Postprint: Dynamic Evaluation of Academic Paper Impact in Online Scientific Communication

Authors: Wang Luona, Duanming Li, Li Xing

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] Construct evaluation indicators and models for academic paper impact in online scientific communication, enriching the content and methods of academic paper impact evaluation. [Method/Process] First, analyze the characteristics of indicators in current academic evaluation research, reference the Altmetrics indicator system, select several major authoritative domestic online media platforms as research platforms, and obtain relevant indicator data. Second, screen indicators through correlation with citation frequency and download volume, and construct an evaluation model using principal component analysis. Finally, verify the model's effectiveness through empirical analysis. [Results/Conclusion] Through research and evaluation of academic paper impact in the network environment, improve the singularity and one-sidedness of traditional academic paper evaluation models, optimize the indicator system for academic paper impact evaluation, and provide new ideas and methods for related research.

Full Text

Preamble

Vol. 62 No. 4 February 2018
Research on Dynamic Evaluation of Academic Papers Impact in Online Science Communication
Wang Luona, Li Duanming, Li Xing
School of Economics and Management, Southwest University of Science and Technology, Mianyang 621010

Abstract

[Purpose/Significance] This study aims to construct evaluation indicators and models for academic paper impact in online scientific communication,

thereby enriching the content and methods of academic paper impact assessment. [**Method/Process**] First, we analyze the characteristics of existing indicators in current academic evaluation research and select several authoritative domestic online media platforms as research platforms based on the Altmetrics indicator system to obtain relevant index data. Second, indicators are screened through their correlation with citation frequency and download volume, and an evaluation model is constructed using principal component analysis. Finally, the model's validity is verified through empirical analysis. [**Result/Conclusion**] By studying and evaluating the impact of academic papers in the network environment, this research improves the uniformity and one-sidedness of traditional academic paper evaluation models, optimizes the indicator system for academic paper impact evaluation, and provides new ideas and methods for related research.

Keywords: paper evaluation; impact research; online science communication
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2018.04.014

Introduction

The number of academic papers in China has been growing continuously. According to the 2017 Chinese S&T Paper Statistics Report [1], SCI database 收录的中国科技论文已达 32.42 万篇 in 2016, accounting for 17.1% of the world's total and ranking second globally for the eighth consecutive year. While this growth has promoted academic development, it has also brought an urgent problem: the evaluation of academic papers. Such evaluation should consider not only academic impact but also social impact. Traditionally, researchers conducted academic exchanges through published works, journal articles, and academic conferences, where information could not be shared in a timely manner, and evaluations were mostly based on citation counts. This has resulted in a certain one-sidedness and limitation in traditional academic paper evaluation systems. How to overcome these limitations is one of the issues we consider. We propose to approach this from the social impact of academic papers, examining their influence in the online environment through three aspects: indicator system construction, evaluation model building, and empirical analysis, hoping to enrich the current evaluation system and provide new ideas and methods for related research.

1 Research Status

Research on academic paper impact evaluation can be traced back to the 1950s, when the renowned American information scientist Garfield proposed the concept of citation frequency [2] to evaluate academic paper impact. Since then, numerous experts and scholars have devoted themselves to academic evaluation work, proposing many insights that have promoted the development of research evaluation.

1.1 Research on Academic Impact

Traditional methods for evaluating academic paper impact include peer review and citation analysis. Peer review is widely used but relies heavily on expert selection; experts' professionalism, disciplinary background, and personal preferences all introduce randomness into evaluation results, making subjective factors unavoidable. Moreover, peer review is time-consuming and labor-intensive, gradually becoming inadequate given today's growing volume of academic papers. Although citation analysis avoids subjectivity, it only considers surface-level citation frequencies without exploring the reasons for citation, and negative citations are also counted in frequency statistics, leading to biased results.

