

文章编号:1674-2869(2015)05-0074-05

大数据下中文期刊论文被引分析

朱丽君¹, 陈金芳²

1. 武汉工程大学图书馆, 湖北 武汉 430074;

2. 绿色化工过程教育部重点实验室(武汉工程大学), 湖北 武汉 430074

摘 要: 为了解中文期刊论文被引情况, 用文献计量学的方法对 2006~2014 年中国知网数据库收录的武汉工程大学期刊论文进行统计, 分别从年度被引分布、被引频次对比及《武汉工程大学学报》被引情况等方面进行了系统分析, 发现中文期刊论文被引率为 62.91%, 篇均被引频次 2.76 次。不同类型、不同学科领域被引用情况不同, 反映出武汉工程大学在相关领域研究中的影响力不断提升, 科研投入和人才引进战略逐渐收到实效。但自 2012 年以来, 被引率递减加大, 零被引篇数增多, 需要制定相应鼓励政策, 重视期刊论文被引率, 进一步提高论文质量。

关键词: 被引率; 被引频次; 期刊论文

中图分类号: G353.1

文献标识码: A

doi: 10.3969/j.issn.1674-2869.2015.05.015

0 引言

在期刊论文文献体系中, 论文之间并不是孤立存在的, 而是通过研究者之间相互交流、借鉴、参照、继承、质疑产生联系, 形成引文, 产生被引率。美国学者 Markusov^[1]和 Smart 等^[2]先后在 1973 年和 1981 年提出“被引率”这一概念。Van Campenhout^[3]提出论文被引率与同行公认的其它指标高度相关。由此, 被引率渐渐成为论文质量高低的参照点, 并成为人们关注的重要参考标志。一般来说, 被引率是指该期刊前 5 年发表论文在统计当年被引论文篇数/该期刊前 5 年发表论文篇数。被引率反映了期刊论文被利用的情况, 被引率越大, 反映期刊论文的被利用率越高。同时反映着科技期刊所载论文的学术价值、资料价值和适用价值。

随着大数据时代的到来, 在数据处理的理念上发生了三大转变: 一是统计全部数据而不是抽样数据, 二是数据要及时有效而不是绝对精准, 三是更注重揭示相关关系而不是探究因果关系^[4]。大数据的核心是预测, 即根据全面的数据预估未来事情的发展趋势^[5]。基于此, 笔者以武汉工程大学发表的中文期刊论文被引数据为例, 探讨武汉工程大学中文期刊论文被引变化态势, 找寻提升被引率的对策。

1 统计说明及方法

自 2006 年 2 月更名为武汉工程大学后, 其论

文发文量稳步增长, 论文整体呈现出化工特色、多学科特色、时代特色和地域特色^[6]。笔者对武汉工程大学 2006~2014 年所发表的中文期刊论文被引情况进行统计, 并分析被引年度分布、被引频次比较及《武汉工程大学学报》被引数据。

1.1 统计范围

本文的主要统计范围为 2006~2014 年发表于正式公开期刊上的以武汉工程大学为作者单位的所有中文期刊论文, 且会议论文和论文集不在统计范围之内。

1.2 数据来源

笔者以中国知网 (China National Knowledge Infrastructure, 简称: CNKI) 的中国学术期刊网络出版总库 (China Academic Journal Network Publishing Database, 简称: CAJD) 所收录的 2006~2014 年中文期刊论文为统计数据源。

1.3 统计方法

进入中国知网 <http://www.cnki.net/>, 选择资源总库下的中国学术期刊网络出版总库, 使用高级检索, 在时间限定中输入 2006~2014 年, 并在作者单位中输入“武汉工程大学”, 将检索结果依据“被引”进行排序统计。

1.4 统计时间

本文各种数据统计时间以 2015 年 3 月 12 日前的“中国知网”中国学术期刊网络出版总库的数据为例。

收稿日期: 2015-03-16

作者简介: 朱丽君 (1964-), 女, 湖南汨罗人, 研究馆员。研究方向: 信息管理。

1.5 相关定义

通常被引率是按学科来区分,如果按单位区分可靠度就会降低.但本文是在大数据背景下统计,只作预测和趋势.因此本文在分析中用到的被引篇数、被引率和篇均被引频次等多个文献计量指标均指某时期数据,即被引率是指武汉工程大学某时期论文的被引篇数÷某时期论文总篇数×100%;篇均被引频次是指武汉工程大学某时期论文被引频次÷某时期论文总篇数.

