

基于文献计量学的民族学与文化学学科半衰期研究 *

张迪 赵越 邓三鸿

(南京大学信息管理学院 江苏南京 210023)

摘要: 半衰期是反映文献老化现象的计量指标,本文以半衰期作为研究方法,对我国民族学与文化学各子学科文献老化情况进行了计量与分析。论文通过检索 CNKI 获取本学科相关的文献,对民族学与文化学各个子学科的引用半衰期和被引半衰期进行了对比分析。

关键词: 学科半衰期; 被引半衰期; 引用半衰期; 民族学与文化学

中图分类号: G250.252 文献标识码: A DOI: 10.11968/tsygb.1003-6938.2015043

Study of the Academic Half-life of Ethnology and Culture Science

Abstract Half-life reflects the aging of literature. The aging of literature in the branches of ethnology and culture science are analyzed quantitatively in this paper. The data are acquired from CNKI and citing half-life and cited half-life are compared and analyzed.

Key words academic half-life; cited half-life; citing half-life; ethnology and culture science

1 引言

文献老化程度研究是文献计量学的重要研究方向,而半衰期则是描述文献老化情况的重要指标。文献老化的研究最早由美国纽约大学的 Gosnell 博士在《大学图书馆中文献老化问题》一文中正式提出^[1],1958 年英国学者 Bernal 在《科技情报传递:用户分析》一文中首次提出用“半衰期”来表征文献老化速度,1960 年 Burton 和 Kebler 在《一些科技文献的“半衰期”》一文中提出用“中值引文年限”来替换“半衰期”概念^[2],也被称为共时半衰期。随之出现了大量研究文献半衰期来揭示文献老化规律的文章。我国学者罗式胜对期刊引用半衰期和期刊被引半衰期的概念、计算方法、意义等进行了系统介绍^[3-5]。白云对 CSSCI 收录的来源期刊发文及引用信息进行研究,考察了我国人文社会科学期刊被引半衰期,得出结论认为被引半衰期与期刊的学术影响度没有明显的相关性而与学科的老化密切相关^[6]。在民族与文化学学科方面,华涛^[7]、邓三鸿^[8]、鞠秀芳^[9]等也对本学科的期刊进行分析,给后来者提供了良好的研究基础。

目前,对于期刊的半衰期研究已相对成熟,被引半衰期已成为评价期刊学术质量的一项重要指标。但综合以往的半衰期研究会发现,大多数的研究只建立在期刊半衰期的基础上,这些研究尚未深入讨论半衰期与学科之间的关系。本文将在以往研究基础上,以学科分类为基础,获取 CNKI 数据库中民族与文化学学科的引用和被引数据,从学科及各子学科引用半衰期、被引半衰期指标的时序变化角度,深入考察民族与文化学学科及各分支的文献老化现状。

2 数据来源

本文数据来源于 CNKI《中国学术期刊网络出版总库》、《中国博士学位论文全文数据库》和《中国优秀硕士学位论文全文数据库》等 10 多个数据库。引用数据包含所有来源数据库的参考文献,涉及期刊论文、学位论文、会议论文、图书、专利、标准、报纸文章等类型文献,被引数据不包括图书与外文。民族与文化学学科是一个横跨《中图法》众多分类的学科,其文章也发表在众多学科的期刊中,这给收集数据提供了挑战。在检索时,根据《中华人民

* 本文系国家社科基金项目“以创新和质量为导向的哲学社会科学评价体系实证研究”(项目编号:12BTQ037)研究成果之一。

收稿日期:2015-01-10;责任编辑:刘全根

共和国学科分类与代码简表》(国家标准 GBT 13745-2009)上所列出的民族和文化学学科各子学科,找到相应的中图分类号,由于分类表与中图法分类有很大差别,在细分子学科中图分类号的时候会有出入,笔者在数据统计过程中严格按照《中华人民共和国学科分类与代码简表》的分类要求,把类似《改革开放 30 年新疆民族关系问题研究综述》的文章分入“民族问题理论”一类中而没有分入“新疆民族研究”中,并且没有重复计算,力求最大可能得到有效数据(见表 1)。

