Scientometrics (2019) 120:155–166
https://doi.org/10.1007/s11192-019-03117-9

Ranking scientific publications considering the aging characteristics of citations

Yanan Wang[1] · An Zeng[1] · Ying Fan[1] · Zengru Di[1]

Received: 20 September 2018 / Published online: 24 May 2019

Abstract
Ranking the significance of scientific publications has been a challenging topic for a long time. So far, many ranking methods have been proposed, one of which is the well-known PageRank algorithm. In this paper, we introduce aging characteristics to the PageRank algorithm via considering only the first 10 year citations when aggregating resource from different nodes. The validation of our new method is performed on the data of American Physical Society journals. The results indicate that taking into account aging characteristics improves the performance of the PageRank algorithm in terms of ranking accuracy for both papers and authors. Though our method is only applied to citation networks in this paper, it can be naturally used in many other real systems and similar improvements are expected.

Keywords PageRank · Citation network · Time bias · Milestone papers · Scholars

Introduction

For a long time, determining how to measure the scientific influence of scientific publications has been one of the focus of scietometrics (Sidiropoulos and Manolopoulos 2006; Zeng et al. 2017). So far, many measures have been proposed, but these methods have some drawbacks (Maslov and Redner 2008; Frey and Rost 2010; Sorzano et al. 2014). When we measure the impact of scientific publications or scientists, the number of citations is a simple but widely used metric (Garfield 1955; Amsterdamska and Leydesdorff 1989). This method is intuitive and easy to use, but it only considers the number of references, and ignores the quality of the references (Maslov and Redner 2008). So the effectiveness of the measurement needs to be improved. To evaluate the scientific impact of a paper, one must consider not only the number of citations that it received, but also the ways by which the paper is being cited. Taking into account the quality of citing papers and other factors will influence the importance of cited papers (Radicchi et al. 2009; Franceschet 2010).

✉ An Zeng
anzeng@bnu.edu.cn

✉ Ying Fan
yfan@bnu.edu.cn

1 School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China

The PageRank algorithm (Brin and Page 1998) embodies this idea. It takes into account the importance of the citing papers and tries to assign high scores to these papers that are cited by some important papers. This algorithm is an iterative process on networks. In each step, the score of each paper is updated by aggregating the resource passed from all the papers that cited it. The final stable scores are used as the indicator of the significance of the papers.

As the PageRank algorithm is widely used in ranking scientific publications (Chen et al. 2007; Ma et al. 2008), a variety of modifications based on it have been proposed (Walker et al. 2007; Su et al. 2011; Fiala 2012; Nykl et al. 2014; Yao et al. 2014; Zhou et al. 2016; Mariani et al. 2016). For instance, Walker et al. (2007) proposed the CiteRank algorithm. This method modifies the PageRank algorithm by initially distributing random surfers exponentially with age in order to account for strong aging characteristics of citation networks. Another variant, called rescaled PageRank (Mariani et al. 2016), deals with the PageRank scores of papers published in a similar time in order to weaken the time bias. The lack of time bias in this metric makes it possible to compare papers of different age on the same scale. The PageRank algorithm has also been applied to rank the importance of different journals (Bollen et al. 2006; Gonzalez-Pereira et al. 2010), scientists (Radicchi et al. 2009; Ding et al. 2009; Yan and Ding 2011; Ding 2011; Fiala and Tutoky 2017), institutions (Yan 2014), countries (Fiala 2012), etc.

It is generally acknowledged that because the number of citations of a paper is accumulated over time, old papers have a longer time span to obtain citations. Consequently, old papers have more advantages over their young peers. The PageRank algorithm exhibits this time bias (Rendner 2004; Chen et al. 2007; Mariani et al. 2015). In this method, old papers tend to get higher scores. In order to solve this problem, several methods have been proposed, such as the aforementioned CiteRank and rescaled PageRank methods. One may wonder if there are any other solutions to this issue. In the past few years, some researches have considered the number of citations within different numbers of years after publication of a paper to measure the importance of a paper when studying the creativity of scientists in different career periods (Wang et al. 2013; Sinatra et al. 2016). Despite that this metric cannot result in a perfectly fair comparison of papers from different years, it weakens the strong bias of citation count to old papers. For instance, Sinatra et al. (2016) used the concept of "c_{10}"(the number of citations within 10 years after the publication of a paper) to measure the importance of a paper. This concept inspires us to apply the main idea behind "c_{10}"to improve the PageRank algorithm by only considering the citations within 10 years after the publication of a paper. In this way, the accumulated advantage of old papers will be eliminated, which makes the latest published papers have the opportunity to come to the fore. The method is denoted as c_{10}-based PageRank (CPRank). In this method, the edges that do not satisfy the time criteria are not deleted, so that the resource that is passed from the downstream nodes (the citing papers) to the upstream nodes (the cited papers) is unchanged. However, the upstream nodes only receive the resource from these downstream nodes that meet the time criteria instead of all downstream nodes. This differs the CPRank algorithm from the classic PageRank algorithm.

We validate our method in the citation network constructed from the data of American Physical Society (APS) journals. By comparing the rankings of influential papers (e.g. Nobel prize winning papers and the Milestone Letters) under different algorithms, we found that the CPRank method can improve the ranking of relatively new papers and weaken the accumulated advantage of old papers. We also ranked authors with the CPRank method. We summed the scores of all papers of an author to represent the score of the authors, and rank them accordingly. The results show that the Nobel prize winning authors gain better rankings in the

CPRank method than in other methods. Moreover, we compared the rankings of the Nobel prize winning papers among all papers of Nobel laureates in different methods. We observed that the CPRank method could achieve a good performance. Specifically, most Nobel laureates' prize winning papers have better rankings in CPRank than in other methods.

