IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 2, 2018, accepted January 21, 2019, date of publication February 4, 2019, date of current version February 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897111

A New Journal Ranking Method: The Reputation Analysis of Citation Behavior Model

LEIBAO ZHANG[1], QI QIAN[1], SHUAI ZHANG[2], WENYU ZHANG[2], AND KUN ZHU[2]
[1]School of Public Finance and Taxation, Zhejiang University of Finance and Economics, Hangzhou 310018, China
[2]School of Information, Zhejiang University of Finance and Economics, Hangzhou 310018, China

Corresponding author: Shuai Zhang (zhangshuai@zufe.edu.cn)

This work was supported in part by the National Social Science Foundation of China under Grant 17NDJC168YB, and in part by the National Natural Science Foundation of China under Grant 51875503 and Grant 51475410.

ABSTRACT The authority of journals is usually a key indicator in guiding people to choose important journals, and the journal ranking is a common way to distinguish the journals' authority. The common ranking methods or models are based on citation data, such as the impact factor (IF), the PageRank algorithm, and the hyperlink-induced topic search algorithm. In this paper, we present a new model, named the reputation analysis of citation behavior (RACB) model, which not only considers the number of citations but also considers the reputation of inter-citation behavior. First, the model fits the function relation of the citation desire index (CDI) based on the citation data of the most reputable journal (named the top journal in the remainder of this paper). Then, the CDI values of the target journals are calculated by using the function relation of the top journal. An improved gray correlation analysis is used to describe the deviation between the CDI and the actual citation rate, named the random citation rate, indicating the inter-citation reputation value of the target journals. Finally, a case study showed that the ranking result of the RACB model has a high similarity with the IF values of 2017, which indicates that the ranking results of the RACB model that considered journal reputation evaluations are more reasonable than those of the traditional PageRank algorithm. This paper proposes a new method of journal ranking from the perspective of reputation evaluation, which can rank the journals more reasonably. It encourages researchers to cite articles more fairly in order to avoid a situation in which the low reputation of the citation behavior affects the quality of the journal that published the article.

INDEX TERMS Journal ranking, citation desire index, random citation rate, improved gray correlation analysis, reputation evaluation.

I. INTRODUCTION

In recent years, knowledge development has intensified academic competition. Journals attract the attention of research institutions or scholars by enhancing the journals' authority. On the other hand, scholars also need to know the authoritative value of related journals in specific fields to understand the frontiers in related fields. In general, more authoritative journals are more likely to publish high-quality articles. Therefore, both institutions and individuals are very concerned about the results of journal rankings.

Until now, there have been many indicators and methods of measuring journal authority. As an index that can be used to measure the average citations of articles, the impact factor (IF) is gradually used by scholars in the field of journal ranking [1]. Journal Citation Reports (JCR) calculates and reports on the IF value of journals indexed by Science Citation Index (SCI) or Social Science Citation Index (SSCI). In 2007, Thomson Reuters (now Clarivate Analytics) added a new indicator to JCR, named the five-year IF, which was calculated on the basis of five years citation data of journals. Hirsch [2] put forward the h-index to measure the academic achievements of journals or authors. The h means that the journal or author has published h articles that were cited at least h times. Some studies have shown that the h-index has a great relationship with the area in which it is located.

In addition to these indicators based on citation data, some scholars applied the PageRank algorithm [3] and the Hyperlink-Induced Topic Search (HITS) algorithm [4] to journal rankings. The PageRank algorithm considers not only

The associate editor coordinating the review of this manuscript and approving it for publication was Neil Yuwen Yen.

the number of journals cited but also the influence of the cited journals. The Scimago Journals Rank (SJR), another important indicator, is calculated by the PageRank algorithm. Compared with the PageRank algorithm, the HITS algorithm considers both the influence of the citing journals and the cited journals.

Although there are many indicators and methods to measure the influence of journals, highly influential journals do not necessarily have a good reputation. Some scholars have studied the reputation of journals. Bergstrom [5] considered that people should evaluate the suitability of the citation data when using it, and make good use of the citation data as much as possible. Ioannidis [6] proposed that inappropriate self-citation eventually affects the accuracy of the measurement, and explored the measures to alleviate in the adverse effects of self-citation. In addition, the reputation of journals can be obtained by peer review, but peer review is mainly based on the subjective judgment of the reviewer and lacks specific data support.

With regard to reputation evaluation, there are relatively mature evaluation methods in the field of e-commerce. Yan *et al.* [7] proposed some behaviors that produce untrue data named a ''collusion collusive,'' which means that some online users perform a series of behaviors that go against their wishes, and are even illegal, for their benefit. For example, in some rating activities, the conspirators ignore the facts, give positive comments to their cooperators, but give negative comments to their competitors. If the initial data is untrue, then the credibility of the results is clearly questionable. When identifying the ''collusion collusive'' phenomenon in the field of e-commerce, it will reduce the reputation of the online users. If a similar ''collusion collusive'' phenomenon appears in a journal citation, we can learn from it to reduce the credibility of the corresponding journals. The key question is how to calculate the reputation values of journals. Wang *et al.* [8] proposed a high-reliability multifaceted reputation evaluation mechanism to solve an incomplete feedback and complicated rating and even identify malicious collusive raters. Many institutions and scholars have studied the reputation evaluation of self-citation behavior, but there are few studies on the reputation of journals' inter-citation behavior.

This study aims to explore the reputation of journals from the perspective of inter-citation behaviors, which makes up for the lack of specific data support in peer reviews, and complements the journal's self-citation reputation. We proposed a new model, named the reputation analysis of citation behavior (RACB) model, to evaluate the inter-citation behavior of journals. The journal reputation can be reflected from the difference between its citation desire index (CDI) and its actual citation rate (ACR). Reputable journals have good motivation for citing articles to ensure the quality of the citation behavior. We consider the reputation evaluation of inter-citation behavior as a new enhancement to the traditional ranking method, and analyze the rationality of the reputation evaluation of inter-citation behavior by comparing the experimental results of various ranking methods.

The remainder of this paper is organized as follows. Section 2 summarizes the analytical results by some scholars about journal rankings, citation analysis, and reputation evaluation. Section 3 describes the proposed journal ranking model based on citation reputation. Section 4 discusses the experimental process and an analysis of the experimental results. Section 5 is the conclusion of the article.

II. RELATED WORK

To avoid the disadvantages of inefficiency and subjectivity in judging the quality of articles, Gross and Gross [9] proposed that a librarian select appropriate journals based on citation data. Cartter and Sawyer [10] mentioned earlier methods of evaluating rankings, but these methods are not mature enough.

A. TWO MAIN METRICS

Kademani and Kalyane [11] studied the publication productivity of an individual scientist by using the IF index. Scholars have reordered the journal rankings in various fields and subjects according to IF. Garfield [12] studied some aspects of IF, including how to calculate the IF, the sources of IF, how IF affects the ranking of journals, and the ways to improve IF. However, Seglen [13] argued that IF is not necessarily a good indicator of journal rankings because it does not represent the citation relation well. For example, there are 40 or 50 references in some articles, and some have only 10, therefore the quality of the references is different.

Maslov and Redner [14] proposed the IF based only on the citation analysis had intrinsic limitations: they gave the same weight to all references, and did not consider the quality of the citation. With the introduction of the PageRank and HITS algorithms, scholars made some improvements using these methods. For example, Ding *et al.* [15] proposed a weighted PageRank algorithm and compared the influence of different damping factors on author rankings. Yan and Ding [16] considered a combination of citation and network topology based on a weighted PageRank algorithm, and compared it with other indexes such as the h-index. The results showed that the weighted algorithm has higher stability under different damping factors by comparing with other traditional methods. Su *et al.* [17] used PrestigeRank algorithm to study the missing data of PageRank, and they ranked journals in the physics field by using the PrestigeRank algorithm and PageRank algorithm. They found that the PrestigeRank algorithm had better robustness according to the results. Zheng *et al.* [18] proposed fuzzy clustering for webpages based on the traditional PageRank algorithm, and the proposed new algorithm not only decreased the time but also improved the accuracy of the search.

