然会大于只被一些低水平期刊引用的论文^[39]。1995 年 H 指数被提出，后来出现对其完善的系列指数；最近又流行“特征因子”(Eigenfactor)。与传统影响因子相比，基于 PageRank 算法的特征因子具有明显的优点。但目前没有一种评价指标是完美的，特征因子也不例外：发量依赖，限于发文数量，规模小的期刊不会得到较高的特征因子分值；对于排名靠后的期刊特征因子的分值很低，使得期刊评价时的区分度不够；另外特征因子有和影响因子相同的一些缺陷，没有脱离原有量化学术评价方法的老路；它也继承了 PageRank 的不足^[40]。

引文分析在科学评价方面的应用存在问题的主要原因有：①文献被引用，不一定反映其重要性。一篇错误观点的论文，可能因为批评指正的需要，被引频次(数量)会很高；相反有些用而不引，作者在论文中不自觉地吸取了别人论文的某些观念，却没有将它当作引用文献。有些学科领域，有不引用论文的传统，引文很少。施引作者认为对研究“有用”的资料，并不是专门对论文

质量的评价，如何区分正面引用、负面引用和引文深度是十分棘手的问题^[41]。②省力法则。著者选用引文受可获得性的影响，选用参考文献以方便为准则，以占有为优先，引文并不一定是最好的。如 Soper 在 1976 年发现，在著者最新发表的论文中，所引用的文献绝大部分是个人藏书，小部分是所属部门或单位图书馆藏书，来自其他来源的文献很少^[42]。目前有许多研究表明，开放获取期刊比传统期刊、同一期刊上开放获取论文比未开放获取论文的引用次数多，也在一定程度上说明这个问题。③时间性问题。许多不朽之作在其发表后许多年，人们才注意并频繁引用。如哥白尼的《天体运行论》、爱因斯坦的《论运动物体的电动力学》和布拉德福的《专门学科的情报源》等，这就是所谓的“睡美人现象”。许多重大的科技进展被融汇为普遍承认的常识，本应很高的被引次数却没有持续下去，如爱因斯坦的相对论、有关 DNA 双螺旋结构的论文就是典型。这给引文评价功能的正确发挥造成了一定难度。④引用偏见。各种原因和目的的伪引、漏引，尤其是为拔高著者身份的自引、互引和为掩遮某种联系的漏引，都会给引文分析带来误差。⑤马太效应的存在。人们往往以“名著”、“权威”作为选择引文的标准，有的确是出于需要，也有的则是为了抬高自己论文的身价。一种期刊因为发表名人的文章而为众人所引用，以至引起连锁反应，结果其引文率很高。这种马太效应的作用，掩盖和影响文献引用的真实性，不能客观地评价论文的价值。如加拿大魁北克大学的 Lariviere 和 Gingras 在 WoS 中寻找重复发表的论文。他们把“重复发表”定义为：(a)两篇文章标题完全一样；(b)第一作者姓名完全一样；(c)两篇文章后面开列的参考文献数量完全一样。他们共找出 4,918 组重复发表的论文，其中 4,532 组论文发表在影响因子有异的刊物上。这些论文组中的两篇文章，多数发表在同一年。对这 4,532 组论文被引状况的对比分析表明，发表在较高影响因子刊物上的论文，要比发表在较低影响因子刊物上的同样论文的被引次数，平均高 1 倍。这说明，文章被引用反映的并非完全是论文的内在价值，还部分反映了发表论文的期刊的影响力和显

示度^[43]。⑥其他方面的问题。引用率依学科领域、语言、时间等的不同而有所不同。有些语言不被广泛使用、或是出现时间较晚,导致引用文献的数量比其他文献少很多;自引在每个作者的引用行为中占有的比例,因引用者的习惯、风格等的不同而表现出明显的差异;有些专门学科的研究领域规模较小,导致论文被引用的地位不平等;综述评论文章,往往被引次数较高,实际上综述文章的创新性和重要性与其高引用率不符^[44]。

