

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

赵蓉英 郭凤娇 谭 洁

摘 要 本文借助 Altmetric Explorer 和 SPSS 统计分析软件获取和分析处理数据,综合运用相关性分析法和主成分分析法,对现有 Altmetrics 评价指标进行定量分析和选取,设计和构建基于 Altmetrics 的学术论文影响力评价框架及模型,并通过实证研究证明评价模型的可行性。通过与传统引文指标评价方法的对比分析,发现 Altmetrics 评价指标同传统的引文指标在评价结果上既具有一致性也存在差异,Altmetrics 测度的影响力偏向社会影响力及社会关注度,较少涉及表征论文质量的学术影响力。基于 Altmetrics 的学术论文影响力评价对传统的基于文献计量的学术论文影响力评价有所补充。图 6。表 12。参考文献 15。

关键词 Altmetrics 学术论文 评价模型

分类号 G250.2

Evaluation of Academic Papers Impact Based on Altmetrics: A Case Study of Chinese Language and Literature

ZHAO Rongying, GUO Fengjiao & TAN Jie

ABSTRACT

Traditional evaluation of academic papers impact is done by analyzing the phenomenon of paper citations. With the development of web 2.0, the online communication is used by more and more researchers to acquire, spread and share research achievements. Altmetrics was born under this background. In this paper, evaluation model based on Altmetrics was constructed to evaluate the influence of academic papers. We hope that it not only overcomes the shortcomings of traditional evaluation methods, but also evaluates the academic papers impact more comprehensively and scientifically by combining with the characteristics of the academic exchanges in the network environment.

In this paper, problems existing in the current study of Altmetrics evaluation indicators were first summarized. Secondly, aiming at the shortcomings and deficiencies of related research, an evaluation model flowchart of academic papers impact based on Altmetrics was constructed, combining with the existing evaluation indicators. Then, all the relevant journal articles about Chinese Language and Literature published in 2012 from Mendeley were selected as the sample documents. Altmetrics Explore and SPSS were used to acquire, analyze and process the data. Synthetically using the correlation analysis and principal component analysis, we quantitatively analyzed and selected evaluation indicators from the existing indicators. Indicators were determined by factor analysis and weights were assigned according to the score

通信作者:郭凤娇,Email:happyguofj@126.com,ORCID:0000-0002-2902-8299 (Correspondence should be addressed to GUO Fengjiao,Email:happyguofj@126.com,ORCID:0000-0002-2902-8299)

coefficient matrix of each principal component. Empirical studies had shown the feasibility and rationality of the evaluation model.

By the comparative analysis with the traditional citation indicators, we find that Altmetrics indicators have consistency in the evaluation results with the citations and h-index, but there are still some differences. Altmetrics evaluation indicators are mainly used to measure the controlled readership of academic papers, the depth of social communication of traditional media and interactions produced by recommendations, comments, etc. It is in favor of social influence and social attention, and less involved in academic impact. Therefore, evaluation of academic papers' impact based on Altmetrics is a complement to traditional evaluation based on citation indicators, and cannot completely replace the traditional evaluation methods.

The feasibility and rationality of the evaluation model based on Altmetrics were verified in this article. A comprehensive model uniting Altmetrics indicators and citation indicators can both measure the social influence and academic impact of academic papers. We look forward to selecting more rational Altmetrics indicators and citation indicators to design a better comprehensive evaluation model, making the evaluation of academic papers impact more comprehensive and effective.

Our research has constructed an evaluation model of academic paper impact based on Altmetrics. As a complement to traditional evaluation model, the model can play the role of predicting highly cited papers. 6 figs. 12 tabs. 15 refs.

KEY WORDS

Altmetrics. Academic papers. Evaluation model.

学术论文是科研人员最重要的科研成果之一,其影响力体现在学术影响力和社会影响力两方面。学术影响力体现了科研人员在所属科研领域中的学术地位以及其研究成果所具有的科研学术价值;社会影响力表现在被学术同行、专家群体外的社会大众所认知和了解的程度。传统的基于文献计量学的学术论文影响力评价通过学术论文的引用和被引用进行分析,但存在时滞过长等诸多弊端。

随着 web 2.0 技术的发展,越来越多的专家学者以及非学术领域群体开始通过在线社交媒体进行学术交流,Altmetrics 正是在这样的时代背景下诞生的。本文基于 Altmetrics 构建评价指标模型,对学术论文影响力进行评价。一方面克服以引文为主要衡量指标的评价方法的缺陷;另一方面结合网络环境下学术交流的特点,基于论文主题内容,更全面、科学地对学术论文进行影响力评价,切实体现其意义和价值。

1 评价模型构建研究

1.1 当前研究存在的不足

关于 Altmetrics 的研究越来越受到关注^[1-4],国内外学者也提出相关评价指标^[5-7],并通过大量的实证研究验证这些指标的可行性^[8-11]。这些指标的提出,不仅拓宽了评价的广度,使评价更全面、具体,而且加深了评价的深度,因为其不仅从学术论文受关注度这个主观性视角去度量学术影响力,更从学术论文内容质量和学术价值方面深层次探讨学术影响力的真正内涵。但是,随着研究的深入,笔者发现这些评价指标不可避免地存在相应弊端,主要体现在以下三个方面。

