

DOI:10.12154/j.qbzlgz.2019.06.003

网络环境下论文影响力综合评价体系构建*

——基于时间因素视角

刘运梅 李长玲 杜德慧 (山东理工大学科技信息研究所 淄博 255049)

摘要 网络环境下,为了有效评价不同年龄的论文综合影响力,文章根据 Altmetrics 不同指标随时间变化的不同表现,提出时间因素视角下的首次他引速度Cs、月被引峰值Ctop、月均使用量Ut、首月讨论量T1等指标,并经过数据标准化、分配权重将各指标结合起来,构建动态、全面的论文影响力综合评价体系,以降低时间因素对指标数值累积的影响,更客观、公正地评价新老文献。文章选择PLOS网站的PLOS使用数据、Scopus引用数据、Twitter讨论数据,以2011年至2017年每年被引频次最高的10篇高被引论文为样本,进行实证研究。研究发现:多维度的动态论文影响力综合评价体系,有效地解决了时间因素的累积效应,在公平客观地评价不同年龄文献的同时,还可以有效识别发文时间较短的高价值新文献。

关键词 学术影响力 社会影响力 使用指标 引用指标 讨论指标

Construction of Comprehensive Evaluation System of Paper Influence under the Network Environment: Based on the Perspective of Time Factor

Liu Yunmei Li Changling Du Dehui

(Institute of Science and Technical Information, Shandong University of Technology, Zibo, 255049)

Abstract In the network environment, in order to effectively evaluate the comprehensive influence of papers in different ages, this paper proposes firstly citation speed Cs, the monthly citation peak Ctop, the monthly average usage Ut and the first month discussion T1 according to the different performances of different indicators over time, and use the standardization and distribution weights to combine the evaluation indicators to build a dynamic and comprehensive evaluation system for the influence of the paper. Taking the PLOS usage index, Scopus citation index and Twitter discussion index on the PLOS website as examples, the top 10 highly cited papers with the highest citation frequency from 2011 to 2017 were selected as samples for empirical research. It is found that the multi-dimensional dynamic evaluation system effectively solves the cumulative effect of time factors. While evaluating the papers of different ages fairly and objectively, it can also effectively identify high-value new papers with short published time.

Keywords academic influence, social influence, use indicators, citation indicators, discussion indicators

1 引言

学术论文作为科学研究成果的文献记录,是科学家最重要的科研成果之一,也是衡量科研水平的重要

评价依据^[1]。论文的影响力主要体现在学术影响力和社会影响力两个方面。其中,学术影响力表征论文在学术界的科研价值与论文质量,目前常用的学术影响力评价指标有期刊影响因子、单篇论文被引频次等,但

*本文系国家自然科学基金重点项目“跨学科潜在知识生长点识别与创新趋势预测研究”(编号:19ATQ006)的研究成果之一。

这些以被引为主的定量评价方法具有时间滞后性、单一性和片面性;社会影响力表征论文在社会大众中的受关注与传播情况^[2],以论文在社交媒体工具、开放存取平台的在线使用、收藏、提及与讨论等指标为代表,然而这些新兴评价指标的覆盖率、系统性、稳定性均有待提高,并不能独当论文影响力综合评价的重任。在这种背景下,极有必要构建基于多维指标的影响力综合评价体系,以挖掘不同层次、不同角度的高价值文献。

2010年,Priem等^[3]提出 Altmetrics,其全面包含了论文在各大文献数据库、社交媒体平台、文献管理工具、新闻媒体中20余种引用、使用、讨论、收藏评价指标,可以提供论学术影响力和社会影响力评价的依据。目前,国内外有关 Altmetrics 论文评价的研究主要集中于以下三个方面:(1)Altmetrics 评价学术论文的意义、价值探讨。Altmetrics 的提出者 Priem 教授^[3]指出“替代计量学是出于分析和理解学术目的、基于社会网络构建的新计量指标研究”;邱均平等^[4]认为 Altmetrics 可以衡量学术论文多方面的影响力,如社会影响力、教育影响力、经济影响力;另外,也有一些学者认为 Altmetrics 作为一种新型学术评价指标,严谨性不高^[5-7]。(2)网络评价指标与传统引用指标关系的研究。有学者就两者的相关性进行实证研究,结果各有不同。Ortega^[8]随机抽取一万名西班牙国家委员会学者的论文进行研究,发现网络评价指标较引用指标表现出更强的平台依赖性,更倾向于反映学术成果的受欢迎程度和网络传播能力;Bornmann^[9]通过荟萃分析法发现,线上文献管理平台与引用指标之间具有较高的相关性,可作为传统引用指标的补充;宋丽萍等^[10]将生物信息学、免疫学两个学科论文的网络评价指标与引用指标进行相关性检验,结果呈现低度相关;余厚强^[11]以百万条学术论文的 Altmetrics 数据为基础,从数值类型、指标种类、用户类别三个角度考察网络评价指标与引用指标的相关性,发现独立用户数量、网络平台类型、学科种类、用户类别对两者相关性均有影响。(3)利用不同的 Altmetrics 数据源,提出学术论文评价指标或方法。赵蓉英等^[12]运用相关性分析法和主成分分析法,对 Altmetrics 指标定量分析和选取,构建学术论文影响力评价框架及模型;由庆斌等^[13]收集 Mendeley 平台中提供的指标数据,运用相关分析法、主成分分析法构建单篇学术论文影响力评价模型,该模型对学术论文的评