In addition to these two methods, scholars have conducted research on various evaluation indicators and methods. Garfield proposed the concept of impact factor in 1972 [3] as an indicator to measure journal quality, but it has the drawback of excessive self-citation. J.E. Hirsch proposed the h-index in 2005 to evaluate researchers' academic output [4], but the h-index fails to adequately reflect highly cited papers. To balance the quantity and quality of citing documents, some scholars [5] constructed evaluation indicators for single-paper academic impact from citation analysis, though this method still cannot avoid limitations arising from different disciplinary attributes. Similarly using citation analysis, Zhao Rongying et al. [6] introduced full-text citation analysis, arguing that it could fundamentally change the development of citation analysis and scientific evaluation. Like peer review, this method is detailed and accurate but time-consuming, making it difficult to implement given the current pace of academic paper production. Other scholars have attempted different approaches; Y.B. Zhou et al. [7] proposed using recursive algorithms to calculate academic paper influence and reputation, suggesting that paper reputation should be given more consideration. Ye Ying and L. Leydesdorff collaborated in 2013 to propose the concepts of academic matrix and academic trace [8], applicable to both single-paper evaluation and evaluation of research groups such as journals and institutions. Tang Jirui et al. [9] applied academic trace and impact rectangle methods to single-work evaluation, comparing them with total citations and h-index, and found that academic trace and impact rectangle provide more comprehensive measurement information.

1.2 Research on Social Impact

Evaluation of academic paper social impact primarily focuses on Altmetrics. Altmetrics is a concept proposed by J. Priem [10] in 2010, referring to the impact of academic papers on social networks, measuring their social influence through the dissemination and communication of academic achievements in the online environment. Since the inception of Altmetrics, numerous scholars have conducted exploratory research. L. Bornmann [11] compared the correlation between Twitter, Mendeley, CiteULike and other indicators with citation analysis, identified the advantages and disadvantages of using Altmetrics to evaluate academic achievements, and recommended combining it with citation analysis.

H.Q. Yu et al. [12] compared domestic Weibo with Twitter, collecting 4.4 million records provided by Altmetric.com, and found that indicators collected from social websites have more obvious immediacy advantages over traditional citation indicators in evaluating academic paper impact. K.H. Chen et al. [13] studied the correlation between scholars' grant applications, awards, and Altmetrics indicators, finding that Altmetrics indicators better reflect scholars' own influence. Zhao Rongying et al. [14] used Altmetric.com tools to construct academic paper impact evaluation indicators for Chinese language and literature, and through empirical research found both commonalities and differences, proposing that Altmetrics-based evaluation indicators can supplement traditional bibliometric indicators. You Qingbin et al. [15] used correlation analysis to screen altmetrics indicators, conducted principal component analysis on indicator data provided by the Mendeley platform, and constructed a paper impact evaluation model.

In summary, existing research on academic paper impact evaluation mostly expands evaluation indicators and innovates evaluation methods, but all have certain one-sidedness. Both domestic and foreign scholars use existing evaluation systems and tools—Altmetrics.com—for research, but there is a lack of a quantitative evaluation system and method for academic paper social impact in the domestic academic environment. Therefore, we propose to start from the social impact of academic papers, referring to the authoritative evaluation tool Altmetrics.com's indicators and platforms, to construct an evaluation indicator system for academic paper social impact in China's online environment, use principal component analysis to build an evaluation model, conduct empirical analysis on the model, and verify the rationality and effectiveness of this method.

2 Construction of the Evaluation Indicator System

Currently, there are five mainstream tools abroad for evaluating paper impact in the online environment, each with different characteristics. Among them, Altmetric.com and PLOS Impact Explorer represent the two tools that track the richest data sources, while Altmetric.com has the most concentrated evaluation objects. Referencing Altmetric.com's data sources [16] better suits this study. Altmetric.com tracks 16 platform data sources; we only list several representative platforms below. Since foreign online environments do not correspond one-to-one with China's, platform selection has slight deviations.