In December 2016, Elsevier B.V. proposed a practical indicator, CiteScore (CS; https://journalmetrics.scopus.com/), which is similar to IF. As the two main indicators to measure the journals' impact, the IF and CS have similar computational rules and coverage, but there are also obvious differences between them [19]. IF is calculated based on two-year

citation data, while CS is calculated based on three-year data. In addition, the number of journals, document types and access to computing the citation data between them are also different.

B. OTHER METRICS

Currently, there are more and more indexes and algorithms for evaluating journals, based on the spirit of rigorous and responsible attitude. People should consider these indicators as a whole. Yu *et al.* [20] divided journal evaluation indicators into three categories: journal impact, timeliness, and journal features. They established a structural equation about the three indexes. Moed [21] introduced the concept of index comparative reports and analyzed the similarities and differences of the indicators so that scholars could use these evaluation indicators more accurately. Cheang *et al.* [22] proposed two multidimensional journal evaluation methods based on the citation.

Among various methods for evaluating journals, meta ordering is popular because of its large number of information sources. Vana *et al.* [23] described a method for creating meta-rankings by using heterogeneous journal ranking. The greatest advantage of meta-ranking is its ability to find journals with similar quality. When researchers conduct interdisciplinary scientific research (IDR), we need an index that is comprehensive and complex. Wagner *et al.* [24] argued that the entropy were the appropriate indicator and provided some insights of measuring IDR. Maity and Hatua [25] proposed a journal quality point model, which contains the physical presentation, reference research, and citation analysis of the three methods for journal evaluation. Yu *et al.* [26] proposed a multiple-link, mutually reinforced journal-ranking (MLMRJR) model. This model contained intra-networks and inter-networks, and combined the PageRank algorithm and HITS algorithm to determine the results of a ranking. In addition, Song *et al.* [27] focused on the relationship between time-varying and journal rankings, and established a journal-time-topic model to study the trends of journal ranking over time. Later scholars were afraid whether the authority value of the journal was unfair based only on the citation data. Jiang *et al.* [28] put forward the question of whether citation analysis is a legitimate evaluation tool, because there are many negative references cited and excessively abnormal self-citations. Each year, Thomson Reuters presented some indexes about citation data, including the self-citation rate and coupling-citation rate. As a result, the journals that had high self-citation or high coupling citations were suppressed or even eliminated.

Garfield [29] proposed a journal ranking method named MutualRank, which ranks researchers, articles, and venues by considering the network relationships among researchers, articles, and venues. Social web metrics, also called ''altmetrics'', were proposed by Priem *et al.* [30]. It can quickly gain the influence of articles by analyzing the comments and bookmarks of related articles on the Internet. The altmetric score is based on network data, which can identify highly-cited articles more accurately than general journal ranking methods [31]. However, with the passage of time, the network data will become more and more ineffective, and the measurement accuracy of MutualRank or altmetric will decrease.

Peer review is another way to evaluate journals. Peer colleagues evaluate the quality of each other's papers, or the quality of the journals. Peers usually hold universal values and norms, which are the basis for them to evaluate their peers [32]. When these norms and values are applied fairly, peer review is understood to be fair [33]. Peer review has always been the principal means for most scientific disciplines to control research quality, but it has been criticized and questioned because peer review lacks unified quantitative standards and relies more on the subjective ideas of the judges. Helmer *et al.* [34] studied gender bias in peer reviewing and found that stronger female participation would increase the quality of scientific research output.

C. THE PROPOSED MODEL

The RACB model proposed in this study is a method combining the traditional ranking method and the journal's inter-citation reputation. On the one hand, the traditional ranking method only considering the citation counts has limitation in representing reputation information [35], because they lack the credibility analysis of citation behavior. On the other hand, the traditional peer reviews are largely dependent on the subjective consciousness of the reviewers. The proposed RACB model not only combines the reputation evaluation of citation behavior, but also uses quantitative data analysis to produce the more objective evaluation results.

The RACB model proposes a reputation evaluation system for a journal's inter-citation behavior. The main idea of this model is that there are two directed CDIs between any two journals, and the CDI is a reasonable citation rate, in theory. The greater the difference between the ACR and CDI of a journal, the lower the reputation value obtained by using the RACB model. The specific details are discussed in the Methods section.

III. METHOD

The RACB model is mainly based on the study of journals' inter-citation behavior to evaluate the reputation of journals. We establish some criteria for journal citation behavior based on the literature analysis, and calculate the reputation value of journals according to these criteria. The final credit value is normalized from 0 to 1, and the closer the reputation value of the journal is to 1, the more consistent the journals are with the citation criteria. Then credit value is used as an important factor in the journal-ranking formula.

The RACB model studies the reliability of journal' inter-citation behavior based on the statistical data of JCR. Oh *et al.* [36] extracted and analyzed 22 motives and 21 factors of citation behavior through the statistical analysis of bibliographic information, then based on the data of the questionnaire, they found that the most important citing motivation is supporting, and the journal reputation is considered

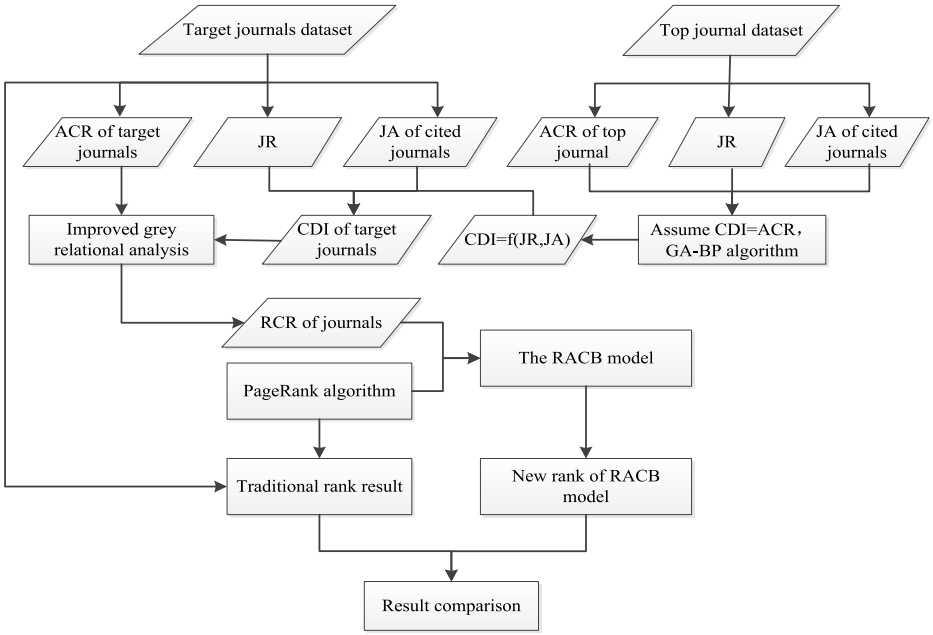


FIGURE 1. The flow chart of the RACB model.

to be the most influential factor. Sendhilkumar *et al.* [37] proposed a new method for calculating the quality score of the articles, whose most important step is to use the cosine similarity to calculate the content correlation between the source article and the cited article. The content relevance between the articles is the primary criterion when references are cited. In addition, scholars also consider the influence of the cited articles (famous authors or high-impact journals). In general, the main motivation of article citation is the content relevance between the articles and the influence of the cited article. The main citation motivation among the journals is similar to that among the articles, that is, the journals relevance and the authority of the cited journals. So, we provide some basic assumptions for the model:

(1) The reputable author cites the articles mainly by considering the relevance of the cited literature with this article and the journal influence of the cited article.