价效果与引用评价模型一致;宋丽萍等^[14]以四个学科为例,利用主成分分析方法对 ALM 指标解析,发现三个科学评价维度(引用维、共享维、利用维)只能以80%的精度描述论文的影响力;Wang等^[15]提出了一个连续、动态的论文评价体系,将论文发表后的时间分为0~6个月、6个月~2年、2~5年、5年以上四个阶段,不同评价指标根据影响力大小在四个时间阶段中动态调整其权重,纳入综合评价体系。

综上所述,(1)目前用 Altmetrics 评价论文影响力的研究文献未考虑时间因素——文献年龄对论文评价的影响。一篇论文正式发表后,各 Altmetrics 指标的数值累计量会有增无减,使得“老文献”更具评价优势。但实际上,具有相同指标数值累计量的“年龄较小论文”具有更大的影响力。(2)多数 Altmetrics 定量研究集中于运用相关性分析、主成分分析等方法,局限于平台的固定化指标数值,将论文影响力看作是一种结果而忽视其过程。然而大部分论文评价指标都具有动态特性,论文在发表后不同时间阶段数值变化和作用发挥不尽相同^[16]。研究发现^[17]:引用指标具有低及时性、高波动性、高持续性的特点,可将学术论文中长期的被引情况作为学术影响力判定指标;使用指标具有高及时性、低波动性、高持续性的特点,可将学术论文整个生命周期的平均被使用情况作为社会影响力判定指标;讨论指标具有低及时性、低波动性、低持续性的特点,可将学术论文首月的讨论情况作为社会影响力判定指标。

因此,针对以上问题,本文根据不同指标随时间变化的不同表现,提出时间因素视角下的首次他引速度、月被引峰值、月均使用量、首月讨论量等指标,并通过标准化、分配权重将各评价分指标结合起来,构建动态、系统、全面的论文影响力综合评价体系,将各项指标的最佳效力发挥时间、数据量的时间变化规律纳入其中,降低时间因素对指标数值累积的影响,从引用、使用、讨论等方面综合、客观、公正评价新老文献。