4.2 科学交流方面

科学交流是知识创新的前提和基础,科学文献则是科学交流的重要载体和传递工具。文献之间引用关系较客观真实地表征了科学的累积性、连续性和继承性,可以描述、解释、预测和评价科学交流行为活动。引文实际上是作者群体在学术共同体认同的学术理念和规则影响下产生的、可以考察分析的一种学术交流现象。众多学者从引文分析视角研究科学交流活动。Borgman 认为文献计量学为研究科学交流的结构和过程,提供了一整套有力的手段和方法,其中尤以引文分析最突出^[45]。Harter 通过引文分析研究了电子期刊在科学交流过程中的影响力^[46]。但是,其中引文分析存在的问题也不能不引起我们的重视,以引文分析为视角的科学交流研究有时把实际情况给扭曲了。MacRoberts 在最近的研究中,用 20 个生物地理科学的案例说明:在作者研究过程中使用了大量的先前成果,但大部分并没有被引用,这有作者的主观原因也有受版面限制等其他因素;数据类的论文比理论分析类论文具有明显更少的被引用现象;但是没有被引用,并不意味着作者没有使用过;而且引用的论著并不比没引用的论著对作者的影响更大,即引用的频率并不能表征其在科学交流中的重要度^[47]。事实上,在科技领域,一旦一个观点被充分广泛地传播,并为众人所知晓和公认后,再引用原始叙述已不再需要了,这时包含有这个广为人知观点的文献即成为一条含蓄的引文,这种现象在科技文献中比较普遍,它也无法从引文角度分析。

在科学交流中,特别是在网络环境下,存在

着多种交流方式,引文视角仅仅是其中的一个方面。20 世纪中期,普赖斯就认为学者 80% 的信息是通过非正式渠道交流获取的,科学研究重要的信息往往是通过会议或者面谈等其他交流方式获取的。1991 年,Hicks 通过地理学知识形象地描述了引文在科学交流中的地位,“引文仅仅是巨大岩石层中的一个细小的区域,它影响非常有限(不到 10%),也十分不具代表性。只是通过 ISI 的盲目推崇,引文分析变成合适的市场化商品,从而成为了极吸引人的可用岩石层”^[48]。在 Hicks 的研究基础上,MacRoberts 绘出了四个层次组成的引文分布地形图,在此基础上经过少量修改,我们得到图 1,较形象地显示了从引文角度分析科学交流的局限。

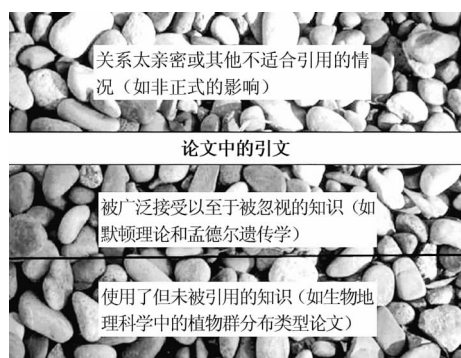


图 1 引文分布地形图

5 结语

科学文献的引证与被引证,是科学发展规律的表现,体现了科学知识的累积性、连续性和继承性。引文分析是图书情报学十分关键的研究内容,由于在科学评价、信息检索和知识发现等方面的重要应用,成为图书情报学影响其他学科领域的重要方面。从理论、方法与应用等方面对引文分析存在的问题进行了较为全面系统的总结,这使我们认识到,引文分析仍处于十分不完善的阶段,再加上引用对象、过程和引文库的不断变化,使得这一问题更加复杂。特别是在国内,引文分析被过度热心而盲目使用、滥用现象较为普遍,导致大量低水平重复性的成果,这是我们进行引文分析必须正视的问题,而引起这些