(1) 缺少一套度量不同数据源的完善的评价指标体系。已有的评价指标体系是根据不同的 Altmetrics 工具提出的,而由于这些研究工具

所属科研机构不同,跟踪监测的数据源也不尽相同。例如,Altmetric.com 针对的是 CiteULike 和 Mendeley 中的学术论文相关指标数据,ALMS 则是针对在 PLoS 上发表的学术论文。这就容易造成误解:评价指标是否是为了检测自己想要检测的数据源而“量身定制”的。如果不能针对所有的数据源进行统一标准的评价,那么评价就会缺乏公平性、科学性,评价工作就会失去意义。

(2) 各指标没有明确的权重。在进行评价时,只有结合实际量化评价对象在不同层面的重要性,才能保证评价的客观性。并不是所有的指标在同等程度地影响学术论文的影响力,也就是说各指标之间的影响程度是存在差异的,需根据不同层级的指标设置相应的权重,不能使用统一标准,而现有评价指标的权重分配不明确。

(3) 各指标间的相关性以及各指标与引用次数的关联程度不明确。现在大多数研究都是

建立在现有的 Altmetrics 评价指标基础之上,而验证这些指标能否作为学术论文影响力评价指标的研究却很少。各指标之间是否存在某种关联关系从而可聚合为某一类指标,反映学术论文在该层面上的影响力,我们不得而知;这些指标对传统学术论文影响力评价指标有何程度的影响,能否作为传统学术论文影响力评价指标的补充,进而更全面、科学、客观地进行影响力评价,还有待我们深入研究。

1.2 评价模型构建流程

针对目前相关研究的不足,笔者结合现有评价指标,构建基于 Altmetrics 的学术论文影响力评价模型,并应用该模型进行实证研究,验证该模型的可行性,以及与传统影响力评价模型相比,评价结果的一致性。评价模型的构建框架如图 1 所示,主要分为三个步骤,即数据源的确定和数据的获取,数据处理,指标和权重的确定。

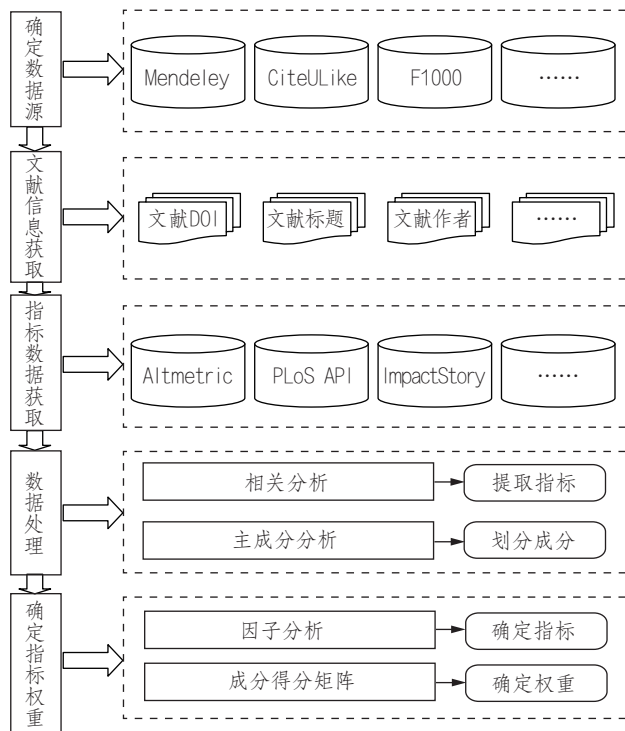


图 1 评价模型构建流程图

(1) 确定数据源并获取数据。由于 Altmetrics 是对学术论文在在线社交网络中基于交流过程的计量,因此这里的数据应包含两部分内容,一部分是学术论文的相关数据信息,另一部分则是学术论文在在线社交网络的交互行为中所产生的计量指标数据。并且这两部分数据必须具有某一文献特征属性,使文献数据源和指标数据源之间建立连接,这一特征属性可以是文献 DOI 等。文献数据可以从专门的文献管理平台获取,而指标数据则可以借助现有的 Altmetrics 工具得到,从而避免因指标数据获取不全、不准确而造成误差。

(2) 数据处理。最终作为研究对象的是各文献记录的指标数据,因此需要对获取的指标数据进行处理,主要包括相关性分析和主成分分析。相关性分析可确定各指标间的关联程度,提取出相关性最强的指标,剔除关联性较弱的指标,从而使指标的代表性更强,评价的意义更突出。主成分分析则可将相关性较强的指标通过因子分析,利用共同因子,聚类成协同影响评价对象的评价层面的影响因子。这些影响因子彼此之间互不相关,是独立的成分,能够综合反映评价对象的不同评价维度。

(3) 确定指标和权重。无论是相关性分析还是主成分分析,最终的目的是确定评价指标,本文最终确定的评价指标就是主成分分析中确定的各主成分。而各个成分对评价对象的影响程度需要