本文的设计思路如下页图1所示。

2 基于时间因素的分指标提出

2.1 引用指标——首次他引速度 C_s 、月被引峰值 C_{top}

一篇论文在另一篇论文中被正式引用即形成论文被引,传统引用指标在反映文献价值方面具有一定的时间滞后性,一篇论文在发表后几个月甚至一年内被

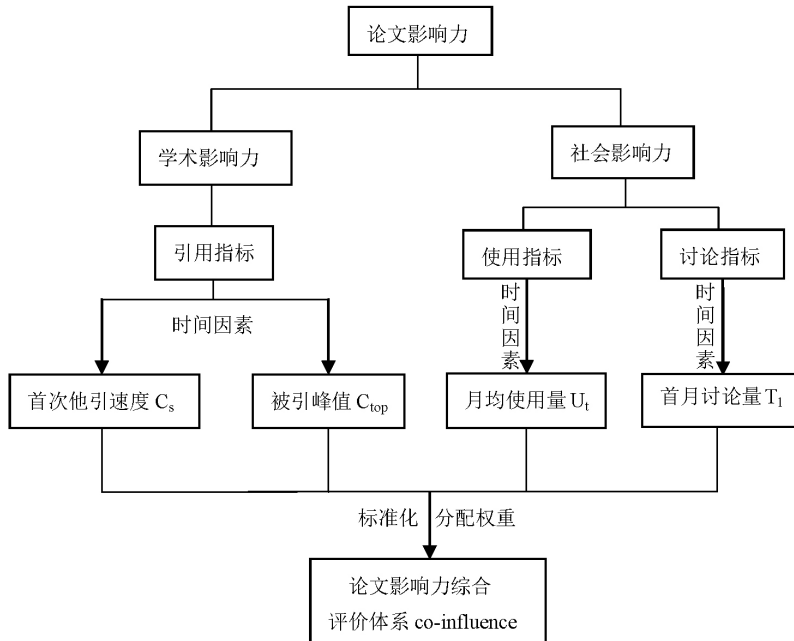


图1 研究思路流程图

引用数据较少,还远远未积累到能够对论文价值做出准确评价的程度^[17]。因此,有必要构建能够规避引用时间滞后问题的引用分指标,提高文献的新颖性、即时性评价效果。

一般来说,文献越长时间未被引用或关注,其价值就认为越低,反之文献首次被引所经历的时间越短,其科学价值和学术影响就越大^[18]。2011年,Egghe等^[19]提出文献首次被引速度 $t_f = t_c - t_p (t_c \geq t_p)$, t_c 表示该文献首次被引用的日期, t_p 表示出版日期。首次被引是论文第一次被引用,受该论文作者的研究连贯性或主观意图影响,部分论文的首次被引可能为论文作者进行的自引行为。考虑到应用首次被引是为了体现论文影响力与受关注程度,本文对首次被引速度 t_f 改进,去除论文首次被引为作者自引的情况,提出第一个引用评价分指标——论文“首次他引速度 C_s ”。一般地,文献首次被引所经历的时间越短,其影响力越大。为了表达首次他引时间与影响力的反向关系,本文令论文首次他引速度 C_s 为一篇论文自正式发表后,首次被他人(除论文所有署名作者外的其他学者)引用的时间(月)取倒数,即:

$$C_s = \frac{1}{(T_c - T_p)/30} = \frac{30}{T_c - T_p} (T_c > T_p) \quad (公式1)$$

其中, T_c 表示该论文首次被他引的日期, T_p 表示论文出版日期。

另一方面,传统引用指标未考虑评价论文的年龄问题,但发表年龄不同、被引频次相同的论文,其影响力是不同的。因此,构建一个能够规避引用数据时间累积问题的引用分指标,在理论和实践上都具有重要意义。

由于论文引用数据不同于使用指标由公众广泛参与形成流量数据,其形成具有较大的限制条件、主客观复杂因素,导致论文各月份的被引频次数值较小、波动较大^[17],形成被引峰值。若使用时间稀释作用的被引均值指标来衡量论文被引,只对达到被引峰值时间前后的文献具有较好的评价效果,难以突显高价值、发表时间较短或较长文献的影响力。因此,本文以月份为单位,用论文“月被引峰值 C_{top} ”表征论文的学术影响力,以弱化时间

累积效应影响,月被引峰值 C_{top} 可表示为:

$$C_{top} = \max(c_1, c_2, c_3, \dots, c_n) \quad (公式2)$$

其中,某篇论文自正式发表至评价时点共经过 n 个月, c_1, \dots, c_n 分别表示该论文发表第一个月到第 n 个月的被引频次。

2.2 使用指标——月均使用量 U_t

使用数据是已正式发表的科学文献被浏览或下载产生的数据^[20],根据文件格式,浏览主要是HTML格式,下载可分为PDF格式和XML格式。在本文前期的研究成果^[17]中发现,总使用量指标能在第一时间对论文的被使用情况做出反应,并且随着时间流逝,一篇论文会由于各种原因被不断使用,其每个月份的使用量在一定数值水平上呈现较为平稳的发展。因此,论文总使用量指标容易受时间累积因素的影响。

月平均使用量作为一种稀释效应,可缓解时间因素的影响,相同总使用量的论文发表时间越长,其稀释效果越明显,可较公平地反映论文在网络上的被使用、被关注情况。因此,本文提出“月均使用量 U_t ”作为使用类分指标,即论文自发表之日起到评价时点的总使用量与相隔月份数之比,数学表达式如下:

$$U_t = \frac{U}{t} = \frac{HTML + PDF + XML}{t} \quad (公式3)$$

其中, U 表示某论文自发表以来获得的总使用量

(即HTML浏览量、PDF与XML下载量之和),论文自发表到评价时点相隔月份数为 t 。

2.3 讨论指标——首月讨论量 T_1

讨论数据是收集科学文献在网络社交媒体(Facebook、Twitter等)、学术型网站平台(Wikipedia、ResearchGate等)上对一篇论文研究内容讨论的数据^[21]。若论文具有传播价值,短时间内会在各种推荐与传播中产生较高的讨论数量,能够及时地体现文献价值。然而,讨论数据的产生伴随着社会公众的新鲜感^[20],在论文发表后不久便会被时间冲淡,后续再较少出现讨论数据,总讨论量基本维持在早期的数据量水平。有研究成果发现,大部分论文的讨论量均集中于论文发表后的第一个月,其余大部分月份的讨论数据基本为零^[17]。