(2) When other conditions are the same, the author is always more willing to refer to articles more relevant within their own fields.

(3) When other conditions are the same, the author is always more willing to refer to articles which appear in a high-impact journal.

(4) For two articles of relevance and where the influences of the corresponding journals are similar, the authors cite them with the same probabilities.

In order to evaluate the reputation of journals, this study introduces the concept of CDI to describe the possibility of a reputable journal to refer to another journal article, CDI is an index that is determined by the journal relevance (JR) and journal authority (JA). Then, we use a random citation rate (RCR) to represent a journal's credibility of reference. Finally, we add the RCR as a new factor to the traditional journal ranking model, and compare the ranking results with those of the traditional ranking method. A flow path of the RACB model is shown in Fig. 1.

A. CITATION DESIRE INDEX

The CDI is used to indicate what kind of reference preferences should be used. It is not mixed with any social emotional elements. Based on the basic assumptions, the value of CDI_{ij} is computed by formula (1):

$$CDI_{ij} = f\left(JR_{ij}, JA_j\right), \tag{1}$$

where i, j denote two journals, JR_{ij} denotes the degree of correlation between i and j, JA_j denotes the JA of journal j, and CDI_{ij} represents the theoretical probability of journal i citing the papers in journal j.

This study aims to calculate the difference between the CDI and the ACR of journals: the greater the difference, the lower the RCR and the lower reputation of the citing behavior, and vice versa. To determine the relationship and related parameters between the CDI, the JR, and the JA, a fitting study should be applied based on the citation data of a reputable journal.

This study fits the predictive function of the CDI based on the relationship between the ACR of top journal, the JR, and the JA of the cited journal (not the top journal).

1) CALCULATION PROCESS OF JR

The aim of this sub-section is to study the calculation process of JR. Salton and Buckley [38] proposed a similarity

calculation between the search query Q and the article D, as shown in formula (2):

$$similarity\,(Q, D) = \frac{\sum_1^t w_{qk} w_{dk}}{\sqrt{\sum_{k=1}^t \left(w_{qk}\right)^2 \cdot \sum_{k=1}^t \left(w_{dk}\right)^2}}, \quad (2)$$

where t represents the number of terms in query Q, and w_{qk} and w_{dk} represent the weight of the term k in query Q and article D, respectively.

For the weight calculation of term t in article d, Zhang *et al.* [39] used TF-IDF [40] to calculate the weight, as shown in formula (3):

$$w_{t,d} = f_{t,d} \cdot log_2 \frac{N}{df_t}, \quad (3)$$

where $f_{t,d}$ is the number of times term t appears in article d, N is the total number of articles in the database, and df_t is the number of articles containing the term t.

The method of calculating JR in this study is borrowed from the idea of formulas (2) and (3). This study directly makes the terms of the entire journal into a term corpus, and considers the importance of the word in the papers, to jointly measure the correlation between the journals. The specific steps are as follows:

(1) All articles of the journals are broken down into lots of terms with frequency indicators by computer program coding. A term frequency-journal matrix $dataset_{m\times n}[a_{ij}]$ is formed, where m represents the total number of terms, n represents the total number of journals, and element a_{ij} of the matrix indicates the frequency of the term i that appears in the journal j. Formula (4) avoids the error caused by the difference in the length of the journals:

$$tf_{ij} = \frac{a_{ij}}{\sum_{i=1}^m a_{ij}}, \quad (4)$$

(2) To give each term a corresponding weight by using the Inverse Document Frequency (IDF), the calculation formula is as follows:

$$idf_{ij} = log_2 \frac{n}{df_i}, \quad (5)$$

where the n is the total journals, and df_i is the number of journals containing the term i.

(3) When the matrix element is weighted, the column vectors $\vec{j}_j$ of the matrix represent the corresponding journal j, and JR can be calculated by calculating the distance between two vectors. The formula is as follows:

$$JR_{j_1 j_2} = similarity\left(\vec{j}_1, \vec{j}_2\right) = \frac{\vec{j}_1 \times \vec{j}_2}{\left|\vec{j}_1\right| \left|\vec{j}_2\right|}, \quad (6)$$

Finally, we can get the JR values between the top journal and the cited journals.

2) CALCULATION PROCESS OF JA

Although the IF reflects results that are not accurate enough, and there are many other indicators to measure the authority of the journal, most scholars still use the IF as the main measure because it has been widely recognized. For example, Poria *et al.* [41] discussed the journal IF on tourism research, and showed the growing importance of IF to journals. This study also directly uses the current IF as the authoritative value of the journal.

3) FITTING PROCESS OF CDI FORMULA

In this study, we need to fit the functional relationship between CDI, JR, and IF based on the citation data of the top journal. Traditional data fitting such as polynomial fitting needs to know the general functional form, which is not suitable for this study. A back-propagation (BP) neural network [42] trains a neural network by using the input and output data of the system, making the BP neural network a black box that can express the unknown function, and then predicts the output target with a trained network.

However, a BP neural network has the disadvantages of slow learning speed and no significant global optimization effect. Lv *et al.* [43] predicted surface subsidence in backfilled coal-mining areas based on the GA-BP algorithm. This hybrid algorithm combined the global optimization characteristics of the genetic algorithm and the fast convergence characteristic of the BP algorithm, and the algorithm has been widely applied in many areas.

In this study, we fit the functional relationship between CDI, JR, and IF using the GA-BP hybrid algorithm developed by Lv *et Al.* [43].

B. RANDOM CITATION RATE

The RCR is an indicator of the degree of reputation used to measure a journal's inter-citation behavior. The first part of the model obtains the predictive function relation formula of CDI, JR, and IF based on the top journal data by the GA-BP hybrid algorithm, and then uses the predictive function to calculate the CDI of these journals. Here, a central hypothesis of this study is as follows: for a fully reputable journal, it actually refers to the literature of the data distribution, which should be very close to the CDI (to allow for a small amount of error). When the difference between theoretical and actual data is too large, we have reason to believe that the author has taken into account other relevant social factors when he or she considered the paper's citation.

First, we can calculate the ACR of each journal based on the citation data over the years in target journals. Then, according to the mathematical formula of the former fitting, the theoretical CDI of each journal to other journals is obtained. The closer the ACR value of a journal to its CDI value, the higher the RCR value of the journal. In other words, the RCR represents the degree of proximity between the ACR and CDI.

From a journal's perspective, it has 10 years of ACR data and CDI data. We can use the method of gray relational analysis with CDI as the reference sequence and ACR as the comparative sequence. Their correlation degree stands for the RCR value of the journal.

The traditional gray relational analysis adopts the equal right treatment method when calculating the correlation degree by formula (7):

$$\gamma_{0i} = \frac{1}{n}\sum_{k=1}^{n} \frac{\min \Delta_i(k) + p \max \Delta_i(k)}{\Delta_i(k) + p \max \Delta_i(k)}, \tag{7}$$

where γ_{0i} is the relativity between comparison sequence i and the reference sequence, $\Delta_i(k)$ represents the difference number sequence of the comparison sequence i and the reference sequence, n is the number of elements of a comparison sequence, and p is the resolution coefficient that weakens the influence of the correlation coefficient distortion caused by a too-large max $\Delta_i(k)$.

This takes the same approach toward all factors, but ignores some factors that are in reality more important. Some scholars considered the result is not accurate or objective, and modified it as formula (8):

$$\gamma_{oi} = \sum_{k=1}^{n} w_{oi}(k) \frac{\min \Delta_i(k) + p \max \Delta_i(k)}{\Delta_i(k) + p \max \Delta_i(k)}, \tag{8}$$

where $w_{oi}(k)$ is the weight of each factor. Here, these factors are mainly about the time of the years. Our assumptions based on the time-aware approach are as follows:

(1) The closer the time distance, the greater the influence. For example, when we calculate the 2017 RCR value of a journal, the data from 2016 has a greater impact on the results than the data from 2015.