通过权重的赋予加以体现。常用的确定权重的方法主要有层次分析法和模糊综合评价法。

2 实证分析

2.1 数据采集

笔者选择免费开放的文献管理平台 Mendeley^[12] 中的文献作为数据源,选取语言学学科中汉语言文学 (Chinese Language and Literature) 这一学科主题的相关学术论文作为目标文献,指定文献类型为期刊论文,论文发表年限为 2012 年(之所以选择这一年限主要考虑到学术论文从发表到被引用有一定时间差),最终获取到 297 篇文献记录。然后借助 Altmetric Explorer^[13] 获取相关指标数据,这里主要探讨的评价指标包括 News Outlets、Blog、Twitter、Predited、Facebook、Google +、Mendeley、CiteULike、Wikipedia、Vedio、Peerreview。通过所获取文献的 DOI,在谷歌学术中检索得到文献的被引频次与 h 指数。

Schubert 将单篇论文的 h 指数定义为:如果引用某论文的所有论著中有 n 篇论著每篇至少被引用了 n 次,则该论文的 h 指数就是 n^[14-15]。单篇论文的 h 指数不仅能体现论文的直接影

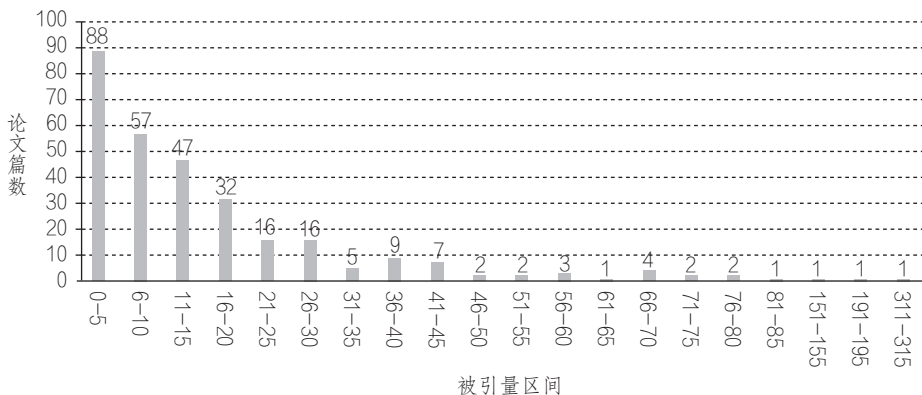


图2 论文被引量统计图

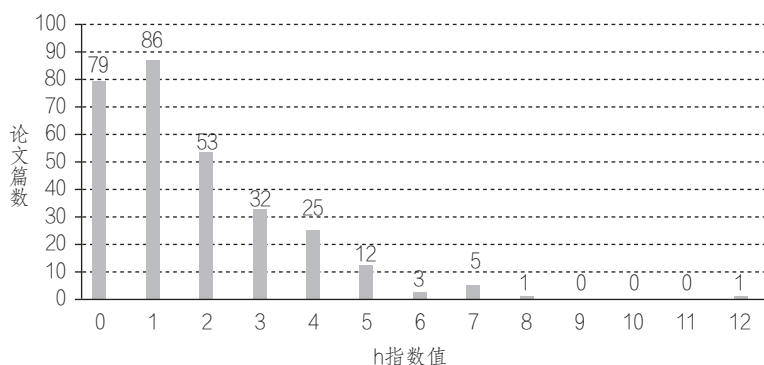


图3 论文h指数统计图

图2与图3的横坐标分别表示被引量区间和h指数值,纵坐标均表示论文篇数,如被引量在0到5次间的论文有88篇,h指数值为0的论文有79篇。笔者采取均值法确定高被引文献,所有文献的被引频次和为5249次,平均篇被引频次为17.67,因此高被引频次的临界值为18。同理所有文献的h指数和为521,平均h指数为1.75,得到高h指数的临界值为2。被引频次大于等于18次的文献有93篇,占总文献数量的31%;h指数大于等于2的文献有132篇,占总文献数量的44%。

通过Altmetric Explorer 不仅可获取各指标数值,还可对目标文献按照各指标要求进行评

价,给出目标文献的Altmetric Score。得分排在前10的文献及其被引频次和h指数见表1。由表1可知,三个排名结果都有一定差异:被引频次最高的文献被引312次,其h指数为12,排在第1位,Total Score却排在第9位;Total Score排在第1位的文献被引39次,h指数为7;Total Score排在第4位的文献,h指数为0;Total Score排在第8位的文献只被引用7次。由此可见,非高被引、非高h指数论文也会有高Altmetric Score值,这就需要对各评价指标与被引量和h指数间的关系进行研究,挖掘出影响被引量的指标。