讨论指标随时间推移发生的变化较大,呈现出快速利用和快速老化的普遍特征,倾向于新文献的发现和传播,而在后期持续性不足。因此,本文提出“首月讨论量 T_1 ”作为讨论分指标,即一篇论文自发表之日起30天时间里,在网络社交媒体、学术型网站平台获得的讨论数量,数学表达式如下:

$$T_i = T_{D1} + T_{D2} + T_{D3} + \dots + T_{D30} = \sum_{Di=1}^{30} T_{Di} \quad (\text{公式4})$$

其中, Di 表示某论文正式发表之后的第 i 天, T_{Di} 表示第 i 天里获得的网络讨论数量。

3 基于时间因素的论文影响力综合评价体系构建

由于以上四个指标(首次他引速度 C_s 、月被引峰值 C_{top} 、月均使用量 U_i 、首月讨论量 T_1)具有不同的量纲和数量级,需要首先对原始数据进行标准化。本文选用离差标准化方法^[22],对指标数据 x_1, x_2, \dots, x_n 进行变换,为了体现数据结果的可比较性,并将其转化为百分制的标准分数,计算公式如下:

$$y_i = \frac{x_i - \min_{1 \leq j \leq n} \{x_j\}}{\max_{1 \leq j \leq n} \{x_j\} - \min_{1 \leq j \leq n} \{x_j\}} \times 100 \quad (\text{公式5})$$

则新序列 $y_1, y_2, \dots, y_n \in [0, 100]$,且无量纲。经过标准化后,以上四个影响力评价指标均转化为统一的标准格式,分别用 $C'_s, C'_{top}, U'_i, T'_1$ 表示。

引用指标表征论文的学术影响力,是论文最高层次的利用,其在论文影响力评价中仍然占据主流地位^[23],因而设置引用分指标(包括首次他引速度 C_s 与月被引峰值 C_{top})在论文影响力综合评价指标中占较大权

重,赋予其权重为0.4;其次,使用指标(月均使用量 U_i)与讨论指标(首月讨论量 T_1)从论文的交流利用情况、在线社交媒体交流情况来度量论文的社会影响力,是对传统引文评价的补充指标,本文赋予月均使用量 U_i 、首月讨论量 T_1 指标的权重系数均为0.3。其中,本文认为首次他引速度与月被引峰值对于论文被引能力的判定同等重要,分别赋予两者0.5的权重。因此,论文影响力综合指标(*co-influence*)计算方法如下:

$$co-influence = 0.4 \times [0.5 C'_s + 0.5 C'_{top}] + 0.3 \times U'_i + 0.3 \times T'_1 \quad (\text{公式6})$$

4 实例验证

4.1 数据来源与指标计算

本文选择PLOS ONE期刊作为数据来源样本,同时选取权威数据库Scopus的引用数据作为引用类指标的代表;使用类指标选择PLOS网站的HTML浏览量、PDF、XML下载量总和;讨论类指标选择全球范围内比较普及和热门的社交软件Twitter的讨论数据。为了较好地突出时间因素对论文评价的影响,并分析基于时间因素的分指标及综合指标的评价效果,本文依次选取PLOS ONE期刊2011年至2017年每年被引频次最高的前10篇高被引论文作为样本,通过分析这70篇论文的引用、使用、讨论指标,以评价样本论文的综合影响力水平。分别给70篇论文编号1~70,数据采集时间为2018年7月1日。

借助Python爬虫工具,在PLOS网站中获取这70篇高被引论文的发表时间、PLOS总使用量、Twitter网站每条用户评论的产生时间,以及Scopus数据库中每条施引文献公开出版时间(部分施引文献的出版时间在Scopus数据库中缺失,通过“全渠道搜索引擎”查阅原文,获取其“published time”)。对获取的数据做如下处理:(1)基于样本论文的发表时间、施引文献发表时间,根据公式1与公式2分别计算得到70篇样本论文的首次他引速度 C_s 、月被引峰值 C_{top} ,结果见下页表1第四、五列。(2)基于样本论文的发表时间、PLOS总使用量,根据公式3计算70篇高被引论文的月均使用量 U_i ,结果见表1第八列。(3)基于每篇样本论文的发表时间、Twitter用户评论产生时间,统计得到70篇样本论文的首月讨论量 T_1 ,见表1第十列。(4)运用公式5,将以上计算得到的70篇论文四个指标数值(首次他引速度 C_s 、