(2) Data with equal time intervals have the same influence. For example, the influence degree of the 2015 correlation coefficient on the RCR value of 2016 is the same as the influence of the 2016 correlation coefficient on the RCR value of 2017.

For this purpose, we design a square weighting function to assign the value of weights as formula (9):

$$w_{0i}(k) = \frac{(Y_k - 2000)^2}{\sum_{k=1}^{n} (Y_k - 2000)^2}, \tag{9}$$

where Y_k represents the years, and $Y_k = 2006 + k$.

Thus, the RCR value of journal i can be calculated by formula (10):

$$RCR_i = \gamma_{0i} = \sum_{k=1}^{n} \frac{(Y_k - 2000)^2}{\sum_{k=1}^{n} (Y_k - 2000)^2} \times \frac{min\Delta_i(k) + p\, max\Delta_i(k)}{\Delta_i(k) + p\, max\Delta_i(k)}, \tag{10}$$

C. JOURNAL RANKING

The traditional method of journal ranking is to construct a journal citation network $G(V, C)$, which shows a reference relation between journals. V is a collection of vertices that represent various journals, and C is the directed edges that represent the number of citations and citations by other journals. Then, we can calculate the PageRank values of each journal by formula (11):

$$PR_i = \frac{1-d}{N} + d \sum_{i,j \in V} \frac{n_{ji} PR_j}{N_{j.out}}, \tag{11}$$

where PR_i represents the PageRank value of journal i, PR_j represents the PageRank value of journal j that cites the articles of journal i, d represents the damping factor, N represents the total number of journals that need to rank, n_{ji} represents the number of times journal j cite journal i, and $N_{j.out}$ is the number of out-links of journal j.

Based on the PageRank algorithm, the RACB model considers the evaluation of journal citation behavior. The RACB model is expressed as formula (12):

$$RACB_i = \frac{1-d}{N} + d \sum_{i,j \in V} \frac{n_{ji} RCR_i PR_j}{N_{j.out}}, \tag{12}$$

where RCR_i is the RCR value of journal i.

The proposed RACB model is more complicated and requires more computational cost than the traditional journal ranking methods. However, in practical application of the proposed model, the model complexity and computational cost is worth the gain, because of following facts: (1) The RACB model ranks journals from the perspective of reputation evaluation of journals' inter-citation behavior, and the ranking results of RACB model are more reasonable than the traditional ranking methods for the same year data, as will be verified in the subsequent case study. (2) At present, the costs of computer hardware resources have been reduced to some extents, and with the development of cloud computing, the cost of time and effort in the required computation will be reduced exponentially.

IV. CASE STUDY

A. DATASET

In this section, we rank journals whose papers are related to ethics as a case study. This case is focused on the study of ethics-related journals. The ethics theme includes medical ethics and business ethics. The keywords of the subdomains contain the term "ethics," so the retrieval method is executed by typing "ethics".

Some scholars choose specific journals in a field to experiment on and demonstrate their methods. For example, Biljecki [44] analyzed 12,436 papers of 20 GIScience journals in the period 2000–2014, and obtained a good experimental result. First, we selected 10 target journals with different characteristics based on the IF value ranking and the publication volume of 10 years so that they can better represent the entire theme of ethics journals. The dataset came from the ISI Web of Science database, and was downloaded on September 9, 2017. The dataset contains 7,682 articles, which are listed in Table 1.

Among the journals of ethics, Muzur [45] proposed that the American Journal of Bioethics is considered the first in the "ethics" and "medical ethics" categories. This is not

TABLE 1. Published information of target journals.

Journal	Number of publications	Percentage
Bioethics	757	10.07%
Developing World Bioethics	262	4.34%
Hastings Center Report	1198	8.10%
Health Care Analysis	275	4.56%
Journal of Bioethical Inquiry	627	10.35%
Journal of Law Medicine and Ethics	1046	13.83%
Journal of Medical Ethics	2176	26.54%
Journal of Medicine and Philosophy	418	6.92%
Kennedy Institute of Ethics Journal	208	3.45%
Science and Engineering Ethics	715	11.84%

only because the journal has a high IF value but also because high-quality articles about bioethics are usually published in this journal. Fan [46] ranked the journals of the History and Philosophy of Science category based on journal influence, and the results showed that the American Journal of Bioethics is the number-one-ranked journal.

This study extracted data from the American Journal of Bioethics (called the top journal), which cited other journals from the dataset, and made a preparation experiment to fit the function relation of the CDI.

B. EXPERIMENT

The RACB model adds a citation reputation evaluation to the journal ranking. We compare journals' theoretical CDIs with ACRs. The smaller the difference, the greater the RCR and the better the credibility. Our focus is on the definition of reasonable inter-citation behavior, that is, an author will mainly consider the relevance between the two journals and the authority of the cited journal when he/she is citing articles.

1) DATA FITTING OF CDI

To obtain the predictive function relationship between CDI, JR, and IF, this study assumed that the citation behavior of a top journal is completely reasonable. The predictive function formula is fitted according to the actual citation data of the top journal. This sub-section adopts the GA-BP hybrid algorithm to fit the predictive function of the CDI, and then we can calculate the CDI values of the target journals.

The curve-fitting function has two inputs and one output. Therefore, we set the BP neural network structure to 2-5-1, which means that the input layer has two nodes, the hidden layer has five nodes, and the output layer has one node. Thus, we have $2 \times 5 + 5 \times 1 = 15$ weights and $5 + 1 = 6$ thresholds.

We downloaded all records of the top journal (2007–2016). After calculating the JR, JA, and ACR, the 10 years of data were obtained, containing the input (JR, IF) and output (ACR). In the GA-BP algorithm, this study took the eight years of data from 2007–2014 as training data, and took the two years of data from 2015–2016 as testing data, to get the predictive function of the CDI. The errors of the curve-fitting results are shown in Fig. 2.

In Fig. 2 (a), the sample in the X-axis represents the testing data, and the function output in the Y-axis represent both predicted output (fitted values) and actual output (actual values). In the four subgraphs of Fig. 2 (b), the X-axis represents the actual output, the Y-axis represents the predicted output. The system has done the regression analysis of four kinds of data (expressed in fitting curves of different colors), and normalized all data within -1 to 1. Regression R values measure the correlation between predicted outputs and actual outputs. When the fitting effect is the best, the regression curve should coincide with the diagonal line ($R = 1$).

2) INTER-CITATION REPUTATION

The CDI values between the target journals can be obtained from the previous section. The next task is to measure the difference between the CDI and ACR in various journals through an improved gray relational analysis. The result (RCR value) is expressed as a value between 0 and 1.

To measure the difference between the CDI and ACR, this study made some improvements in the traditional gray relational analysis, as follows: (1) The input data was not sequences of one dimension but two groups of two-dimensional sequences by vectors; we used the Euclidean distance formula to calculate the difference of sequences. (2) The traditional gray relational analysis does not consider the weights of different factors; here, we considered the influence of the factor about the years. A closer year from the present gave more weight to the result.

Other parameters in this section contained p and d, p is the resolution coefficient in formula (10) and d is the damping factor in formula (11). To reflect the better experimental results, this study made an adaptive analysis of the p value, and compared the experimental results when $p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$, and 0.7 respectively.

In this study, we compare the variances of the reputation values of target journals and the Spearman's rank correlation coefficient of journals' RCR values under different p values. The greater the variance values, the more significant the reputation difference. We can see from the data in Table 2 that the value of p has not changed the reputation ranking results of target journals, but will affect the significance of the reputation (the greater the p value, the smaller the variance of the reputation value of the journal). Table 3 illustrates the Spearman's rank correlation coefficient results of RCR values under the different p values, and when p is 0.3, 0.4, or 0.5, the correlation is higher. By combining the variance values, we chose $p = 0.3$ as the experimental data for further analysis.