表 1 2010~2012 年民族学与文化学学科发表的文章数

学科代码	学科名称	文章数		
		2010 年	2011 年	2012 年
85010	民族问题理论	2933	3003	3132
85020	民族史学	1166	1007	834
85030	蒙古学	7	4	13
85040	藏学	0	1	0
85042	新疆民族研究	0	0	0
85050	文化人类学与民俗学	1075	1018	1202
85060	世界民族研究	972	935	1177
85070	文化学	592	634	777
85099	民族和文化学其他学科	506	466	455

对比各子学科 2010~2012 年度的发文数,可以发现,蒙古学、藏学、新疆民族研究这几个子学科三年的发文数都非常低,文章数量太少难以准确计算学科半衰期,所以本文的研究将舍弃这些数据量很小的子学科及中图分类法中未列出的部分二、三级学科。

3 民族学与文化学学科半衰期的计算与分析

3.1 民族学与文化学学科引用半衰期的计算与分析

3.1.1 学科引用半衰期的计算

某学科的引用半衰期是指该学科在统计当年所引用的全部文献中,按照时间划分,较新文献的引用数达到总引用数 50%时的年数^[4]。具体计算方法是:统计某一年内所有发表文献的参考文献,按照出版年的降序排列,统计同一出版年文献数占所有文献数的百分比,然后计算每个出版年的累计百分比,在计算过程中,我们把计算当年记为第 0 年。公式如下:

$$HL = X + (50\% - A) / (B - A) \quad (1)$$

其中, X 为累积百分比最接近 50%的那年经过的年数; A 为累积百分比最接近 50%的那年的累积百分比; B 为累积百分比刚超过 50%的那年的累积百分比。

本文以 2012 年文化学子学科的数据为例来说明如何计算学科引用半衰期(见表 2)。

表 2 文化学学科 2012 年所引参考文献相关数据

出版年代	2012 年引用该年次数	累积次数	累积百分比
2008~2012	970	970	17.13
2007	263	1233	21.77
2006	296	1529	27.00
2005	327	1856	32.77
2004	360	2216	39.13
2003	295	2511	44.34
2002	255	2766	48.84
2001	267	3033	53.56
2000	243	3276	57.85
1900~1999	2387	5663	100

最接近 50%的引用累计百分比是在 2002 年达到的 48.84%,距 2012 年为 10 年,则对应公式(1)中参数的值分别为: X=10, A=48.84%, B=53.56%, 由此可计算出文化学子学科引用半衰期为:

$$HL = 10 + \frac{50\% - 48.84\%}{53.56\% - 48.84\%} = 10.25$$

依次本文可统计计算得出民族学和文化学各子学科的半衰期(见表 3)。

表 3 民族学与文化学学科 2010~2012 年引用半衰期

学科代码	学科名称	引用半衰期(年)		
		2010	2011	2012
85010	民族问题理论	7.14	7.91	7.55
85020	民族史学	11.07	11.26	11.18
85050	文化人类学与民俗学	9.95	9.36	10.47
85060	世界民族研究	6.42	6.28	6.60
85070	文化学	9.71	9.07	10.25
85099	民族学和文化学其他学科	9.45	11.40	9.49
平 均		8.96	9.21	9.26

3.1.2 学科引用半衰期的分析

可以看出,民族学与文化学学科的各个研究领域的引用半衰期都较长,都在 6 年以上,“民族史学”更是在 11 年以上,相比较于管理学学科 5.8 年的引用半衰期长了不少,这些都是因为本学科的学科性质较偏理论,学者们在研究的时候需要比较多地参考成熟的理论成果,也需要借助这些成果来印证其理论研究,所以引用半衰期较大。而且笔者在引用数据中发现各个子学科皆引用了大量的图书,这也使得引用半衰期较大。

观察发展趋势也可以发现,民族学与文化学各个子学科的半衰期相对稳定,波动在适当的范围内。唯有学科代码为 85099 的“民族学和文化学其他学科”在 2011 年的时候波动太大,对其数据进行分析可以发现,该学科中大部分的文献是属于没有参考文献的人物传记类文献(其中 2010 年共 506 篇文章,329 篇为传记,占 65.02%;2011 年共 466 篇文章,291 篇为传记,占 62.45%;2012 年共 455 篇文章,232 篇为传记,占 50.99%),这样就造成了少量的文献有大量的引用,那么这些文献的研究课题的轻微改变就会造成学科半衰期的大幅波动。