表1 70篇样本论文的影响力综合指标计算数据

序号	出版年份	总被引频次	首次他引速度 C_s	月被引峰值 C_{top}	发表月数	总使用量	月均使用量 U_t	总讨论量	首月讨论量 T_1	影响力综合指标 co-influence	序号	出版年份	总被引频次	首次他引速度 C_s	月被引峰值 C_{top}	发表月数	总使用量	月均使用量 U_t	总讨论量	首月讨论量 T_1	影响力综合指标 co-influence
1	2011	1705	0.78	150	87	136155	1565	0	0	28.407	36	2014	266	0.3	19	54	34673	642.09	40	26	5.131
2	2011	1080	0.15	67	85	56399	663.52	0	0	9.721	37	2014	281	0.39	21	50	59617	1192.34	62	47	7.215
3	2011	965	0.24	67	80	45683	571.04	0	0	10.537	38	2014	258	0.21	20	52	15163	291.6	0	0	3.767
4	2011	972	0.3	63	83	37780	455.18	0	0	10.613	39	2014	333	1.14	16	54	30977	573.65	8	5	14.009
5	2011	714	0.3	48	84	50193	597.54	0	0	8.464	40	2014	230	0.36	14	47	7313	155.6	0	0	4.525
6	2011	626	0.18	29	90	65439	727.1	0	0	4.86	41	2015	248	1.02	20	32	36597	1143.66	10	10	13.64
7	2011	456	0.18	35	87	76346	877.54	0	0	5.824	42	2015	203	0.21	13	42	13627	324.45	5	4	2.697
8	2011	448	0.24	39	87	33659	386.89	0	0	6.488	43	2015	178	0.12	15	41	33227	810.41	35	28	2.721
9	2011	445	1.89	25	85	42020	494.35	0	0	23.259	44	2015	191	0.24	9	38	17482	460.05	1	1	2.71
10	2011	465	0.27	27	84	44245	526.73	0	0	5.533	45	2015	140	0.15	9	41	2505	61.1	0	0	1.42
11	2012	1804	0.66	95	75	79994	1066.59	12	6	19.399	46	2015	135	0.3	8	42	169040	4024.76	19	7	5.246
12	2012	825	0.15	47	70	37822	540.31	4	3	6.698	47	2015	122	0.48	8	38	123169	3241.29	248	221	10.772
13	2012	710	0.21	38	78	69587	892.14	0	0	6.475	48	2015	150	0.42	14	36	42463	1179.53	41	24	5.971
14	2012	654	0.42	30	75	101066	1347.55	10	4	8.027	49	2015	135	0.6	17	35	16523	472.09	25	15	8.129
15	2012	535	0.33	23	78	42193	540.94	0	0	5.652	50	2015	114	1.35	15	38	136410	3589.74	1506	1359	41.907
16	2012	468	0.81	19	78	56957	730.22	0	0	10.543	51	2016	87	0.18	11	26	15372	591.23	18	17	2.5
17	2012	424	0.36	25	77	22498	292.18	0	0	5.908	52	2016	147	0.24	12	30	16296	543.2	11	8	3.317
18	2012	394	0.21	16	73	23791	325.9	7	6	3.258	53	2016	62	0.21	8	25	3099	123.96	0	0	1.731
19	2012	370	0.6	27	75	16943	225.91	0	0	9.074	54	2016	60	0.3	9	28	10346	369.5	22	17	3.548
20	2012	502	0.12	28	78	46415	595.06	7	4	3.84	55	2016	62	0.18	9	26	5028	193.38	12	11	1.788
21	2013	1005	0.12	44	64	70562	1102.53	30	21	6.791	56	2016	53	0.12	5	29	3317	114.38	0	0	0.305
22	2013	706	0.24	36	61	36342	595.77	8	7	6.436	57	2016	54	0.3	9	28	297600	10628.57	507	432	17.066
23	2013	525	0.57	43	62	98242	1584.55	224	134	13.918	58	2016	62	0.15	7	29	24415	841.9	54	20	1.913
24	2013	474	0.42	37	61	24753	405.79	14	5	8.554	59	2016	56	0.57	7	31	5336	172.13	1	1	5.696
25	2013	419	0.48	26	65	49107	755.49	79	37	8.29	60	2016	147	0.51	17	25	24094	963.76	7	5	7.179
26	2013	432	0.15	20	56	45374	810.25	83	52	4.09	61	2017	80	0.51	12	18	18010	1000.56	16	14	6.58
27	2013	417	0.24	31	56	12592	224.86	0	0	5.566	62	2017	33	0.15	4	18	10057	558.72	32	30	1.563
28	2013	377	0.3	22	62	87490	1411.13	218	119	7.581	63	2017	34	0.21	9	19	4568	240.42	24	1	2.219
29	2013	371	0.33	18	59	90584	1535.32	57	51	6.271	64	2017	27	0.24	5	19	19137	1007.21	56	52	3.137
30	2013	319	0.27	23	59	413622	7010.54	1881	546	18.367	65	2017	44	0.51	9	10	489920	48992	1871	1706	65.35
31	2014	453	1.29	17	44	305404	6941	926	811	33.926	66	2017	24	0.18	3	18	13821	767.83	3	0	1.221
32	2014	337	0.21	22	55	89823	1633.15	90	85	6.283	67	2017	25	0.36	3	19	84243	4433.84	392	336	11.316
33	2014	306	0.27	11	54	15347	284.2	18	15	3.354	68	2017	27	0.33	7	18	11257	625.39	34	34	3.841
34	2014	374	0.69	31	45	15197	337.71	13	8	10.615	69	2017	23	0.6	4	17	3542	208.35	1	0	5.796
35	2014	265	0.21	17	55	82645	1502.64	5	3	4.098	70	2017	47	0.15	6	19	15099	794.68	0	0	1.343