After we obtained the RCR values of the journals, we conducted two rankings of the journals by using the traditional PageRank algorithm and the RACB model respectively. Yu *et al.* (2017) determined a Spearman's rank correlation coefficient of different damping factors. The result showed that $d = 0.85$ was considered the most reasonable value. Finally, we discussed the rationality of the influence of the RACB model on the journal rankings.

C. RESULTS AND ANALYSIS

In this section, we compare the results of various ranking methods, including the traditional PageRank algorithm,

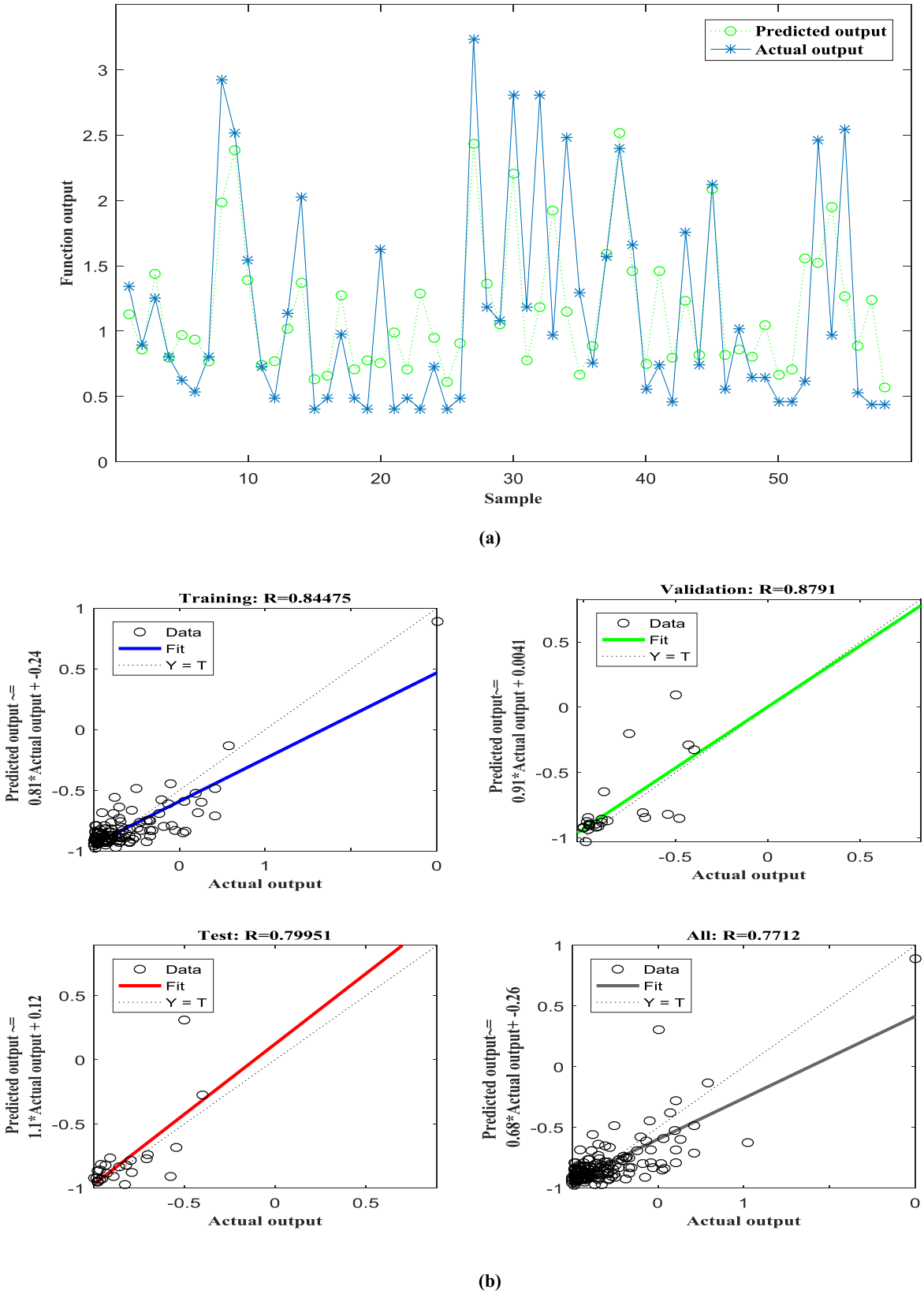


FIGURE 2. The fitting effect diagram of GA-BP algorithm.

TABLE 2. The RCR values under different p values.

Journals	p=0.1	p=0.2	p=0.3	p=0.4	p=0.5	p=0.6	p=0.7
Bioethics	0.7240	0.7371	0.7479	0.7570	0.7647	0.7713	0.7771
Developing World Bioethics	0.6558	0.6908	0.7192	0.7427	0.7625	0.7794	0.7941
Hastings Center Report	0.5713	0.6018	0.6263	0.6466	0.6636	0.6780	0.6905
Health Care Analysis	0.8961	0.9052	0.9129	0.9194	0.9250	0.9299	0.9342
Journal Of Bioethical Inquiry	0.6952	0.7235	0.7469	0.7666	0.7835	0.7981	0.8108
Journal Of Law Medicine and Ethics	0.5575	0.6249	0.6730	0.7093	0.7379	0.7611	0.7803
Journal Of Medical Ethics	0.7283	0.7486	0.7626	0.7731	0.7812	0.7879	0.7934
Journal Of Medicine And Philosophy	0.7258	0.7380	0.7481	0.7567	0.7641	0.7704	0.7760
Kennedy Institute Of Ethics Journal	0.6978	0.7140	0.7273	0.7383	0.7476	0.7556	0.7625
Science And Engineering Ethics	0.7199	0.7454	0.7667	0.7846	0.8000	0.8133	0.8249
Variance value	0.0089	0.0067	0.0055	0.0047	0.0042	0.0039	0.0037

TABLE 3. Spearman's rank correlation coefficient of journals RCR values under different p values.

The p values	0.1	0.2	0.3	0.4	0.5	0.6	0.7
0.1	1.0000						
0.2	0.9911	1.0000					
0.3	0.9688	0.9931	1.0000				
0.4	0.9379	0.9754	0.9945	1.0000			
0.5	0.9019	0.9505	0.9802	0.9956	1.0000		
0.6	0.8633	0.9213	0.9601	0.9841	0.9964	1.0000	
0.7	0.8239	0.8898	0.9365	0.9679	0.9872	0.9971	1.0000

the RACB model, the IF value of 2016, the IF value of 2017, the latest five-year IF value, the CS of 2016, and the CS of 2017, and analyze the rationality of the RACB model considering the reputation evaluation. We also compare the relationship between the RCR value, the self-citation rate, and the h-index of the target journals to identify the reasonableness and effectiveness of the RCR value.

In the fundamental bibliometric methods, compared with IF, Eigenfactor score considers not only the number of citations but also the quality of citations, however, Eigenfactor score is calculated based on the reference data of the last five years, while the PageRank algorithm and the RACB model is calculated based on the reference data of the last one year. Similar to Eigenfactor score, the SJR also considers the quality of citation, but the SJR and other methods (like Eigenfactor score, IF, *et al.*) mainly differ in that they use the different original data sets. SJR is calculated based on Scopus database, while the other methods are calculated based on Web of Science database. The h-index is mainly used to measure high-quality papers or scholars, but difficult to measure the general low-cited papers. Therefore, the above three fundamental bibliometric methods are not appropriate for the comparison with the proposed RACB model.