First, we conducted an initial screening of online science communication platforms in China's network environment. The screening criterion was to search for academic papers on each platform and check whether the retrieved content was paper-related, using entries that explicitly mentioned the paper as the standard. After preliminary investigation, we found that major news websites, encyclopedia-type websites, social networking sites like Douban and Renren, document sharing platforms like Aizhike, and Blog Garden 几乎没有对学术论文的分享和提及, so these were not considered in indicator construction.

In the Altmetrics.com tool, News Outlets are foreign news media websites. In China's online environment, mainstream news media rarely mention academic papers, and the academic exchange content mainly involves medical popularization. Therefore, we selected ScienceNet as the news media platform, which covers both news highlights and academic information and paper dynamics, meeting the requirements for tracking paper social impact. For document management platforms, we selected Baidu Wenku, Docin, Doc88, and 360doc. For encyclopedia platforms, we selected MBA Think Tank Encyclopedia to correspond with Wikipedia abroad. For social media platforms, foreign platforms include Blog, Twitter, and Facebook, with Blog tracking over 11,000 academic and non-academic blogs. Considering user numbers and influence, we selected the most representative Sina Weibo, Sina Blog, WeChat Official Accounts, and Zhihu; hereafter, "blog" and "microblog" refer to Sina Blog and Sina Weibo. For academic community forums, F1000 is an online service that recommends research papers based on opinions from top scholars in biology and medicine; we selected DXY.cn and XiaoMuChong, where DXY.cn focuses on biomedical knowledge exchange and XiaoMuChong is a comprehensive academic forum.

Through expert consultation and referring to Yu Houqiang [17] and Zhao Rongying's [18] classification of altmetrics indicators, we propose that the evaluation indicator system be divided into three levels: reading, acquisition, and recommendation, tracking three major measurement indicators of literature on each platform. Specific indicators and platform selections are shown in Table 1 and Table 2 .

3 Empirical Analysis

3.1 Data Acquisition

We used the Chinese database CNKI as the data source. Since there is a time lag between publication and citation, we selected all journal papers published on CNKI in 2015 (retrieval date: August 22, 2017). From each discipline category on CNKI, we chose the 10 papers with the highest citation frequencies, obtained their citation counts and download volumes on CNKI, removed duplicate papers, and obtained a total of 91 valid papers. Based on these papers from CNKI, we tracked specific indicator values for each paper on various platforms, obtained indicator data through web scraping, and manually screened entries that mentioned complete paper titles with high matching degrees. These data served as subsequent research samples.

For traditional bibliometric modeling of the obtained papers, due to large differences in download volume sample data, we used the range transformation method for data normalization, as shown in Formula (1):

$$ij = (Xij - minxj) / (maxxj - minxj) \quad \text{公式 (1)}$$

Where X_{ij} represents sample data, min x_j is the minimum value in the column, max x_j is the maximum value in the column, and X _{ij} is the

normalized data. Weight allocation used the analytic hierarchy process, with experts scoring to obtain judgment matrices, thereby determining weights of 0.75 for citation volume and 0.25 for download volume. The final academic paper impact ranking based on citation analysis is shown in Table 3 (only top 10 papers displayed).

3.2 Data Processing

We normalized the data. Since indicator data had different units, we used the range transformation method for normalization, consistent with the method above. Results are shown in Table 4 (only partial results displayed). Before using data from each platform, we conducted correlation tests to confirm which platform data could be used for subsequent research. Using SPSS data analysis software, we tested correlations between data and citation/download volumes. Pearson test results are shown in Table 5 .

According to Table 5, the platforms most strongly correlated with citation and download volumes are Weibo, Baidu Wenku, and Docin, while the weakest correlations are with Zhihu and MBA Think Tank Encyclopedia, with coefficients below 0.2. Other platforms fall in between. ScienceNet has a correlation coefficient of 0.548 with citation volume but only 0.137 with download volume, showing almost no correlation. Additionally, the correlation between download volume and citation volume in Table 5 is also low, indicating that download volume and citation volume are not directly proportional, and their impact on literature cannot be generalized. The N value of 91 in all cases represents the sample size.