月被引峰值 C_{top} 、月均使用量 U_t 、首月讨论量 T_1)转化为统一的标准格式,得到 C'_s 、 C'_{top} 、 U'_t 、 T'_1 ,并通过公式6分配的权重,将四个评价分指标综合起来,计算得到70篇高被引论文的影响力综合指标co-influence,见表1第十一列。

4.2 结果分析

(1)首次他引速度 C_s 、月被引峰值 C_{top} 、月均使用量 U_t 均能较好地缓解时间因素对指标数值累积的影响,相对较公平、客观评价不同年龄的论文。观察表1可以发现,随着论文出版年份由2011年至2017年越来越远,70篇样本论文的总被引频次、总使用量逐渐减小。2011年十篇论文的总被引频次分布区间为[445,1705],而2017年十篇论文的分布区间已降到[23,80];2011年

至2017年间,大部分论文的总使用量数值由十万级降到万级再降到千级,各年间论文使用量差别明显。因此,应用总被引频次、总使用量的评价方法使“老文献”更具评价优势,忽略了“新文献”在较短时间内反映的学术价值与影响力。

反观70篇论文的首次他引速度、月被引峰值与月均使用量可以发现,各年份间论文的三项指标数值分布均匀,除月被引峰值随文献出版时间出现较小幅度递减外,其余两项指标数值与文献出版时间并无明显关联。2011年出版的“老论文”与2017年出版“新论文”的首次他引速度均分布在0.15以上;2011年出版的十篇论文月均使用量分布在[386.89,1565]区间内,而2017年出版的十篇新论文月均使用量可高达48992

次。因此,首次他引速度、月被引峰值与月均使用量这三项指标在“新老文献”之间起到了较好地平衡作用,有效缓解了文献出版时间对评价结果的影响。

(2)首次他引速度 C_s 、月被引峰值 C_{top} 与月均使用量 U_i 能有效识别发文时间较短的潜在高价值文献。被引频次相同,发表时间不同的论文,其价值是不同的;同时,在网络平台上被使用次数相同,发表时间不同的论文,其影响力大小也是不同的。其中,发表时间相对较短的“新文献”在未来时间里可能会得到更多被引与关注,是潜在的高价值文献。例如,发表于2017年的论文65,在发表后仅10个月较短时间内,累计PLOS平台总使用量489920篇次,月均使用量也高达48992次/月,远远超出2011年发表的老文献近10倍。另外,在2017年发表的文献61,总被引频次仅为80次,在70篇样本论文里并不占优势,但在其发表后的18个月里,被引峰值高达12次,并且在发表后两个月便被他人引用,未来的价值有待可期。因此,通过首次他引速度、月被引峰值与月均使用量的时间稀释效应,不仅能较公平地评价新老文献,还能识别潜在的高价值新文献。