1) ANALYSIS OF THE JOURNAL RANKING RESULTS

Table 4 and Table 5 list the numeric size and ranking results of the target journals under various ranking methods. The ranking results of the traditional PageRank algorithm are close to those of the IF of 2016, but differ greatly from the IF rankings of 2017. The difference is caused by the change in the original dataset. The results of CS of 2016 are very close to those of CS of 2017, but they both differ greatly from those of traditional ranking methods or the RACB model. After considering the reputation evaluation of the target journals' inter-citation behavior, the ranking results of the RACB model were changed slightly compared to the traditional PageRank algorithm. Compared with the traditional PageRank algorithm, the ranking results of the RACB model are closer to those of IF of 2017.

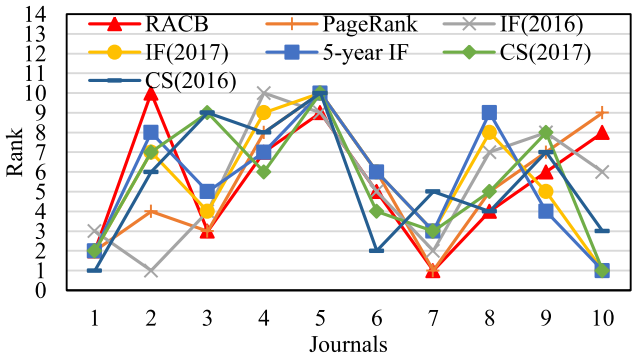


FIGURE 3. The journals ranking with seven methods.

Fig. 3 shows the differences in the ranking results of each method more intuitively. To better illustrate the relationship between these methods, this study conducts a Spearman's correlation analysis for the seven ranking methods. The results are shown in Table 6.

We can see from the Table 6, the correlation coefficient between the latter two is as high as 0.879. This is mainly because CS is calculated based on the reference data of the last three years. The data sets for calculating CS of 2016 and CS of 2017 are the same in two years, which leads to insignificant difference in their results. However, the RACB model proposed in this paper is calculated based on one year's journal citation data, and IF is based on two years' citation data. As for RACB model, the IF value is more comparable than CS value, therefore, in the subsequent comparative analysis,

TABLE 4. The calculation results of the target journals.

Journal number/ Value	PageRank	IF(2016)	IF(2017)	5-year IF	RACB	CS(2016)	CS(2017)
1	0.2604e-04	1.75	1.562	1.676	0.1436e-04	1.43	1.35
2	0.1829e-04	1.769	0.898	1.212	0.0248e-04	0.91	0.85
3	0.2198e-04	1.731	1.345	1.358	0.124e-04	0.72	0.66
4	0.0469e-04	0.875	0.82	1.222	0.0385e-04	0.77	0.91
5	0.0250e-04	1.204	0.817	0.868	0.0275e-04	0.64	0.56
6	0.1582e-04	1.613	1.223	1.315	0.0916e-04	1.33	1.15
7	0.3107e-04	1.764	1.529	1.54	0.1743e-04	0.96	1.17
8	0.1688e-04	1.293	0.871	0.98	0.0964e-04	0.98	1.05
9	0.1101e-04	1.129	1.333	1.455	0.0678e-04	0.82	0.75
10	0.0306e-04	1.454	2.229	1.813	0.0301e-04	1.12	1.57

TABLE 5. The ranking results of the target journals.

Journal number/ Rank	PageRank	IF(2016)	IF(2017)	5-year IF	RACB	CS(2016)	CS(2017)
1	2	3	2	2	2	1	2
2	4	1	7	8	10	6	7
3	3	4	4	5	3	9	9
4	8	10	9	7	7	8	6
5	10	9	10	10	9	10	10
6	6	5	6	6	5	2	4
7	1	2	3	3	1	5	3
8	2	7	8	9	4	4	5
9	4	8	5	4	6	7	8
10	3	6	1	1	8	3	1

TABLE 6. Spearman's rank correlation coefficient among different methods.

Methods	RACB	PageRank	IF(2016)	IF(2017)	5-year IF	CS(2016)	CS(2017)
RACB	1.000						
PageRank	0.746	1.000					
IF(2016)	0.333	0.806	1.000				
IF(2017)	0.479	0.442	0.539	1.000			
5-year IF	0.442	0.309	0.346	0.952	1.000		
CS(2016)	0.358	0.333	0.418	0.564	0.515	1.000	
CS(2017)	0.358	0.261	0.333	0.661	0.673	0.879	1.000

the IF value is used as the main comparative index of the proposed RACB model.

For the experimental results, this paper mainly includes seven ranking methods: the RACB model, traditional PageRank algorithm, the IF value of 2016, the IF value of 2017, latest five-year IF value, the CS of 2016 and the CS of 2017.

The traditional PageRank algorithm and the IF value of 2016, which is calculated based on the citation data of 2016 year, are the contrast methods of the proposed RACB model. In this study, we choose the IF value of 2017 and latest five-year IF value as the standard ranking methods, because of following two reasons: (1) although there are many journal ranking methods, the IF ranking of JCR reports is generally accepted in academic circles; (2) the IF value only considers the citation counts in the calculation and it is not appropriate to use the IF value ranking of the same year as the standard citation method, while the RACB model is an improved method which combines the journal's inter-citation reputation and the citation counts.

Traditional method only based on citation counts has difficulty in identifying the future trend of journals' influence, but the proposed RACB model can predict the future changes of journals' influence according to the journals' inter-citation reputation. As we can see from Table 5, the journal entitled Developing World Bioethics ranks fourth in the traditional PageRank algorithm, ranks first in the IF value of 2016, but drops to seventh in IF value of 2017. Its ranking result is only 10th in the RACB model because its lower inter-citation reputation is identified after its citation behavior is analyzed by the RACB model. This phenomenon shows that the RACB model considering the inter-citation reputation can effectively predict the future changes of the journals' influence, therefore its ranking results are more reasonable.

From the Table 5, we can see that the ranking results of the IF value of 2016 based on the 2016 citation data have a significant difference from the ranking results of the IF value of 2017. This is mainly owing to changes in the citation data of the target journals during the year. The 2017 journal citation data are affected by the reputation of the journals in 2016 or even earlier. For example, in 2016, if the journal A colluded with journal B which has low correlation with journal A in citation, then the citation behavior will affect the journal reputation, and will affect researchers' citation choices in the next year. There are many other factors that lead to the change of citation data in 2017, such as the journals' self-citation, the journals' inter-citation reputation and so on. When we are ranking journals based on the traditional method and taking into account one or more of these factors, the ranking results will be closer to the IF ranking of 2017.

TABLE 7. The ranking results of the 10 journals.

Rank of RCR	Journal name	RCR values	The h-index	Total number of published	Self-citation rate	Total number/h-index
5	*Bioethics*	0.7479	27	757	6.2264%	28.0370
8	*Developing World Bioethics*	0.7192	16	262	4.2963%	16.3750
10	*Hastings Center Report*	0.6263	26	1198	4.6070%	46.0769
1	*Health Care Analysis*	0.9129	18	275	4.8157%	15.2778
6	*Journal of Bioethical Inquiry*	0.7469	14	627	16.6774%	44.7857
9	*Journal of Law Medicine and Ethics*	0.6730	29	1046	8.7502%	36.0690
3	*Journal of Medical Ethics*	0.7626	38	2176	9.6077%	57.2632
4	*Journal of Medicine And Philosophy*	0.7481	23	418	19.6476%	18.1739
7	*Kennedy Institute of Ethics Journal/*	0.7273	17	208	5.6466%	12.2353
2	*Science And Engineering Ethics*	0.7667	26	715	14.9801%	27.5000

Notes: **(1)** All the indicators are based on citation data of 2007-2016 years on the ISI Web of Science; **(2)** The smaller the ratio of total number to h-index is, the more high quality articles the journal has.