3.3 Evaluation Model Construction

Based on the above analysis, 10 indicators were used for principal component analysis: Weibo, Blog, Baidu Wenku, XiaoMuChong, Docin, 360doc, WeChat Official Account, Doc88, DXY.cn, and ScienceNet. Before principal component analysis, we conducted KMO test and Bartlett's sphericity test to confirm that the data met the conditions for principal component analysis. The results showed a KMO value of $0.876 > 0.5$ and a P-value of 0, indicating significant correlations between variables and that the data sample was suitable for principal component analysis.

Table 6 shows the total variance explained by principal components. The first three components have eigenvalues greater than 1, with a cumulative variance contribution rate of 65.783%, indicating that the two extracted common factors can represent 65.783% of the original variables and basically reflect the original data.

Table 7 shows the rotated component matrix using the varimax method. WeChat Official Account, Blog, Weibo, Baidu Wenku, XiaoMuChong, Docin, 360doc, and Doc88 constitute Component 1, while DXY.cn and ScienceNet

constitute Component 2. These two principal components form the academic paper impact evaluation model.

Based on the proportion of common factor variance extracted in Table 6 to the cumulative total variance, we can calculate the weight of each principal component: $\lambda_1 = 0.794$, $\lambda_2 = 0.206$. From Table 8 component score coefficient matrix and λ values, we can derive the scoring formulas for each common factor component F_1 and F_2, where X_1 to X_{10} represent data sources from Blog, Weibo, and other platforms:

$$F_1 = 0.09 \times X_1 + 0.184 \times X_2 + 0.182 \times X_3 + 0.281 \times X_4 + 0.179 \times X_5 + 0.257 \times X_6 + 0.143 \times X_7 + 0.135 \times X_8 - 0.219 \times X_9 - 0.119 \times X_{10}$$

$$F_2 = 0.145 \times X_1 + 0.014 \times X_2 + 0.009 \times X_3 - 0.282 \times X_4 + 0.04 \times X_5 - 0.185 \times X_6 + 0.083 \times X_7 + 0.014 \times X_8 + 0.511 \times X_9 + 0.415 \times X_{10}$$

The final academic paper impact score calculation formula is: $F = 0.794 \times F_1 + 0.206 \times F_2$.

4 Empirical Results and Analysis

We applied the above evaluation model to the 91 sample papers, calculating each paper's social impact score in the online environment and comparing it with citation analysis. Results are shown in Table 9 (only top 10 papers by impact score F displayed).

Comparison of paper rankings in Table 9 reveals that the overall impact ranking in online scientific communication does not differ significantly from citation analysis results, yet it reflects the problem that some papers have low citation frequency but high social impact. First, except for the second paper, other papers ranked high in F-value do not have high citation frequency or download volume, indicating that high download volume and high citation do not necessarily mean high impact, possibly due to immediacy factors or topic popularity. Second, examining papers with high F-values shows that these papers achieved widespread dissemination and exchange in Component 1, with document sharing platforms like Docin and Baidu Wenku having the highest proportion of paper impact, followed by social media sites like Weibo and Blog, while academic communities like XiaoMuChong and DXY.cn have lower usage rates, mainly serving as literature request platforms. Analyzing Component 2 represented by Zhihu and MBA Think Tank Encyclopedia reveals that these two platforms have very few discussions on academic papers compared to other document sharing websites or online communities. Zhihu focuses on question answering, while MBA Think Tank Encyclopedia is a business management-focused encyclopedia platform. Although currently rarely involving paper-related topics, they may become windows for communication between researchers and non-academic users in the future.

Finally, from the perspective of academic paper development, evaluating social impact will promote exchange between academia and non-academia. Incorporating the online scientific communication process of papers into paper evaluation considers both academic users' information exchange processes and creates new opportunities for non-academic users to understand academic development. Academic development cannot rely solely on researchers; understanding societal needs and listening to discussions on academic issues in the online environment better facilitates contemporary academic development. The academic impact evaluation model addresses the closed and delayed nature of previous academic exchange processes and represents a reference method for future research evaluation.