(3)首次他引速度 C_s 、月被引峰值 C_{top} 这两项引用类指标与月均使用量 U_i 、首月讨论量 T_1 分别反映不同类型的影响力。观察表1可以看出,首次他引速度 C_s 、月被引峰值 C_{top} 、月均使用量 U_i 、首月讨论量 T_1 这四项指标对70篇样本论文的评价结果并不具有明显一致性。利用SPSS软件,对样本数据中的四项指标进行Pearson相关性分析。结果显示: U_i 与 T_1 的相关性系数为0.803,在双尾概率0.01水平上显著相关,而 U_i 与 C_s 、 C_{top} 相关系数分别为0.097、-0.097, T_1 与 C_s 、 C_{top} 相关系数分别为0.331、-0.138,均在双尾概率0.01水平上不相关。因此,从相关性来看,使用类指标与讨论类指标存在正相关关系,而这两者与引用类指标并不相关。

这说明,基于网络平台的月均使用量 U_i 、首月讨论量 T_1 与首次他引速度 C_s 、月被引峰值 C_{top} 评价论文的影响力侧重点不同。引文类指标具有科学评价系统的惯性,引证行为蕴含着施引作者对论文的定性价值判断^[5],因此,首次他引速度、月被引峰值这两项引文类指标体现文献的学术影响力。而基于网络平台的月均使用量、首月讨论量指标主要从论文的交流、利用、讨论情况出发,能够捕捉到论文在广阔的范围和不同的读者群体中的影响力,更侧重于社会关注度,体现论文的社会影响力。

(4)影响力综合指标(co-influence)将论文被引用、被使用、被讨论的动态变化特征、最佳效力发挥时间囊括其中,形成了多维度的动态学术评价体系,相对及时、全面地识别出高影响力、高价值文献。论文影响力综合评价体系将各项指标的最佳效力发挥时间、指标数值时间变化特征纳入其中,用论文首次被引速度、整个生命周期的月被引峰值作为引用类指标,用论文整个生命周期的平均被使用情况作为使用类指标,用论文初期的讨论情况作为讨论类指标。因此,影响力综合指标(co-influence)有效减少了引用时间滞后问题、使用数据可伪性问题、讨论数据快速利用快速老化问题带来的影响,是一个多维度的动态学术评价体系,能相对较客观、及时、全面地识别出高影响力、高价值文献。

以影响力综合指标(co-influence)排序,分析表1中综合影响力最高的前十篇论文(论文序号为65、50、31、1、9、11、30、57、39、23)各项指标,可以发现,这些论文在引用、使用或讨论指标中表现出相对较高的水平,各项指标的排名较靠前,具有极高的影响力与关注度。以影响力综合数值最高的论文65为例,其发表时间距样本数据获取时间仅10个月,但在发表后两个月便被其他学者引用,在第四个月被引用9篇次,月均使用量达到48992次,并且在发表后第一个月累积Twitter讨论量1706次。该论文在短时间内得到了学术界、社会各界广泛的关注,并产生了显著的影响力,是当之无愧的高影响力、高被关注论文。再如2011年出版的论文1,其影响力综合得分高达28.407,虽然由于2011年PLOS平台未收集Twitter讨论数据,导致其讨论指标为0。但其首次他引速度、月被引峰值、月均使用量均较高,在发表后1.3个月便被他人引用,月被引量高达150篇次,月均使用量达1565次,该论文产生的价值在距今七年的时间里不曾间断,是名副其实的高价值论文。

5 结语

在利用Altmetrics指标评价论文过程中,不能孤立地看待某个指标的最终数值结果,而有必要综合考量文献年龄与指标的动态变化特征、规律。因此,本文基于时间因素视角,提出了首次他引速度 C_s 、月被引峰值 C_{top} 、月均使用量 U_i 与首月讨论量 T_1 四项指标,并通过标准化、分配权重,建立论文影响力综合评价体系(co-influence)。同时,以期刊PLOS ONE 2011—2017年的70

篇高被引论文为数据样本,进行实证研究,发现 co-influence 指标是一个多维度的动态学术评价体系,能相对及时、全面、客观地识别出高水平、高影响力文献。

因此,本文构建的影响力综合指标,一方面可避免新文献在评价中的劣势地位,较公平地对新老文献进行评价,有利于优化青年科学家成长的评价环境;另一方面,为数字时代的科学评价提供了一个新的视角,有利于激励科学家追求数量与质量并重的科研成果,并紧跟时代步伐,积极参与、投身各大网络平台、社交媒体,扩大自己科研成果的可见度与传播面。另外,该影响力综合评价方法也可以应用于对文献集合体的评价,例如:对中高龄人才与青年人才之间的评价、对新老期刊的评价等。

作为一项探索性的研究,本文还存在一些不足之处:研究结论是针对开放获取期刊的高被引论文的一般性规律,并非一定适应于所有学科领域或其他非高被引论文,若要验证本文的结论是否同样适用于其他论文,还应进一步基于大样本数据进行实证研究和深入分析。同时,以经验划分指标权重是否合理,是否还有其他更好的划分方法,这些问题在今后的研究中需要深入探讨。

参考文献

[1] 刘运梅,李长玲,刘小慧. 基于合著者贡献大小分配权值的p指数探讨[J]. 图书情报工作, 2016(21):81-86.