表1 Altmetric Score值排名前10的文献及其指标值

序号	标题	被引 频次	h 指数	Total Score
1	Growth of Language-related Brain Areas After Foreign Language Learning	39	7	142
2	Positivity of the English Language	37	4	123
3	A Comprehensive Survey of Retracted Articles from the Scholarly Literature	55	3	80
4	Interactive Language Learning by Robots;The Transition from Babbling to Word Forms	12	0	65
5	Music and Early Language Acquisition	26	3	61
6	Second Language Processing Shows Increased Native-like Neural Responses After Months of No Exposure	29	4	60
7	Revisiting Mental Simulation in Language Comprehension;Six Replication Attempts	23	2	43
8	Language Justice for Sign Language Peoples;The UN Convention on the Rights of Persons with Disabilities	7	1	31
9	A Review and Synthesis of the First 20 Years of PET and fMRI Studies of Heard Speech,Spoken Language and Reading	312	12	30
10	Grip Force Reveals the Context Sensitivity of Language-induced Motor Activity During “Action Words” Processing;Evidence From Sentential Negation	14	3	28

此外,从获取到的指标数据来看,总分排名前 10 的文献有 8 篇文献的 News Outlets 指标值为非 0,具体数值见表 2,排在 10 名之后的文献这项指标均为 0,可以假设这项指标具有一定的重要性。另外,在 News Outlets 指标相同的情况下,在 Blog、Twitter 上的提及次数,以及 Mendeley 读者数这几项指标数据越大,对最后得分的影

响程度越大。值得注意的是,要想获得较高的综合得分,必须要在每个较为重要的指标中都获得较高的数值,这也体现了一篇学术论文的综合影响力。那么这些指标中哪些对最后的得分起决定作用,这些指标间是否存在相关性,需要进一步分析,为评价指标的选取以及评价模型的构建奠定基础。

表 2 Altmetric Score 排名前 10 的指标数据表(部分)

序号	Total Score	News Outlets	Blog	Twitter	Facebook	Google+	Mendeley	CiteULike
1	142	1	4	118	6	12	109	0
2	123	3	2	100	4	5	52	2
3	80	2	4	66	3	0	50	6
4	65	2	3	34	3	2	30	0
5	61	1	2	44	7	0	97	2
6	60	1	1	51	1	3	59	0
7	43	0	2	45	0	1	39	0
8	31	3	1	0	0	0	10	0
9	30	0	1	21	3	1	154	6
10	28	1	0	23	1	2	22	0

2.2 数据处理

(1)相关性分析。笔者利用 SPSS 统计分析软件进行皮尔逊(Pearson)相关性分析,衡量各指标与被引量和 h 指数间的相关程度,由此得到与被引量和 h 指数最相关的评价指标,构建评价模型。各指标与被引量和 h 指数间的相关系数如表 3 所示。从表 3 可以看出,Mendeley 读者数与被引量和 h 指数之间的相关系数均高达 0.6 以上;CiteULike 均高于 0.3;Blog、Twitter、Facebook 在 0.2 左右;Wikipidia 与被引量和 h 指数之间的相关系数分别为 0.309 和 0.167;News Outlets 和 Google+ 在 0.1 范围。而 Reditted、Highlighted、Vedio、Peerreview 这四个指标与被引量间的相关系数值均小于 0.1,并且绝大部分文献记录这几个指标值为 0,因此在后续研究中将不再考虑这四项目标。

(2)主成分分析。主成分分析可以将具有相关性的指标归纳组合成彼此独立、共同影响目标对象的不同类别的指标,使同一类别中各指标相关性更强,不同类别间各指标相关性减弱,而且可以找出各类指标中起关键作用的细分指标。笔者对前文中利用相关性分析提取的 Mendeley、CiteULike、Wikipedia、Blog、Twitter、Facebook、News Outlets、Google+八个指标进行主成分分析。进行主成分分析的前提是,各指标或变量间是相互关联的,并且存在的共性成分越多越好,这样才能保证分析的可行性。因此,笔者先利用 SPSS 中的因子分析方法得到各指标的相关系数矩阵(见表 4)。从表 4 可以看出各指标间具有一定的相关性,而且每个指标都有多个强相关的指标。

表 3 各指标与被引量和 h 指数的相关性系数

	Citations			h 指数		
	Pearson 相关	显著性(双尾)	N	Pearson 相关	显著性(双尾)	N
Citations	1		297	0.674 **	0	297
h 指数	0.674 **	0	297	1		297
News Outlets	0.117 *	0.045	297	0.127 *	0.029	297
Blog	0.222 **	0	294	0.270 **	0	294
Twitter	0.190 **	0.001	296	0.265 **	0	296
Reddited	0.022	0.705	297	0.084	0.148	297
Facebook	0.229 **	0	294	0.208 **	0	294
Google+	0.097	0.094	297	0.223 **	0	297
Mendeley	0.601 **	0	297	0.606 **	0	297
CiteULike	0.371 **	0	297	0.342 **	0	297
Wikipedia	0.309 **	0	297	0.167 **	0.004	297
Highlighted	0.066	0.258	297	0.074	0.206	297
Vedio	0.018	0.755	297	0.041	0.483	297
Peerreview	-0.045	0.439	296	-0.081	0.165	296

注: ** 表示相关性在 0.01 水平上显著(双尾)