2 统计结果分析

2.1 论文被引的年度分布

在 2006~2014 年间,武汉工程大学共发表中文期刊论文 9 321 篇,被引篇数 5 864 篇,被引率为 62.91%,被引频次为 25 752 次,篇均被引频次 2.76 次,如表 1 所示.由表 1 可知,论文总量在逐年稳步缓慢上升;被引篇数在 2010 年达到最高点的 870 篇后,从 2012 年开始加大了下降的幅度;被引频次也出现同样现象.

表 1 武汉工程大学论文被引年度分布表

Table 1 Annual citation distribution of papers published by Wuhan Institute of Technology

年份	论文篇数/篇	被引篇数/篇	被引率/%	被引频次/次	篇均被引频次/次
2006	722	584	80.89	3 718	5.15
2007	1 007	831	82.52	4 718	4.69
2008	926	762	82.29	4 420	4.77
2009	1 056	836	79.17	4 148	3.93
2010	1 113	870	78.17	3 591	3.23
2011	1 131	794	70.20	2 540	2.25
2012	1 069	637	59.59	1 697	1.59
2013	1 215	457	37.61	789	0.65
2014	1 082	93	8.60	131	0.12
合计	9 321	5 864	62.91	25 752	2.76

2.2 论文被引随时间变化情况

武汉工程大学中文期刊论文被引率为 62.91%,其中,2007 年论文被引率最高,2014 年论文被引率最低.随着论文发表后时间的推延,被引机会或被引频率均会增加,因此论文发表后时间长短不同,其论文被引率也不同,但这并不是绝对的.总体上是发表时间越久,被引率越高.发表时间越近的期刊论文被引用的机会越大,如表 2 所示.从表 2 中可以看出,在两个多月时间里,各年度论文被引篇数和被引频次均有变化.2007 年论文被引篇数增加了 5 篇,论文被引频次增加 145 次,而

2012 年论文被引篇数增加 44 篇,论文被引频次增加 232 次.从时间分布来看,并不是随着论文篇数的增加,其被引篇数也增长,2011 年论文总篇数达到 1 131 篇,但其被引篇数却只有 794 篇.对个体论文来说,其被引大多集中在发表后 5 年内,如果 5 年内没有被引用,则今后被引用的概率也不大^[7].一般来说,中文期刊自然科学基金论文发表后的前 4 年,是论文由零被引转向非零被引的关键时段^[8].另外,从《武汉工程大学学报》在 2007~2014 年间论文被引情况中也可以看出论文被引用多集中在发表后的 5 年内,如表 3 所示.

表 2 武汉工程大学论文被引短时变化表

Table 2 Citation changes in short term of papers published by Wuhan Institute of Technology

年份	论文篇数/篇	2015 年 1 月 5 日统计数		2015 年 3 月 12 日统计数	
		被引篇数/篇	被引频次/次	被引篇数/篇	被引频次/次
2007	1 007	826	4 573	831	4 718
2008	926	756	4 189	762	4 420
2009	1 056	821	3 899	836	4 148
2010	1 113	861	3 418	870	3 591
2011	1 131	757	2 258	794	2 540
2012	1 069	593	1 465	637	1 697
2013	1 215	377	570	457	789

表 3 《武汉工程大学学报》被引年度分布表

Table 3 Annual citation distribution of papers published by *Jorunal of Wuhan Institute of Technology*