In this study, the ranking results of the RACB model which consider the journals' inter-citation reputation on the basis of the traditional PageRank algorithm, are closer to the IF value ranking of 2017. This means that the journals' reputation evaluation of inter-citation behavior is one of the factors affecting future citation data.

We will study more factors that influence the future citation data, and constantly improve our ranking methods. When we consider enough factors, the new ranking method will accurately predict the IF trend of target journals in the future.

2) ANALYSIS OF THE JOURNAL REPUTATION RESULTS

This study mainly explored the influence of the inter-citation behavior reputation on journal ranking. The results also show that the ranking of highly reputable journals will rise. To identify the different characteristics of high- and low-reputation journals, this study lists some other relevant data of target journals in Table 7. These indicators include the total number of publications, h-index, the self-citation rate of journals, and the ratio of the total number of publications to the h-index. The total number of publication of journals is generally related to factors such as the journal review cycle and the strict degree of review. The h-index indicates the number of high-quality articles in a journal. The comparison of the h-index will be more scientific on the basis of the same number of publications. Thus, we added the ratio of the published number to the h-index for comparison purpose in Table 7.

The reputation evaluation of journal citation behavior includes self-citation behavior and inter-citation behavior. The proposed RACB model studies the reputation problem of the journals' inter-citation behavior. The ratio of the total number to the h-index can be used to approximately measure the quality of the journals. Although the calculation of the h-index is based on the citation data between journals, the citation behavior of high-quality articles is more credible than that of general articles.

We can see in Table 7 that the two target journals, *Hastings Center Report* and *Health Care Analysis*, have relatively low self-citation rates (4.6070% and 4.8157%). The experimental results of the proposed RACB model show a clear difference between their RCR values (0.9129 and 0.6263). In general, the citation behavior of *Health Care Analysis* is more reputable than that of the *Hastings Center Report*. The ratio of the total published number to the h-index also confirms this conclusion. The ratio of the*Hastin Center Report* is 46.0769, the value is far greater than the ratio of reputable journal (the *Health Care Analysis* is 15.2778). This phenomenon indicates that the reputation of the journal's citation behavior affects the author's choice of where to submit his or her manuscript, especially for an author who usually writes high-quality articles.

V. CONCLUSIONS

Scholars study the authoritative ranking of journals or papers, generally based on the cited times. However, a single index not only makes the measurement inaccurate but also produces false information easily. Thomson Reuters proposed the concept of self-citation distortion, and investigated and studied journals with high self-citation and data anomalies, giving warnings to journals with serious self-citation or coupling-citation and even excluding them from the SCI (or SSCI).

This study proposed a new method for journal rankings named the RACB model. The method considers journals' reputation evaluations of inter-citation behavior based on the PageRank algorithm. Most journal ranking methods are based on citation data, but if there is a phenomenon similar to ''collusion citation'' when an author cites the articles, the results of these ranking methods will have a larger error than the real ranking of journals. This study analyzed and compared the ranking results of the PageRank algorithm and the RACB model, which were based on citation data from 2016. Moreover, this study found that the ranking result of the RACB model considering inter-citation reputation evaluations has a stronger correlation to the IF ranking of 2017. This also means the journals' reputation of inter-citation will affect the author's citation choice, and the ranking method that

considers the reputation evaluation can get more reasonable ranking results.

The proposed RACB model can be used to evaluate a journal's reputation and also evaluate the author's reputation, even in an article. The expectation of this study is that the reputation evaluation of inter-citation behavior can be used as commonly as citations in bibliometric in the future. The inter-citation data will be more authentic and reliable if this can be achieved.

However, the study has some limitations. It is not perfect for calculating the JR value between journals with the RACB model, which is difficult to use in large-scale journal rankings. The accuracy of the fitting formula of the CDI has a significant correlation with the journal category, i.e., the more frequent the interdisciplinary studies are, the lower the accuracy of the CDI in this field. Moreover, the RACB model does not consider the factor of self-citation. All work in this study is based on inter-citation of the journals only. Our future work will combine the inter-citation and the self-citation of journals in the ranking model. Considering that the proposed model is more complicated and requires more computational cost than the traditional journal ranking methods, we will need to improve it to make it both effective and simple in the future. In the future, the possibility of the integration among the linear mixed model, eigenvector scoring algorithm and the proposed RACB model can be explored so that the journals' reputation can be formulated more comprehensively. In addition, the partial correlation analysis using possible confounders (prestige, open-access option, etc.) can also be considered for the possible integration with the present journals' inter-citation behavior so that the journals' reputation can be modeled more flexibly.

SUPPORTING INFORMATION

The data used in this paper have been uploaded to Figshare (https://doi.org/10.6084/m9.figshare.7653818.v1). The raw data are divided into two parts: one is the ten year citation data of the top journal (D1.rar), and the other is the citation data of the ten selected journals (D2.rar). These raw data are downloaded from Web of Science. The remaining data include the matrix after the word frequency extraction of top journal (D3.rar), the preprocessing results of top journal data (D4.xlsx), the matrix after the word frequency extraction of ten journals (D5. rar), the JR and IF of ten journals (D6.xlsx), the ACR values of ten journals (D7.xlsx), the CDI and ACR of ten journals (D8.xlsx), RCR values under different p values (D9.xlsx), the inter-citation data of ten journals (D10.xlsx), and the ranking results of various ranking methods (D11.xlsx).

REFERENCES

[1] E. Garfield, "Citation analysis as a tool in journal evaluation," *Science*, vol. 178, no. 4060, pp. 471–479, 1972.

[2] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 46, pp. 16569–16572, 2005.

[3] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, nos. 1–7, pp. 107–117, Apr. 1998.

[4] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *J. ACM*, vol. 46, no. 5, pp. 604–632, 1999.

[5] C. Bergstrom, "Eigenfactor: Measuring the value and prestige of scholarly journals," *College Res. Libraries News*, vol. 68, no. 5, pp. 314–316, 2007.

[6] J. P. A. Ioannidis, "A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation," *J. Psychosomatic Res.*, vol. 78, no. 1, pp. 7–11, 2015.

[7] S.-R. Yan, X.-L. Zheng, Y. Wang, W. W. Song, and W.-Y. Zhang, "A graph-based comprehensive reputation model: Exploiting the social context of opinions to enhance trust in social commerce," *Inf. Sci.*, vol. 318, pp. 51–72, Oct. 2015.

[8] M. Wang, G. Wang, Y. Zhang, and Z. Li, "A high-reliability multi-faceted reputation evaluation mechanism for online services," *IEEE Trans. Services Comput.*, to be published.

[9] P. L. K. Gross and E. M. Gross, "College libraries and chemical education," *Science*, vol. 66, no. 1713, pp. 385–389, 1927.

[10] A. M. Cartter and R. A. Sawyer, "An assessment of quality in graduate education," *Phys. Today*, vol. 9, no. 8, p. 75, 1966.

[11] B. S. Kademani and V. L. Kalyane, "Bibliometric indicators for publication productivity analysis of an individual scientist," *Int. J. Scientometrics Inform.*, vol. 2, no. 4, pp. 49–58, 1996.

[12] E. Garfield, "The meaning of the impact factor," *Int. J. Clin. Health Psychol.*, vol. 3, no. 2, pp. 363–369, 2003.

[13] P. O. Seglen, "Why the impact factor of journals should not be used for evaluating research," *Brit. Med. J.*, vol. 314, no. 7079, pp. 498–502, 1997.

[14] S. Maslov and S. Redner, "Promise and pitfalls of extending Google's PageRank algorithm to citation networks," *J. Neurosci.*, vol. 28, no. 44, pp. 11103–11105, 2008.