Through obtaining evaluation indicators for academic paper impact in China's online environment, we studied the characteristics of academic dissemination in online scientific communication and constructed an academic impact evaluation indicator system and model. This model references the Altmetrics indicator system and authoritative research, providing a theoretical foundation. However, since foreign social network environments differ from China's, indicators cannot be applied mechanically. We established a localized indicator system based on previous research, constructed an evaluation model for academic paper impact in online scientific communication, and proposed a new indicator system and model that enriches existing research content and methods. However, this study has some limitations, mainly in data source acquisition: the sample size is not large enough, the problems reflected are not comprehensive, and the manual screening process inevitably introduces some subjectivity. The data sample requires further investigation.

References

[1] Institute of Scientific and Technical Information of China. Status of Chinese S&T Paper Output [EB/OL]. [2017-10-31] http://conference.istic.ac.cn/cstpcd2017/document/正文 2017_2 国际.pdf.

[2] GARFIELD E. Citation indexes for science: a new dimension in documentation through association of ideas [J]. Science, 1955, 122(3159): 103-111.

[3] GARFIELD E. Citation analysis as a tool in journal evaluation [J]. Science, 1972, 178(4060): 471-479.

[4] HIRSCH JE. An index to quantify an individual's scientific research output [J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(46): 16569-16572.

[5] He Chunjian. Construction of evaluation indicators for single-paper academic impact [J]. Library and Information Service, 2017, 61(4): 98-107.

[6] Zhao Rongying, Zeng Xianqin, Chen Bikun. Full-text citation analysis: a new development in citation analysis [J]. Library and Information Service, 2014, 58(9): 129-135.

[7] ZHOU YB, LV LY, LI M H. Quantifying the influence of scientists and their publications distinguishing between prestige and popularity [J]. New journal of physics, 2012, 14(3): 33033-33049.

[8] Xue Fei, Leydesdorff, Ye Ying. Exploration of multivariate indicators for academic evaluation [J]. Journal of Library Science in China, 2017, 43(4): 63-73.

[9] Tang Jirui, Ye Ying. Comparative study on academic trace and impact rectangle of single works [J]. Journal of Library Science in China, 2015, 41(2): 4-16.

[10] PRIEM J. Scholarship: beyond the paper [J]. Nature, 2013, 495(7442): 437-440.

[11] BORNMANNEL. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics [J]. Scientometrics, 2015, 103(3): 1123-1144.

[12] YU HQ, XU SM, XIAO TT. Global science discussed in local altmetrics: Weibo and its comparison with Twitter [J]. Journal of informetrics, 2017, 11(2): 466-482.

[13] CHEN KH, TANG MC, WANG CM, et al. Exploring alternative metrics of scholarly performance in the social sciences and humanities in Taiwan [J]. Scientometrics, 2015, 102(1): 97-112.

[14] Zhao Rongying, Guo Fengjiao, Tan Jie. Research on academic paper impact evaluation based on Altmetrics: taking Chinese language and literature as an example [J]. Journal of Library Science in China, 2016, 42(1): 96-108.

[15] You Qingbin, Wei Bo, Tang Shanhong. Construction of paper impact evaluation model based on altmetrics [J]. Library and Information Service, 2014, 58(22): 5-11.

[16] Altmetric. What outputs and sources does Altmetric track? [EB/OL]. [2017-08-20] https://help.altmetric.com/support/solutions/articles/6000060968-what-data-sources-does-altmetric-track.

[17] Yu Houqiang, Qiu Junping. Theoretical research on stratification and aggregation of altmetrics indicators [J]. Library Journal, 2014, 33(10): 13-19.

[18] Zhao Rongying, Guo Fengjiao. Altmetrics: a new perspective for academic impact evaluation [J]. Journal of Intelligence, 2017, 35(1): 14-18.

Author Contributions

Wang Luona: Data collection and processing, paper writing and revision;
Li Duanming: Paper supervision;
Li Xing: English translation.

Note: Figure translations are in progress. See original paper for figures.