年份	论文篇数/篇	被引篇数/篇	被引率/%	被引频次/次	篇均被引频次/次
2007	173	141	81.50	750	4.34
2008	211	177	83.89	1 012	4.80
2009	290	258	88.97	1 541	5.31
2010	335	290	86.57	1 461	4.36
2011	323	249	77.09	836	2.59
2012	209	151	72.25	333	1.59
2013	202	84	41.58	156	0.77
2014	183	25	13.66	34	0.19
总计	1 926	1 375	71.39	6 123	3.18

2.3 论文被引频次高低比较

不同的文献类型、引文窗口、期刊论文主题等都会对期刊论文的被引周期发生影响^[9]。

笔者统计了 2006~2014 年间武汉工程大学中文期刊论文最高被引频次和最低被引频次,并统计了 2006~2014 年间《武汉工程大学学报》被引频次最高的前 10 篇论文,分别列于表 4 和表 5 中。

在 2006~2014 年度中,武汉工程大学中文期刊论文最高被引频次为 150 次,最低被引频次为 1 次。共有 5 864 篇论文被引用,其中有 1 738 篇论文仅被引用 1 次。2006 年最高被引频次为 150 次的论文为“国外主流写作理论对我国外语写作教学的启示”;2014 年最高被引频次为 8 次的论文是“MOOC 课程模式及其对高校的影响”。从年度最高

表 4 2006~2014 年度最高和最低被引频次对比列表

Table 4 Contrast between the highest and the lowest citation frequencies from 2006 to 2014

年份	最高被引 次数/次	论文题名	最低被引 数/次	最低被引 篇数/篇	零被引 篇数/篇
2006	150	国外主流写作理论对我国外语写作教学的启示	1	114	138
2007	55	超级电容器导电聚合物电极材料的研究进展	1	168	176
2008	95	煤矿安全监控系统的现状与发展趋势	1	129	164
2009	73	一种基于文件过滤驱动的 Windows 文件安全保护方案	1	207	220
2010	32	国外循环农业发展模式及启示	1	227	243
2011	26	房地产预警指标体系及综合预警方法研究	1	278	337
2012	19	考虑战略顾客行为时的供应链性能分析与协调	1	256	432
2013	10	MOOC 下的高职教育改革探索	1	285	758
2014	8	MOOC 课程模式及其对高校的影响	1	74	989

表 5 《武汉工程大学学报》被引频次最高的前 10 篇论文

Table 5 The top 10 papers of citation frequencies from *Journal of Wuhan Institute of Technology*

被引频次/次	题名	发表年份
37	上市公司内部控制信息披露质量实证研究	2009
36	武汉市大学生创业状况调查研究	2009
36	DC-DC 转换电路在光伏发电 MPPT 中的应用	2008
32	我国磷资源开发利用及趋势	2011
32	数字化校园建设中宿舍管理系统的设计与实现	2008
30	基于梁格法及板壳有限元法的立交桥空间分析	2008
29	基于虚拟化技术构建高校分布式云计算数据中心	2011
28	我国物流信息化状况及启示	2009
28	超声波辅助合成离子液体及其性能研究	2007
27	国外大学生职业发展与就业教育对我国高校的启示	2009

被引频次的这9篇论文内容特征看,多数是综述性论文,涉及外语教学、材料、煤矿安监、网络安全、循环农业、房地产预警、管理科学及教学改革等,其中MOOC网络学习模式尤其体现了当前的研究热点,也与我国乃至世界当前的技术发展有关。从表4中还可以看到,自2009年后,最低被引篇数均在200篇以上,零被引篇数也呈逐年上长趋势,但随着时间推移,这个动态数据会逐渐缩小的。