表 4 各指标间的相关系数矩阵

	News Outlets	Blog	Twitter	Facebook	Google+	Mendeley	CiteULike	Wikipedia
News Outlets	1.000	0.535	0.546	0.432	0.412	0.149	0.111	-0.035
Blog	0.535	1.000	0.696	0.539	0.526	0.455	0.299	0.126
Twitter	0.546	0.696	1.000	0.613	0.765	0.403	0.259	0.007
Facebook	0.432	0.539	0.613	1.000	0.543	0.349	0.176	0.035
Google+	0.412	0.526	0.765	0.543	1.000	0.326	0.037	-0.009
Mendeley	0.149	0.455	0.403	0.349	0.326	1.000	0.332	0.053
CiteULike	0.111	0.299	0.259	0.176	0.037	0.332	1.000	0.176
Wikipedia	-0.035	0.126	0.007	0.035	-0.009	0.053	0.176	1.000

笔者用 KMO 和 Bartlett 的球形度进行检验, 为 0.801。一般情况下 KMO 值在 0—1 之间, 而且越接近 1 说明指标间的共同因子越多, 共性越

强,越容易提取公共因子。这样可以验证本文样本数据并进行主成分分析。

表 5 可行性检验结果表

取样足够度的 Kaiser-Meyer-Olkin 度量		0.801
Bartlett 的球形度检验	近似卡方	884.946
	df	28
	Sig.	0.000

虽然样本数据能够进行主成分分析,各指标间也具有共性,但这些共性能否准确反映各指标的原始信息还需要进行验证。如何选取主成分将各个强关联的指标聚类到一起是关键的核心问题,直接决定评价模型的构建。各指标的公因子方差见表 6,从公因子方差表中可以看到各指标的公因子方差介于 0.4—0.9 之间,表明提取出的公因子能够反映各指标的原始信息。

表 6 各指标的公因子方差表

	初始	提取
News Outlets	1.000	0.511
Blog	1.000	0.713
Twitter	1.000	0.823
Facebook	1.000	0.590
Google+	1.000	0.687
Mendeley	1.000	0.463
CiteULike	1.000	0.621
Wikipedia	1.000	0.434

表 7 是各指标解释的总方差。从解释的总方差表中可以看到,只有两个成分的初始特征值大于 1,而且这两个成分的累积方差较大,达到 60%以上,表明八个原始指标可以划分到两个主成分中。

表 7 各指标解释的总方差表

成分	初始特征值			提取平方和载入			旋转平方和载入
	合计	方差的百分比 (%)	累积百分比 (%)	合计	方差的百分比 (%)	累积百分比 (%)	合计
1	3.592	44.901	44.901	3.592	44.901	44.901	3.513
2	1.249	15.609	60.510	1.249	15.609	60.510	1.536
3	0.892	11.153	71.663				
4	0.735	9.183	80.846				
5	0.534	6.673	87.519				
6	0.482	6.028	93.547				
7	0.342	4.278	97.826				
8	0.174	2.174	100.000				

为了进一步验证主成分划分的正确性,用 SPSS 提供的碎石图进行主成分解析(见图 4)。利用碎石图分析主成分数,主要判断的依据是图中曲线拐点位置。拐点之前包括拐点在内的

点数就是主成分数。从图 4 可以清晰地看到拐点出现的置前区域有且仅有两个点,再一次验证了主成分数。

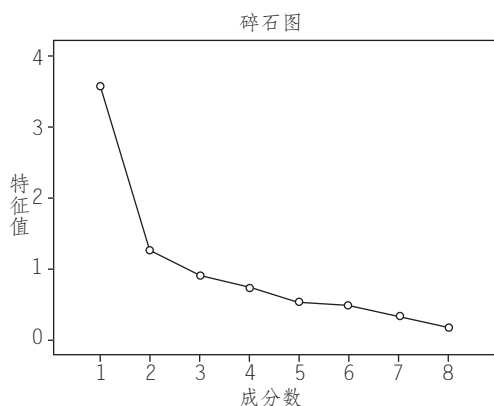


图 4 各指标特征值碎石图

2.3 确定指标和权重

(1) 确定评价指标。在用 SPSS 进行因子分析时,可以通过 Oblimin 算法得到各变量与主成分之间的相关系数矩阵即载荷矩阵,通过系数值的大小可以确定各主成分所应包含的指标。成分结构表见表 8,从表 8 中可以看到 News Outlets、Blog、Twitter、Facebook、Google+ 这几个指标与主成分 1 密切相关,Mendeley、CiteULike、Wikipedia 这三个指标与主成分 2 的相关系数较高。同时还发现主成分 1 中所包含的指标主要是测度学术论文在在线社交媒体和传统媒体中,通过在线社交过程中的交互行为以及媒介传播所产生的影响力,可以认为是媒介交流传播因子;

主成分 2 则是对学术文献所拥有的读者数以及与该学术文献内容相关的搜索查询,从社会人群量的累加上侧面反映影响力,可以认为是读者数量因子。

表 8 各主成分所包含指标的结构表

	成分	
	1	2
News Outlets	0.660	-0.274
Blog	0.839	0.093
Twitter	0.899	-0.123
Facebook	0.762	-0.098
Google+	0.777	-0.288
Mendeley	0.368	0.573
CiteULike	0.356	0.703
Wikipedia	0.078	0.654