[15] Y. Ding, E. Yan, A. Frazho, and J. Caverlee, "PageRank for ranking authors in co-citation networks," *J. Assoc. Inf. Sci. Technol.*, vol. 60, no. 11, pp. 2229–2243, 2009.

[16] E. Yan and Y. Ding, "Discovering author impact: A PageRank perspective," *Inf. Process. Manage.*, vol. 47, no. 1, pp. 125–134, 2011.

[17] C. Su *et al.*, "PrestigeRank: A new evaluation method for papers and journals," *J. Inform.*, vol. 5, no. 1, pp. 1–13, 2011.

[18] W. Zheng, S. Mo, P. Duan, and X. Jin, "An improved PageRank algorithm based on fuzzy C-means clustering and information entropy," in *Proc. 3rd IEEE Int. Conf. Control Sci. Syst. Eng. (ICCSSE)*, Beijing, China, Aug. 2017, pp. 615–618.

[19] F. Fernandez-Llimos, "Differences and similarities between Journal Impact Factor and CiteScore," *Pharmacy Pract.*, vol. 16, no. 2, p. 1282, 2018.

[20] Y. Liping, C. Yuqing, P. Yuntao, and W. Yishan, "Research on the evaluation of academic journals based on structural equation modeling," *J. Inform.*, vol. 3, no. 4, pp. 304–311, 2009.

[21] H. F. Moed, "Comprehensive indicator comparisons intelligible to non-experts: The case of two SNIP versions," *Scientometrics*, vol. 106, no. 1, pp. 51–65, 2016.

[22] B. Cheang, S. K. W. Chu, C. Li, and A. Lim, "A multidimensional approach to evaluating management journals: Refining PageRank via the differentiation of citation types and identifying the roles that management journals play," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 12, pp. 2581–2591, 2014.

[23] L. Vana, R. Hochreiter, and K. Hornik, "Computing a journal meta-ranking using paired comparisons and adaptive lasso estimators," *Scientometrics*, vol. 106, no. 1, pp. 229–251, 2016.

[24] C. S. Wagner *et al.*, "Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature," *J. Inform.*, vol. 5, pp. 14–26, Jan. 2011.

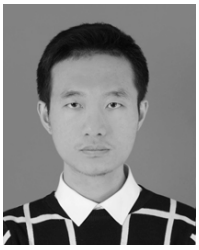
[25] B. K. Maity and S. R. Hatua, "Designing a model to evaluate scholarly publications with special reference to social sciences in India," *Scientometrics*, vol. 109, no. 3, pp. 2031–2048, 2016.

[26] D. Yu, W. Wang, S. Zhang, W. Zhang, and R. Liu, "A multiple-link, mutually reinforced journal-ranking model to measure the prestige of journals," *Scientometrics*, vol. 111, no. 1, pp. 521–542, 2017.

[27] M. Song, S. Kim, and K. Lee, "Ensemble analysis of topical journal ranking in bioinformatics," *J. Assoc. Inf. Sci. Technol.*, vol. 68, no. 6, pp. 1564–1583, 2017.

[28] X. Jiang, X. Sun, Z. Yang, H. Zhuge, and J. Yao, "Exploiting heterogeneous scientific literature networks to combat ranking bias: Evidence from the computational linguistics area," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 7, pp. 1679–1702, 2016.
[29] E. Garfield, "Is citation analysis a legitimate evaluation tool?" *Scientometrics*, vol. 1, no. 4, pp. 359–375, 1979.
[30] J. Priem, D. Taraborelli, P. Groth, and C. Neylon. (2010). *Altmetrics: A Manifesto*. [Online]. Available: http://altmetrics.org/manifesto/
[31] R. Costas, Z. Zahedi, and P. Wouters, "Do 'altmetrics'' correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective," *J. Assoc. Inf. Sci. Technol.*, vol. 66, no. 10, pp. 2003–2019, 2015.
[32] R. K. Merton, *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL, USA: Univ. Chicago Press, 1973.
[33] T. R. Tyler, "Psychological perspectives on legitimacy and legitimation," *Annu. Rev. Psychol.*, vol. 57, pp. 375–400, Jan. 2006.
[34] M. Helmer, M. Schottdorf, A. Neef, and D. Battaglia, "Gender bias in scholarly peer review," *Elife*, vol. 6, Mar. 2017, Art. no. e21718.
[35] L. Waltman, "A review of the literature on citation impact indicators," *J. Inform.*, vol. 10, no. 2, pp. 365–391, 2016.
[36] Y.-J. Oh, H.-J. Oh, C.-H. Kim, and Y. Kim, "A study on the citation behavior by academic background of researchers," *J. Korean Soc. Inf. Manage.*, vol. 33, no. 1, pp. 247–268, 2016.
[37] S. Sendhilkumar, E. Elakkiya, and G. S. Mahalakshmi, "Citation semantic based approaches to identify article quality," in *Proc. Int. Conf. (ICCSEA)*, Delhi, India, May 2013, pp. 411–420.
[38] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Inf. Process. Manage.*, vol. 24, no. 5, pp. 513–523, 1988.
[39] W. Zhang, T. Yoshida, and X. Tang, "A comparative study of TF*IDF, LSI and multi-words for text classification," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2758–2765, 2011.
[40] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval," *J. Documentation*, vol. 28, no. 1, pp. 11–21, 1972.
[41] Y. Poria, Z. Schwartz, and M. Uysal, "IF you can keep your head: The unintended consequences of the Impact Factor on tourism research," *Tourism Manage.*, vol. 51, pp. 300–302, Dec. 2015.
[42] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.
[43] W. Lv, M. Wang, and X. Zhu, "Model for prediction of surface subsidence coefficient in backfilled coal mining areas based on genetic algorithm and BP neural network," *J. Comput. Methods Sci. Eng.*, vol. 16, no. 4, pp. 745–753, 2016.
[44] F. Biljecki, "A scientometric analysis of selected GIScience journals," *Int. J. Geographical Inf. Sci.*, vol. 30, no. 7, pp. 1302–1335, 2016.
[45] A. Muzur, "Journal issue review: The American journal of bioethics 14, No. 8 (2014)," *JAHR*, vol. 6, no. 1, pp. 133–134, 2015.
[46] K. W. Fan, "Examining the validity of impact factors for journals in the history and philosophy of science category of science citation index," *Sci. Technol. Libraries*, vol. 34, no. 4, pp. 329–338, 2015.

LEIBAO ZHANG received the Ph.D. degree in economics from the Shanghai University of Finance and Economics, China, in 2005. He is currently a full-time Professor with the School of Public Finance and Taxation, Zhejiang University of Finance and Economics, China. He has published about more than 40 papers in the national core economic journals, covering fiscal and taxation theories and policies, and performance management of public expenditure. He has completed about seven books and teaching materials (including cooperation).



QI QIAN is currently pursuing the M.S. degree with the Zhejiang University of Finance and Economics, China. His current research interest includes bibliometrics.

SHUAI ZHANG received the Ph.D. degree in mechanical engineering from Zhejiang University, China, in 2005. He is currently a full-time Professor with the School of Information, Zhejiang University of Finance and Economics, China. He has published more than 30 papers in international journals in the recent ten years, covering bibliometrics, supply chain management, business intelligence, data mining, E-government, and manufacturing informatization.

WENYU ZHANG received the B.S. degree from Zhejiang University, China, in 1989, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2002. He is currently a full-time Professor with the School of Information, Zhejiang University of Finance and Economics, China. He has published more than 40 papers in international journals and more than 20 papers in international conference proceedings in the recent ten years, covering supply chain management, digital library, bibliometrics, concurrent engineering, distributed manufacturing, business intelligence, business analytics, data mining, multi-agent technology, and semantic Web.

KUN ZHU is currently pursuing the bachelor's degree with the Zhejiang University of Finance and Economics, China. His current research interest includes bibliometrics.