2010~2012 年民族学与文化学学科的平均半衰期分别为 8.96 年、9.21 年、9.26 年,可以看出除了学科代码为 85020 的“民族史学”与 85060 的“世界民族研究”,其余的学科引用半衰期都接近平均值。民族史学作为一个研究民族历史的子学科,其学科性质决定了该学科的经典文献利用率较高,半衰期较长。目前学术界普遍认为,“世界民族研究”主要侧重于国外民族政治理论、民族关系、民族问题、民族政策等民族政治方面的研究^[10]。那么这门学科的研究和现在的世界形势有关,世界上冲突不断,针对每个冲突来由的研究,对于每一个新的民族政策的研究等等导致了这门学科的引用半衰期相对较短。

3.2 民族学与文化学学科被引半衰期的计算与分析

3.2.1 学科被引半衰期的计算

被引半衰期是指某学科的论文在某年被引用的全部次数中,较新的一半被引论文发表的时间跨度,从被引半衰期可以看出该学科的文献对整个学科领域的影响情况^[3]。其计算方法与学科引用半衰期类似:

$$HL_2 = X + (50\% - A) / (B - A) \quad (2)$$

同样,我们以 2012 年文化学子学科的数据为例来说明如何计算学科被引半衰期(数据见表 4)。

表 4 文化学学科 2012 年被引参考文献相关数据

出版年代	2012 年被引次数	累积次数	累积百分比
2008~2012	1110	1110	32.80
2007	452	1562	46.14
2006	337	1899	56.10
2005	157	2056	60.74
1990~2004	1329	3385	100

最接近 50% 的被引累积百分比是在 2007 年达到的 46.14%,距 2012 年为 5 年,则对应公式(2)中参数的值分别为: $X=5$, $A=46.14\%$, $B=56.10\%$,由此可计算出文化学子

学科被引半衰期为:

$$HL_2 = 5 + \frac{50\% - 46.14\%}{56.10\% - 46.14\%} = 5.39$$

本文经过统计计算得出民族学与文化学各子学科的被引半衰期(见表 5)。

表 5 民族学与文化学各子学科 2010~2012 年被引半衰期

学科代码	学科名称	被引半衰期(年)		
		2010	2011	2012
85010	民族问题理论	4.33	4.50	4.82
85020	民族史学	5.33	5.58	6.61
85050	文化人类学与民俗学	4.12	4.64	4.92
85060	世界民族研究	3.52	3.74	4.20
85070	文化学	3.69	4.98	5.39
85099	民族学和文化学其他学科	4.07	4.60	4.55

3.2.2 学科被引半衰期的分析

可以发现,民族学与文化学学科各个子学科的被引半衰期在逐年增加,年平均半衰期增长最高达到 11.90%,说明该学科的文献扩散速度在减慢,学科研究逐渐走向成熟,学科影响力正在不断加强,文献老化速度在减慢。

同样可以发现,“民族史学”子学科的被引半衰期最高,这是因为学者在研究时更喜欢引用该学科较为成熟、历经时间检验的文献,故而该学科的半衰期较长。而“世界民族研究”相对与时俱进的探索对于其他学科有一定的现实意义,所以学者们在做研究时更喜欢引用该学科较新的文献,所以该学科的被引半衰期较短。

3.3 民族学与文化学学科引用半衰期和被引半衰期的比较

学科引用半衰期和学科被引半衰期都是反映学科发展的重要指标,两者既联系紧密,又有一定的区别。故本文对比和分析了民族学与文化学学科各个研究领域引用半衰期和被引半衰期(2010~2012 年民族学与文化学学科的引用半衰期和被引半衰期数据对比见表 6),试图发现其差异以及联系,并探讨产生差异和联系的原因。

可以发现,民族学与文化学学科各个子学科引用半衰期均要大于被引半衰期,仔细分析发现产生这个现象的一个重要原因是因为数据源的问题,数据来源于 CNKI 数据库,然而其被引库中并不包括图书信息和外文文献,而在引用文献中却有很多的图书和外文文献。图书著作的生命周期较一般的期刊论文长,一般发表很多年后仍然会被引用,外文文献则是由于语言等方面原因,生命周期较一般的文献更长,这在一定程度上使得引用半

表6 民族学与文化学 2010~2012 年
被引半衰期与引用半衰期对比

学科	2010 年		2011 年		2012 年	
	引用半衰期	被引半衰期	引用半衰期	被引半衰期	引用半衰期	被引半衰期
85010	7.14	4.33	7.91	4.50	7.55	4.82
85020	11.07	5.33	11.26	5.58	11.18	6.61
85050	9.95	4.12	9.36	4.64	10.47	4.92
85060	6.42	3.52	6.28	3.74	6.60	4.20
85070	9.71	3.69	9.07	4.98	10.25	5.39
85099	9.45	4.07	11.40	4.60	9.49	4.55