由此得到基于 Altmetrics 的学术论文影响力评价模型,如图 5 所示。该模型主要包含媒介交流传播和读者数量两大模块。其中,媒介交流传播模块包含 News Outlets、Blog、Twitter、Facebook、Google+ 这几个指标,而读者数量模块则有 Mendeley、CiteULike、Wikipedia 这三个指标。

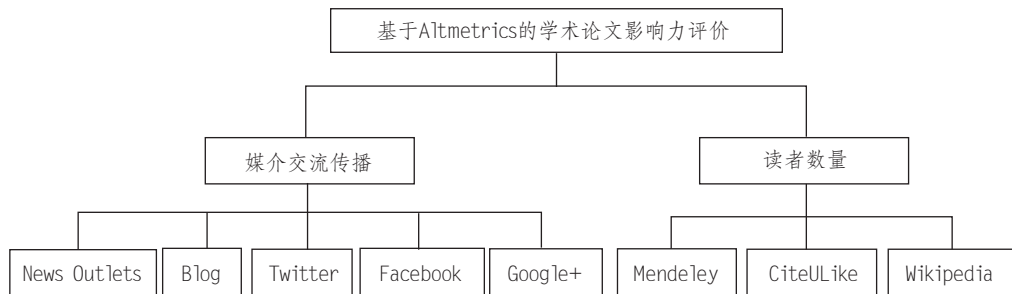


图 5 学术论文影响力评价模型

(2) 确定指标权重。在确定了各项评价指标之后,可对各指标在评价过程中的重要程度进行度量,即分配权重。通过 SPSS 得到各成分

的得分系数矩阵(见表 9)。由此可以写出各成分的得分公式,并将成分因子的方差贡献率占累积贡献率的比重作为各成分因子的权重。笔

者将第一个影响力因子即媒介交流传播记为 AIF_1 , 第二个影响力因子读者数量记为 AIF_2 , 将基于 Altmetrics 的学术论文影响力记为 AIF。

表 9 成分得分系数矩阵

	成分	
	1	2
News Outlets	0.221	-0.141
Blog	0.216	0.152
Twitter	0.264	-0.005
Facebook	0.223	0.000
Google	0.255	-0.140
Mendeley	0.102	0.332
CiteULike	-0.006	0.562
Wikipedia	-0.075	0.498

从表 9 中得到 AIF_1 、 AIF_2 的得分公式为:

$$AIF_1 = 0.221N + 0.216B + 0.264T + 0.223F + 0.255G + 0.102M - 0.006C - 0.075W$$

$$AIF_2 = 0.152B - 0.141N - 0.005T - 0.140G + 0.332M + 0.562C + 0.498W$$

其中, 公式中的各大写英文字母表示各指标, 如 B 代表 Blog 这项指标, T 代表 Twitter, 依此类推。

从表 7 各指标解释的总方差表可以知道, AIF_1 的方差贡献率占累积贡献率的比重为 0.742, 而 AIF_2 的比重为 0.258。也就是 AIF_1 的权重记为 $\lambda_1 = 0.742$, AIF_2 的权重记为 $\lambda_2 = 0.258$ 。最终的学术论文影响力 AIF 的计算公式为 $AIF = \lambda_1 \times AIF_1 + \lambda_2 \times AIF_2$ 。

3 评价结果分析

3.1 各项指标分析

基于 Altmetrics 的学术论文影响力排名前 10 文献的 AIF 值及其 AIF_1 值和 AIF_2 值见表 10。从表中可以看出, AIF 值最高的达到 44.304 分, 比第 2 名高 10 多分; 后续排名间的差距较小, 最多相差 2—3 分, 特别是第 2 名与第 3 名间仅相

差 0.358 分; 第 7 和第 10 名的 AIF_1 值和 AIF_2 值虽然都不高, 但二者都比较均衡, 因此综合排名靠前, 而其他文献的 AIF_1 与 AIF_2 值偏差较大, 总有一个指标值偏高, 从而带动综合排名。由此可以认为, 综合排名的高低不仅取决于单个指标, 指标间的均衡也很重要。

表 10 文献 AIF 值排名(前 10 位)

序号	AIF_1	AIF_2	AIF
1	47.753	34.385	44.304
2	22.281	54.905	30.698
3	34.954	17.069	30.340
4	20.172	49.596	27.763
5	23.712	33.271	26.178
6	24.463	19.968	23.303
7	20.907	18.924	20.395
8	14.889	33.694	19.741
9	12.814	31.996	17.763
10	16.545	12.887	15.601

另外就单个指标而言, AIF_1 值前 10 位的排名结果与综合排名即 AIF 值排名结果在具体的排名次序上有差异, 但也具有一致性; 而 AIF_2 值前 10 位排名结果却与 AIF 值排名结果存在较大差异。由此可见, AIF_1 值在整个评价指标中占据重要地位。