从《武汉工程大学学报》最高被引频次前10篇论文看,论文的发文时间也多集中在2008、2009和2011年间,其内容特点也多倾向于综述性或教学、就业相关。

3 结 语

a.统计结果表明,在2006~2014年,武汉工程大学发表的中文期刊论文被引率为62.91%。从年度被引率看,2007年最高,2014最低,2006~2011年被引率均高于平均被引率62.91%。2012年以后,被引率降为60%以下,而且每年递减较大,虽然发表时间越久,越有可能被引,被引率也就越高,但这并不是绝对的,比如2007年的被引率既高于2006年被引率,也高于2008年被引率。当然,不同学科,其被引率也不同。譬如,我国高被引科学家主要集中在化学、工程、材料科学领域^[10]。文中的论文被引数据是年度平均值,只是代表了一种论文被引趋势,并不能明示具体学科的论文被引情况。

b.从期刊论文被引频次高低比较看,在2006~2014年共发文9321篇,其中被引用的5864篇,零被引用的3457篇。零被引用论文篇数占全部论文篇数37.09%。2004年英国学者Weale等^[11]提出零被引率(non-cited rate)概念,作为一种期刊质量反向评价指标。杨利军等研究发现,零被引率与我国学者在写作时,参考文献偏向于引用非中文期刊类型的文献,而对中文期刊论文的引用不足有关^[12],这在某种程度上反映了国内目前缺乏科研自信的学术氛围,以为外文文献就是好文献,引用外文文献就能提高文章质量的盲从思维^[13],当然也与发文的时间周期、内容质量和研究热点有关。

1次引用频次的论文为1738篇,占被引篇数29.64%(近三分之一),占全部论文篇数18.65%。由此说明论文内容质量还需要进一步提高。

c.从年度最高被引频次看,最高被引频次的9篇论文既有共性又各有特色,其共性多呈现在论文的综述性质,其个性多与专业的普及性、发展性和时代的需求性有关,如外语教学、电极材料、煤矿

安全、网络安全、循环农业、房地产预警、企业管理等均体现了这些特性,尤其近来随着互联网发展而出现在MOOC网络学习模式,更是引起人们的广泛关注。因此,要想提高论文被引频次,在论文选题上应该首选紧跟时代热点,结合现实问题的综述性文章。

d.从《武汉工程大学学报》被引频次前10的论文来看,其多体现地域、单位特色及重点专业研究特色,被引论文发文的时间多在5年内,被引率保持逐年上升趋势。

e.虽然论文被引用客观地反映了其他作者对该文的认同,但不同类型的文章、不同学科领域,引用的情况也不一样。武汉工程大学被引用高的论文一方面反映出本校在相关领域研究中的影响力不断提升的事实,另一方面也反映本校近几年科研投入、人才引进、团队组建等方面所收到的实效。

f.不同学科领域、不同类型的文章,其引用的情况也不一样。在大规模数据背景下,对各个学科引用情况归一化处理后,平均来看引用得越多,说明关注度越高,在一定程度上也可以表明科研成果的水平越高。比如综述类文章引用率就比较高,人文科学引用往往比较少,且人文社科领域往往通过专著而非论文来体现学术水平^[10]。从武汉工程大学年度最高被引频次论文和《武汉工程大学学报》最高被引频次前10篇论文来看,提高论文被引率的对策就是论文的内容要具有新颖性和时代性,研究要具备突破性,结论要具有数据证明性和区域行业通用性特点等,且类型多是“研究综述”、“发展趋势”等综述性的论文。另外,还需制定相应的被引率奖励政策,鼓励作者写出高质量的期刊论文。

致 谢

感谢武汉工程大学图书馆对本文的数据支持!