[2] 王雯霞,刘春丽. 不同学科间论文影响力评价指标模型的差异性研究[J]. 图书情报工作, 2017(13):108-116.

[3] Priem J, Taraborelli D, Groth P. Altmetrics: a manifesto [EB/OL]. [2016-10-27]. <http://altmetrics.org/manifesto>.

[4] 邱均平,余厚强. 基于影响力产生模型的替代计量指标分层研究[J]. 情报杂志, 2015(5):53-58.

[5] Colquhoun D P A. Why You Should Ignore Altmetrics and Other Bibliometrics Nightmares [EB/OL]. [2016- 10- 21]. <http://www.dscience.net/?p=6369>.

[6] Liu J, Adie E. Five challenges in Altmetrics: a toolmaker's perspective[J]. Bulletin of the American Society for Information Science and Technology, 2013, 39(4):31-34.

[7] Thelwall M, Haustein S, Larivière V, et al. Do Altmetrics work? Twitter and ten other social Web services[J]. PLoS One, 2013, 8(5):e64841.

[8] Ortega J L. Relationship between altmetric and bibliometrics indicators across academic social sites: the case of CSIC's members[J]. Journal of Informetrics, 2015,9(1):39-49.

[9] Bornmann L. Alternative metrics in scientometrics: a meta-analysis of research into three Altmetrics[J]. Scientometrics, 2015,103(3):1123-1144.

[10] 宋丽萍,王建芳,孙 斌. 相关性视角下的 WoS 与 Scopus 之比较[J]. 图书情报工作, 2012(4):22-26.

[11] 余厚强. 替代计量指标与引文量相关性的大规模跨学科研究——数值类型、指标种类与用户类别的影响[J]. 情报学报, 2017(6):606-617.

[12] 赵蓉英,郭凤娇,谭 洁. 基于 Altmetrics 的学术论文影响力评价研究——以汉语言文学学科为例[J]. 中国图书馆学报, 2016(1):96-108.

[13] 由庆斌,韦 博,汤珊红. 基于补充计量学的论文影响力评价模型构建[J]. 图书情报工作, 2014(22):5-11.

[14] 宋丽萍,王建芳,刘 芮. 基于主成分分析的科学评价维度研究——以 PLoS ONE 为例[J]. 图书情报工作, 2014(17):119-124.

[15] Wang X, Fang Z, Yang Y. Continuous dynamic and comprehensive article-level evaluation of scientific literature [J]. Eprint Arxiv, 2014(1411):7004.

[16] 王贤文,方志超,王虹茵. 连续、动态和复合的单篇论文评价体系构建研究[J]. 科学学与科学技术管理, 2015(8):37-48.

[17] 刘运梅,李长玲. 时间序列视角下 PLOS ALM 指标特性识别模型构建与应用[J]. 情报资料工作, 2019,40(5):6-15.

[18] Egghe L. A heuristic study of the first-citation distribution[J]. Scientometrics, 2000,48(3): 345-359.

[19] Egghe L, Bornmann L, Guns R.A proposal for a First-Citation-Speed-Index[J]. Journal of Informetrics, 2011,5(5):181-186.

[20] 王贤文,方志超,胡志刚. 科学论文的科学计量分析:数据、方法与用途的整合框架[J]. 图书情报工作, 2015(16):74-82.

[21] Priem J, Groth P,Taraborelli D. The Altmetrics Collection[I/OL]. [2015-06-23]. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0048753>.

[22] 陈小山,陈国福,张 瑞. 基于因子分析和 SEM 模型的期刊评价指标结构关系研究[J]. 情报科学, 2016, 34(10):61-64.

[23] 邱均平,余厚强. 替代计量学的提出过程与研究进展[J]. 图书情报工作, 2013, 57(19):5-12.

[作者简介]刘运梅,女,1995年生,山东理工大学科技信息研究所硕士研究生。

李长玲,女,1969年生,山东理工大学科技信息研究所教授,硕士生导师。

杜德慧,女,1995年生,山东理工大学科技信息研究所硕士研究生。

收稿日期:2018-08-03