问题的原因和相应对策也值得我们深入思考。

目前我们应该以谨慎的态度看待引文分析,对其方法和结果应有客观评价,在充分了解其优缺点基础上进行完善与发展。对引文分析存在的问题从以下几方面进行改进:第一,引文数据库功能进一步加强,为大规模的引文分析奠定基础,也为监督、验证引文分析的正确性提供依据,甚至可以提供引文分析的自动统计处理功能,大多数计量分析方面的功能不需要人工承担。例如,将目前主要用于检索和分析的引文索引库改造为专门用于评价的数据库,这需要对引文进行深度加工,研制出信息和知识含量更大的数据库^[49]。第二,计算机网络技术的深入发展,为新研究方法的提出、可视化软件的研制提供了条件。可以说是计算机的出现,才使得引文分析兴起,随着相关技术的发展,引文分析将进一步完善。随着互联网在学术传播系统的应用及学术信息电子化的发展,新兴的引文分析工具与方法不断涌现。异构异源引文数据互链分析越来越多,引文计量粒度越来越小,可实现引文分析结果可视化的自动智能生成。第三,随着社会文明的进步,科学伦理、科学道德的健全与法制化,著录规范化和标准化的加强,学术共同体不断研讨和交流,取得共识,使得引文数据的信度和效度提高。第四,由于引文分析影响的广泛性,全新的学术传播模式也在不断出现,越来越多的学科都在研究这个问题,从哲学、社会学、心理学、文献学等角度进行深入探讨,引文分析的发展会越来越科学。

参考文献:

- [1] Kaplan N. The norms of citation behavior; Prolegomena to the footnote [J]. *American Documentation*, 1965, 16(3): 179-184.
- [2] Mulkay M J. Methodology in the sociology of science: Some reflections on the study of radio astronomy [J]. *Social Science Information*, 1974, 13(2): 107-119.
- [3] Nicolaisen J. Citation analysis [J]. *Annual Review of Information Science and Technology*, 2007 (13): 609-641.
- [4] Smith L C. Citation analysis [J]. *Library Trends*, 1981(30): 83-106.
- [5] White H D. Reward, persuasion, and the Sokal hoax: A study in citation identities [J]. *Scientometrics*, 2004, 60(1): 93-120.
- [6] Mac Roberts M H, Mac Roberts B R. Testing the Ortega hypothesis: Facts and artifacts [J]. *Scientometrics*, 1987(12): 293-295.
- [7] Moed H F, Garfield E. Basic scientists cite proportionally fewer "authoritative" references as their bibliographies become shorter [C]//*Proceedings of the 9th International Conference on Scientometrics and Informetrics*, 2003, 190-196.
- [8] Harter S P. Psychological relevance and information science [J]. *Journal of the American Society for Information Science*, 1992, 43(9): 602-615.
- [9] Wouters. The citation culture [OL/M]. [2010-06-09]. www.garfield.library.upenn.edu/wouters/wouters.pdf.
- [10] Nicolaisen J. Social behavior and scientific practice: Missing pieces of the citation puzzle [OL/J]. [2010-06-09]. http://vip.db.dk/dbi/sampling/phd_dk.htm.
- [11] 邱均平. 信息计量学[M]. 武汉: 武汉大学出版社, 2007: 20-56.
- [12] Brooks T A. Evidence of complies citer motivations [J]. *Journal of the American Society for Information Science*, 1986, 37(1): 34-36.
- [13] Liu M. A study of citing motivation of Chinese scientists [J]. *Journal of Information Science*, 1993 (19): 13-23.
- [14] Baird L M, Oppenheim C. Do Citations Matter? [J]. *Journal of Information Science*, 1994, 20 (1): 12-13.
- [15] Kousha K, Thelwall M. How is science cited on the web? A classification of Google unique web citations [J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (11): 1631-1644.
- [16] Vaughan L, Shaw D. Web citation data for impact assessment: A comparison of four science disciplines [J]. *Journal of the American Society for Information Science and Technology*, 2005, 56 (10): 1075-1087.
- [17] Brooks T A. Private acts and public objects: An investigation of citer motivations [J]. *Journal of the American Society for Information Science*, 1985, 36(4): 223-229.
- [18] Bornmann L, Daniel H D. What do citation counts measure? A review of studies on citing behavior [J]. *Journal of Documentation*, 2008, 64(1): 45-80.