衰期被拉长。另外一个可能的原因是,因为民族学与文化学学科是一个横跨《中图法》众多分类的学科,其文章也发表在众多学科的期刊中,学科内和学科外的引用与被引不均等造成了这样的情况。民族学与文化学学科的学科性质决定了它会去引用其他学科相对成熟、年代较远的成果,而其他学科引用本学科则偏向于引用学科热点和最新研究成果,这也导致了引用半衰期比被引半衰期要长。

学科的引用半衰期和被引半衰期都是表征文献老化规律的指标,但是两者在研究对象、半衰期的长短意义上却完全不同。作者的引用是一种主观行为,而被引是一种客观的结果。笔者发现,引用半衰期的趋势波动对于被引半衰期的趋势基本没有影响,如“民族学与文化学其他学科”的引用半衰期在 2010 年时出现波动,而被引半衰期则始终平稳。产生这种情况的原因是学科内外引用被引的不一致性以及 CNKI 的被引库缺失了部分数据。

对于一个学科而言,其引用和被引大多发生在该学科内的文献中,一篇文献引用的同时也产生了相对应文献的被引行为。一个相对保守的学科,这个学科的被引数值理论上是无限接近于其学科引用数值的。正是因为学科间的引用关系、互引关系以及单项引用关系,才会出现偏差,但是一个学科有一定的学科范围限定,不可能出现偏离该学科大幅引用其余学科的情况,理论上偏差不会太大。一个学科某年的引用半衰期变长,则说明该学科发表的文献较多引用更早的文献,那么这一年的被引数据中较早的一定会更多,被引半衰期也会变长。所以从本质上看,学科引用半衰期和被引半衰期存在着紧密的联系。

为了进一步认证笔者的想法,笔者在原来的基础上将参考文献中的图书、外文数据和中文期刊论文的数据进行了分离,计算出中文期刊,学位论文等中文学术资源

的引用半衰期(见表 7)。

表7 民族学与文化学文献 2010~2012 年
被引半衰期与引用半衰期(新)对比

学科	2010 年		2011 年		2012 年	
	引用半衰期	被引半衰期	引用半衰期	被引半衰期	引用半衰期	被引半衰期
85010	4.54	4.33	4.57	4.50	5.07	4.82
85020	8.18	5.33	8.19	5.58	8.52	6.61
85050	4.70	4.12	5.25	4.64	5.80	4.92
85060	4.80	3.52	4.20	3.74	5.13	4.20
85070	5.20	3.69	4.66	4.98	6.38	5.39
85099	4.74	4.07	4.87	4.60	5.18	4.55

从数据来看,大体上的发现是学科的引用半衰期和被引半衰期具有一定的正相关性,随着引用半衰期的增长,被引半衰期也随之增长。

4 结语

本文按照《中华人民共和国学科分类与代码简表》分类表,以 CNKI 收录的引用与被引信息为数据源,观测 2010 年至 2012 年的引用半衰期和被引半衰期的变化,对我国民族学与文化学学科的文献老化情况进行了统计分析。得出民族学与文化学学科的引用半衰期约为 8.0 年,被引半衰期约为 6.0 年。引用半衰期大于被引半衰期,引用半衰期较稳定但普遍较高,同时被引半衰期在缓慢增加,说明该学科整体老化较慢。在深入探讨各子学科的时候发现,一个学科越偏向于理论,其半衰期越大,这说明这些领域知识更新慢,老化较慢;一个学科越是偏向于实证和热点,其半衰期越小,说明这些领域知识主题变化频繁,更新较快,老化较快。某个学科可以通过与其他学科的交流产生新的理论方法,这样是保持一个学科活力的重要手段,这种相互引用关系也会影响文献老化速度,而通过文献半衰期研究者可以轻易得出其中的规律,对学科的发展产生积极的影响。

参考文献:

- [1]Gosnell C F. The Rate of Obsolescence in College Library Book Collections, as Determined by an Analysis of Three Select Lists of Books for College Libraries [D]. New York:New York University, School of Education, 1943.
- [2]Burton R E, Kebler R W. The (下转第 98 页)