参考文献:

- [1] MARKUSOV V A. Comparative study of citation rates of Soviet scientific and technical publication[J]. Nauchno-Tekhnicheskaya Informatsiya Seriya, 1973(1):27-31.
- [2] SMART J C, ELTON C F. Structural characteristics and citation rates of education journals[J]. American Educational Research Journal, 1981, 18(4):399-413.
- [3] VAN CAMPENHOUT G, VAN CANEGHEM T. Article contribution and subsequent citation rates: evidence from European accounting review[J]. European Accounting Review, 2010, 19(4):837-855.
- [4] [英] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶. 大数

- 据时代:生活、工作与思维的大变革[M]. 盛杨燕,周涛,译.杭州:浙江人民出版社,2013:17-18,94.
- [5] 柴英,马婧.大数据时代学术期刊功能的变革[J]. 编辑之友,2014(6):28-31.
- [6] 朱丽君. 学科发展对发文量的影响[J]. 武汉工程大学学报,2011,33(12):99-103.
- ZHU Li-jun. Influence of discipline development on the number of published paper[J]. Journal of Wuhan Institute of Technology, 2011, 33(12):99-103. (in Chinese)
- [7] 盖双双,张诗乐,刘雪立,等. 论文被引率在科技期刊评价中的地位 and 作用——基于 36 种 SCI 眼科学期刊问卷调查的实证研究[J]. 中国科技期刊研究, 2014, 25(1):39-43.
- GAI Shuang-shuang, ZHANG Shi-le, LIU Xue-li, et al. Roles of papers cited rate in scientific journal evaluation: empirical study based on the questionnaire of 36 ophthalmologic journals on the SCI database[J]. Chinese Journal of Scientific and Technical Periodicals, 2014, 25(1):39-43. (in Chinese)
- [8] 董建军. 中国知网收录的国家自然科学基金论文的被引情况分析[J]. 中国科技期刊研究, 2012, 23(5): 776-778.
- DONG Jian-jun. Citation analysis of papers sponsored by National Natural Science Foundation of China in CNKI[J]. Chinese Journal of Scientific and Technical Periodicals, 2012, 23(5):776-778. (in Chinese)
- [9] 董建军,盛丽娜,郑成铭,等. 期刊年被引论文率对期刊评价作用的探讨[J]. 中国科技期刊研究, 2014, 25(8):1000-1004.
- DONG Jian-jun, SHENG Li-na, ZHENG Cheng-ming, et al. Effect of cited paper rate per year on journal evaluation[J]. Chinese Journal of Scientific and Technical Periodicals, 2014, 25(8):1000-1004. (in Chinese)
- [10] 齐芳. 如何看待全球高被引用率排名——访中科院文献情报中心张晓林研究员[N]. 光明日报, 2014-8-18(6).
- [11] WEALE A R, BAILEY M, LEAR P A. The level of non-citation of articles within a journal as a measure of quality: a comparison to the impact factor[J]. BMC Medical Research Methodology, 2004(4):14-21.
- [12] 杨利军,万小渝. 引用习惯对我国期刊论文被引频次的影响分析——以情报学为例 [J]. 情报科学, 2012, 30(7):1093-1096.
- YANG Li-Jun, WAN Xiao-Yu. The impact of cite habits to the citations of journal articles in china: A case study of information science[J]. Information Science, 2012, 30(7):1093-1096. (in Chinese)
- [13] 孔朝霞, 陈璐. 引用习惯对中文期刊零被引率的影响分析——以《中国医学科学院学报》10 年载文为例[J]. 科技与出版, 2013, 32(6):126-129.

Citation analysis on Chinese periodicals under big data

ZHU Li-jun¹, CHEN Jin-fang²

1. Library, Wuhan Institute of Technology, Wuhan 430074, China;

2. Key Laboratory of Green Chemical Process (Wuhan Institute of Technology), Ministry of Education, Wuhan 430074, China

Abstract: The papers of Chinese periodicals of Wuhan Institute of Technology in CNKI from 2006 to 2014 were selected to study the citation situation, including annual cited rates distribution, citation frequencies and contrast on cited rates. The results show that the cited rate is 62.91% and the average cited frequency per paper is 2.76. This indicated an improvement of academic impact and a great effect of talent introduction and scientific research input of Wuhan Institute of Technology. However, cited rates have decreased and non-cited rates have increased since 2012, which calls for relevant policy-making, attaching great importance to cited rates and journal quality improvement.

Keywords: cited rates; citation frequency; periodicals

本文编辑: 苗 变