3.2 影响力评价对比分析

AIF 值排名前 10 位的文献及其被引频次和 h 指数见表 11。AIF 值排在第 2 位的文献同时也是被引量最高和 h 指数最高的文献, 且这 10 篇文献都是高被引、高 h 指数文献, 而在 Altmetric Score 前 10 名中则存在非高被引、非高 h 指数文献(见表 1)。这在一定程度上反映了该评价结果与传统基于引文关系的评价结果具有一致性。为了进一步验证这一结论, 笔者进行对比检验, 同高被引论文(被引频次大于等于 18)的确定方法相同, 将 AIF 值大于均值 4.72 的文献确认为高 AIF 值文献。统计得出, 高 AIF 值

的 97 篇文献中有 63 篇高被引文献、72 篇高 h 值文献;132 篇高 h 指数文献中同样也有 72 篇指数文献;93 篇高被引文献中含有 63 篇高 AIF 高 AIF 值文献。

表 11 AIF 值排名前 10 文献及其指标值

序号	标题	被引频次	h 指数	AIF 值
1	Growth of Language-related Brain Areas After Foreign Language Learning	39	7	44.304
2	A review and Synthesis of the First 20 Years of PET and fMRI studies of Heard Speech, Spoken Language and Reading	312	12	30.698
3	Positivity of the English Language	37	4	30.340
4	The Communicative Function of Ambiguity in Language	78	4	27.763
5	Music and Early Language Acquisition	26	3	26.178
6	A Comprehensive Survey of Retracted Articles from the Scholarly Literature	55	3	23.303
7	Second Language Processing Shows Increased Native-like Neural Responses After Months of No Exposure	29	4	20.395
8	Language-selective and Domain-General Regions Lie Side by Side Within Broca's Area	43	7	19.741
9	Bilingual Effects on Cognitive and Linguistic Development; Role of Language, Cultural Background, and Education	78	5	17.763
10	Revisiting Mental Simulation in Language Comprehension; Six Replication Attempts	71	3	15.601

前文列出了基于被引量和 h 指数的论文量统计,这里将 AIF 值划分区间,论文量的分布如图 6 所示。从图 6 中可以看到,基于 AIF 值的论文量分布趋势与基于被引量和 h 指数的论文量分布趋势一致,绝大多数论文的 AIF 值较低,随

着值的增大,论文量减少。AIF 值大于均值的文献有 97 篇,占总论文数量的 33%,这与高被引文献的比例 31% 基本接近,但两者都低于高 h 指数文献的比例。

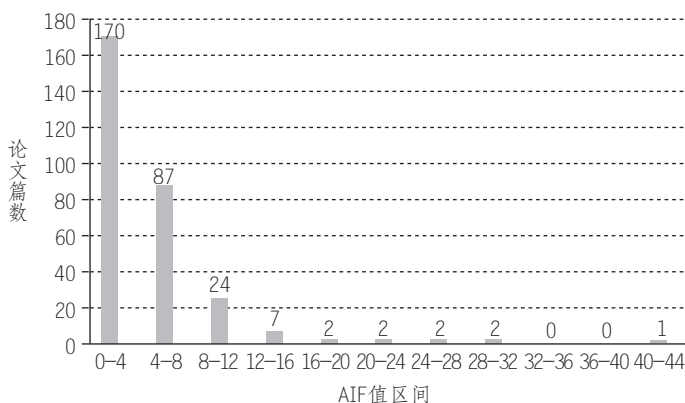


图 6 基于 AIF 值的论文量分布图

此外,笔者通过被引量、h 指数与 AIF 值之间的肯德尔系数检验评价结果的一致性(见表 12),将这三个变量值导入 SPSS 中进行相关性分析,得到基于肯德尔系数的两变量间的相关系数值分别为 0.450 和 0.414,虽然具有一定相

关性,但相关程度不高,仅处于中度相关范围。这也说明,基于 Altmetrics 的学术论文影响力评价对传统的基于文献计量学的学术论文影响力评价有所补充,但还是存在一定差异,不能完全替代。

表 12 AIF 与被引量和 h 指数的相关系数表

			Citations	h 指数	AIF
Kendall 的 tau_b	Citations	相关系数	1.000	0.595 **	0.450 **
		显著性(双尾)	.	0.000	0.000
		N	297	297	297
	h 指数	相关系数	0.595 **	1.000	0.414 **
		显著性(双尾)	0.000	.	0.000
		N	297	297	297
	AIF	相关系数	0.450 **	0.414 **	1.000
		显著性(双尾)	0.000	0.000	.
		N	297	297	297

注:**表示相关性在 0.01 水平上显著(双尾)

4 结论

虽然基于 Altmetrics 的学术论文影响力评价指标能够同传统的引文指标一样,表征学术论文的影响力,并且二者的评价结果在一定程度上具有一致性,但前者不能完全取代后者。

首先,二者对影响力评价的侧重点有所不同。Altmetrics 评价指标主要是从学术文献的读者数,以及它们通过在线社交媒体进行交流过程中的推荐、评论等交互行为和传统媒介的社会传播深度来度量,其影响力偏向社会影响力及社会关注度,而较少涉及与论文质量相关的学术影响力。传统引文指标则直接度量学术论文被其他科研人员引用的情况,表明科研成果的学术价值,是学术影响力的直接体现,但无法体现学术论文的社会影响力。