引用半衰期。其被引半衰期相较引用半衰期稍短,但也将近10年,也能反映历史学论文的持续利用时间较长,对经典文献关注度高,学科发展脱离了初创阶段的高增长新引用,文献有了历史的积淀,参照价值更高,对其他学科的影响力更大,能反映人文科学基础性的学科特质。被引半衰期短于引用半衰期的原因一是由于其客观的基础学科性质,二是数据源本身不规范及资源缺失引起。而历史学各个子学科中综合引用和被引半衰期的变化模式可大致分为成熟型、过渡型以及波动型。成熟型适宜用半衰期进行评价,过渡型则需结合其他定性分析,而对于波动型,半衰期数值的意义则有待商榷。

参考文献:

- [1] Burton, R. E, Kebler, R. W. Half-life of Some Scientific & Technical Literatures[J]. American Documentation, 1960 (11):18-22.
- [2] 罗式胜.文献半衰期的类型及其应用[J].情报学报, 1997, 16(1):62-67.
- [3] 郭红梅, 邵艳娥, 何钦成.文献半衰期研究现状[J]. 医学信息学杂志, 2011, 32(1):56-59.
- [4] 汪继南. 社会科学各学科引文的高峰值调查及引文评价的时段性要求[J]. 情报资料工作, 2005(3):81-82.
- [5] 白云.中国人文社会科学期刊被引半衰期分析研究[J]. 云南师范大学学报(哲学社会科学版), 2006, 38(4): 127-130.
- [6] 姜春林, 刘则渊. 历时态角度看人文社会科学引文峰值及其引文评价的时间选择[J].情报科学, 2006, 24(9): 1343-1347.
- [7] 游毅, 索传军, 钟晶晶. 科技期刊长期老化趋势与影响因素实证研究——基于4门学科期刊30年引文数据的统计分析[J]. 图书情报工作, 2011, 55(12):140-144.
- [8] 邱均平, 宋艳辉, 杨思洛. 国内人文社会科学文献老化规律对比研究——基于Web新形势下的研究[J].中国图书馆学报, 2011, 37(5): 26-35.
- [9] 苏新宁.中国人文社会科学图书学术影响力报告[M]. 北京: 中国社会科学出版社, 2011.
- [10] 陈云鹏, 张宝林.标准文献半衰期算法及实证研究[J]. 情报杂志, 2012, 31(8):94-97.
- [11] 俞培果.科技文献老化的历时研究[J].情报业务研究, 1993, 10(2):122-126.

作者简介: 张磊(1992-),男,南京大学信息管理学院硕士研究生。

(上接第85页)

- “half-life” of some scientific and technical literatures [J]. American documentation, 1960, 11(1): 18-22.
- [3] 罗式胜.期刊被引半衰期的概念及其应用[J].情报科学, 1994(1):23-25.
 - [4] 罗式胜. 期刊引用半衰期的概念——一种与期刊被引半衰期对应的指标[J].图书与情报, 1995(2):20-21.
 - [5] 罗式胜.关于期刊被引半衰期概念的进一步说明[J]. 国外情报科学, 1996(1):44.
 - [6] 白云.中国人文社会科学期刊被引半衰期分析研究[J]. 云南师范大学学报(哲学社会科学版), 2006(4):127-130.
 - [7] 华涛, 苏新宁, 白云.中国民族研究影响力评价(2002-2004)——民族研究概况与期刊被引分析[J].新疆大学学报(哲学社会科学版), 2006(5):71-78.

- [8] 邓三鸿.民族学期刊学术规范量化指标分析[J].西南民族大学学报(人文社科版), 2008(9):60-65.
- [9] 鞠秀芳.民族学期刊学术影响分析——基于CSSCI(2004-2006)[J]. 西南民族大学学报(人文社科版), 2008(9):66-77.
- [10] 包胜利.2006-2009年中国世界民族研究综述[A].中国世界民族学会.第九届中国世界民族学会会员代表大会暨学术讨论会论文集上册[C].中国世界民族学会, 2010.

作者简介: 张迪(1990-),男,南京大学信息管理学院硕士研究生,研究方向:信息分析、数据挖掘;赵越(1990-),男,南京大学信息管理学院研究生,研究方向:数据分析与信息可视化;邓三鸿(1975-),男,南京大学信息管理学院副教授,研究方向:科学计量、数据分析。