- [19] 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(6):62-65.
- [20] Simkin M, Roychowdhury V P. Read before you cite [J]. *Complex Systems*, 2003(14):269-274.
- [21] Prabha C G. Some aspects of citation behavior: A pilot study in business administration[J]. *Journal of the American Society for Information Science*, 1983, 34(3):202-206.
- [22] MacRoberts M H, MacRoberts B R. Problems of citation analysis: A critical review [J]. *Journal of the American Society for Information Science*, 1989, 40(5):342-349.
- [23] 中国科学技术信息研究所. 2003 年度中国科技期刊论文统计与分析[M]. 北京:科学技术文献出版社,2005:121.
- [24] 金标. 12 种优秀图书馆学期刊的引文规范性调查[J]. 图书情报工作,2000(9):88-89.
- [25] 戴泳,张四新. 学术论文故意漏引现象分析[J]. 现代情报,2005(4):48-50.
- [26] 周丽萍. 学术期刊参考文献著录规范研究盲点分析[J]. 浙江海洋学院学报(人文科学版), 2009(2):115-118.
- [27] 丁学东. 文献计量学基础[M]. 北京:北京大学出版社,1993:198.
- [28] Edge D. Why I am not a co-citationist [J]. *Society for Social Studies of Science Newsletter*, 1977 (2):13-19.
- [29] 李玉进. 引文分析存在问题的再认识[J]. 天津外国语学院学报,2001(3):60-62.
- [30] Larivière V. The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities [J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(8):1997-1004.
- [31] 郑德俊,叶继元. 基于合作模式的引文数据库发展策略[J]. 大学图书馆学报,2005(1):79-83.
- [32] 李志鹏. 中文社会科学引文索引的不足与改进[J]. 情报资料工作,2005(5):108-111.
- [33] 赵党志. 信息计量学与网络计量学[M/OL]. [2010-07-24]. http://www.ualberta.ca/~dzhao/DangzhiZhao_Chapter10preprint.pdf.
- [34] 邱均平,李晔君,李江. 共链分析的缺陷及其解决方案研究[J]. 情报理论与实践,2008(2):140-143.
- [35] Zunde P. Structural models of complex information sources [J]. *Information Storage and Retrieval*, 1971(7):1-18.
- [36] 侯汉清.《科学引文索引》与科技评价工作——《引文索引法的理论及应用》译后记[J]. 江西图书馆学刊,2005(2):3-4.
- [37] 叶继元. 首届人文社会科学评价学术研讨会综述[J]. 学术界,2009(4):301-304.
- [38] 赵基明. 影响因子突变的提醒[N]. 科学时报,2010-06-30(A1).
- [39] 任胜利. 特征因子(Eigenfactor):基于引证网络分析期刊和论文的重要性[J]. 中国科技期刊研究, 2009, 20(3):415-418.
- [40] 窦曦曦,祁延莉. 特征因子与论文影响力指标初探[J]. 大学图书馆学报,2009(6):57-62.
- [41] 叶继元. 如何看待引文数据中的负面引用[N]. 光明日报,2007-06-12(11).
- [42] Soper. Characteristics and use of personal collections [J]. *Library Quarterly*, 1976, 46(4):397-415.
- [43] Larivière V, Gingras Y. The impact factor's matthew effect: A natural experiment in bibliometrics [J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(2):424-427.
- [44] 谢文兵. 综述文章高引用率:盛名之下其实难副 [OL]. [2010-06-09]. <http://news.sciencenet.cn/htmlnews/2010/4/230822.shtm>.
- [45] Borgman C, Furner J. Scholarly communication and bibliometrics [J]. *Annual Review of Information Science and Technology*, 2005, 36:3-72.
- [46] Harter S P. The impact of electronic journals on scholarly communication: A citation analysis [J]. *the Public-access Computer Systems Review*, 1996,7(5):5-34.
- [47] MacRoberts M H, MacRoberts B R. Problems of citation analysis: A study of uncited and seldom-cited influences [J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(1):1-12.
- [48] Hicks D, Potter J. Sociology of scientific knowledge: A reflexive citation analysis of science disciplines and disciplining science [J]. *Social Studies of Science*, 1991(21):459-501.
- [49] 叶继元. 引文的本质及其学术评价功能辨析[J]. 中国图书馆学报,2010(1):35-39.
- 杨思洛** 湘潭大学公共管理学院知识资源管理系讲师,武汉大学信息管理学院博士研究生。
通讯地址:湖南湘潭大学公共管理学院。邮编:411105。
(收稿日期:2010-06-09;修回日期:2010-07-25)