其次,现有的 Altmetrics 评价指标也存在问

题,通过前文的主成分分析可了解到,现有的诸多 Altmetrics 指标看似比较全面,但没有重点。很多指标由于社交工具、社交形式以及所产生的社交结果不具备大众性,导致相关指标数据不宜参与到评价中,例如,实证中的 Reditted、Vedio、Highlighted 等指标在获取到的数据中比较稀少,只有少数几条文献记录可以参考。因此,在进行评价时我们需选出有综合性和代表性的指标。

最后,基于 Altmetrics 的学术论文影响力评价有其自身的作用和价值。作为对传统影响力评价没有涉猎的社会影响力层面的补充,能够发挥预测高被引论文的作用。通过实证发现,AIF 值(基于 Altmetrics 的学术论文影响力)高的学术论文更加受大家关注,在一定程度上反映了其学术影响力,有效避免了由于假引用、马太效应等造成的高被引现象,为挖掘具有引用价值和高影响力的文献提供了新的视角和途径。

参考文献

- [1] Eysenbach G. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact [J]. Journal of Medical Internet Research, 2011, 13(4): e123.
- [2] Haustein S, Peters I, Bar-Ilan J, et al. Coverage and adoption of altmetrics sources in the bibliometric community [J]. Scientometrics, 2014, 101(2): 1145-1163.
- [3] 黄芳. 补充计量学及其在生物医学领域的应用[J]. 中华医学图书情报杂志, 2014, 23(7): 15-20. (Huang Fang. Altmetrics and its application in biomedical field [J]. Chinese Journal of Medical Library and Information Science, 2014, 23(7): 15-20.)
- [4] 赵蓉英, 汪少震, 陈志毅. 补充计量学及其分析工具之探究[J]. 情报理论与实践, 2015, 38(6): 29-34. (Zhao Rongying, Wang Shaozhen, Chen Zhiyi. An exploratory research on Altmetrics and its analysis implements [J]. Information Studies: Theory & Application, 2015, 38(6): 29-34.)
- [5] Patterson M. Article-level metrics at PLoS—addition of usage data [EB/OL]. [2014-12-12]. <http://blogs.plos.org/plos/2009/09/article-level-metrics-at-plos-addition-of-usage-data>.
- [6] Shuai X, Pepe A, Bollen J. How the scientific community reacts to newly submitted preprints; article downloads, Twitter mentions, and citations [J]. PLoS One, 2012, 7(11): e47523.
- [7] Thelwall M, Haustein S, Larivière V. Do altmetrics work? Twitter and ten other social web services [J]. PLoS One, 2013, 8(5): e64841.
- [8] Mohammadi E, Thelwall M. Assessing non-standard article impact using F1000 labels [J]. Scientometrics, 2013, 97(2): 383-395.
- [9] 刘春丽, 何钦成. 不同类型选择性计量指标评价论文相关性研究——基于 Mendeley、F1000 和 Google Scholar 三种学术社交网络工具[J]. 情报学报, 2013, 32(2): 206-212. (Liu Chunli, He Qincheng. Study on correlation of different Altmetrics indicators for paper evaluation—based on three academic social networking tools: Mendeley, F1000 and Google Scholar [J]. Journal of the China Society for Scientific and Technical Information, 2013, 32(2): 206-212.)
- [10] 刘春丽. 基于 PLoS API 的论文影响力选择性计量指标研究[J]. 图书情报工作, 2013, 57(7): 89-95. (Liu Chunli. Study on the Altmetric indicators of paper impact based on PLoS API [J]. Library and Information Service, 2013, 57(7): 89-95.)
- [11] 王睿, 胡文静, 郭玮. 高 Altmetrics 指标科技论文学术影响力研究[J]. 图书情报工作, 2014, 58(21): 92-98. (Wang Rui, Hu Wenjing, Guo Wei. Research on academic influence of high Altmetrics sci-tech papers [J]. Library and Information Service, 2014, 58(21): 92-98.)
- [12] Mendeley [EB/OL]. [2014-12-10]. <http://www.mendeley.com>.
- [13] Altmetric: we make article level metrics easy [EB/OL]. [2014-12-12]. <http://altmetric.com>.
- [14] 杨建林, 严明. 单篇期刊论文 h 指数的实证研究——以图书情报学为例[J]. 图书情报工作, 2010, 54(12): 145-148. (Yang Jianlin, Yan Ming. Empirical research on Schubert's h-index of a single paper—taking the Library and Information Science as examples [J]. Library and Information Service, 2010, 54(12): 145-148.)
- [15] Schubert A. Using the h-index for assessing single publications [J]. Scientometrics, 2009, 78(3): 559-565.

赵蓉英 武汉大学信息管理学院教授, 博士生导师。湖北 武汉 430072。

郭凤娇 武汉大学信息管理学院博士研究生。湖北 武汉 430072。

谭洁 武汉大学信息管理学院硕士研究生。湖北 武汉 430072。

(收稿日期: 2015-07-07; 修回日期: 2015-07-27)