Method

Data collection

In order to test the performance of our algorithm, we constructed the citation network from the data of American Physical Society (APS) journals. Since our analysis needs papers' publication time, we delete those papers lacking time information and corresponding papers which have a citation relationship with them. In the end, for each paper in our database, the information about its DOI, authors, publication time and the DOI of its citing papers is all available. In total, the data contains 481,768 papers published in the period from 1893 to 2010 with 5,002,807 citations. With the information mentioned above, we construct a citation network in which nodes represent papers and edges represent citation relationships among papers. The citation network is directed and acyclic. The outdegree of a node represents the number of papers cited by it and the indegree of a node is the number of papers citing it. Among all papers in our database, there are 87 Milestone Letters published from 1958 to 2002 and we consider 23 Nobel Prize winning papers published from 1961 to 2000 serving as a benchmark set to investigate the performance of different algorithms.

The CPRank method

To begin with, we briefly describe the classic PageRank algorithm. The citation network could be described via adjacency matrix A with the element $A_{ij} = 1$ if paper i cites paper j and $A_{ij} = 0$ if there is no citation relationship between paper i and paper j. The PageRank algorithm can be expressed mathematically as follows:

$$s_j(t) = c + (1-c)\sum_{i=1}^{N}\left(\frac{A_{ij}}{k_i^{\mathrm{out}}}(1-\delta_{k_i^{\mathrm{out}},0}) + \frac{1}{N}\delta_{k_i^{\mathrm{out}},0}\right)s_i(t-1) \tag{1}$$

where $\delta_{a,b} = 1$ when $a = b$, and $\delta_{a,b} = 0$ otherwise. In Eq. (1), N is the total number of nodes in the network, k_i^{out} is the number of outgoing links of node i, that is, the outdegree, $s_j(t)$ is the PageRank score of node j through t steps iterative process and the same for s_i, c is the return probability. The PageRank algorithm describes a random walk process on a directed network, where the score s_i captures the frequency of a particular node to be visited by a random walker. The parameter c ($0 \le c \le 1$) represents the probability for a walker to jump to a random node from the current node, and the $(1-c)$ is the probability for the random walker to continue walking through the directed links. The typical value of c is approximately 0.15 (Brin and Page 1998). Accordingly, we fix $c = 0.15$ in our study. Usually the initial configuration is to set $s(0) = 1$ for all nodes. The final score of each node is defined as the steady value when $s_i(t)$ keeps convergent. Based on the result of each node's final score, we can obtain the final ranking of nodes in PageRank by sorting the scores in descending order.

The essential difference between the CPRank method and the PageRank algorithm is that when counting an article's total number of citations the CPRank method only considers the citations within 10 years after their publication. We adopt this idea in order to weaken the cumulative advantage in terms of time for those old papers. In CPRank, a new adjacency matrix B_{ij} is introduced. The element $B_{ij} = 1$ if paper i cites paper j and paper i was published no more than 10 years after the publication of paper j, and $B_{ij} = 0$ otherwise. The CPRank method can be specifically expressed as the following formula:

$$s_j(t) = c + (1-c)\sum_{i=1}^{N}\left(\frac{B_{ij}}{k_{i,A}^{\mathrm{out}}}(1-\delta_{k_i^{\mathrm{out}},0}) + \frac{1}{N}\delta_{k_i^{\mathrm{out}},0}\right)s_i(t-1) \tag{2}$$

where $k_{i,A}^{\mathrm{out}}$ is the outdegree of node i in the network represented by the adjacency matrix A, rather than in the network represented by the adjacency matrix B. In this method, the resource distribution process from the downstream nodes to the upstream nodes is different with the PageRank algorithm. The downstream nodes do not pass scores to the upstream nodes that do not meet the time criteria. The illustration of the algorithm is shown in Fig. 1.

Results

Basic properties of the CPRank method

In this section, we will apply the APS citation network to test the performance of CPRank. To begin our analysis, we discuss the basic properties of the CPRank method as well as the PageRank algorithm so as to make a comparison between them. One of the basic properties is the convergence of the method. In order to find the result, we define a quantity *error* at step t as $\sum_i |s_i(t) - s_i(t-1)|$ where $s_i(t)$ is the score of node i at iteration step t. The results of *error* in different iteration steps are shown in Fig. 2a. For the convenience of presentation, PageRank is denoted as PR, and CPRank is denoted as PR_{c10} in the following figures. PageRank's convergence is guaranteed by the linear-algebra properties of the algorithm. Hence, provided that the number of iterations is large enough, the algorithm will converge for any given tolerance value. We find that CPRank can also converge to a small tolerance value after a few iteration steps. Its converging speed is similar to PageRank. Given this result, we set the iteration step $t = 100$ in our simulation for the reason that the scores of nodes have attained the stable status. Another basic property we consider

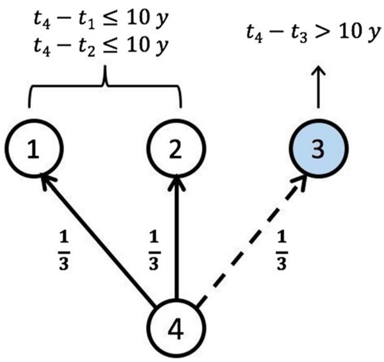

Fig. 1 (Color online) The illustration of the CPRank algorithm. It shows the distribution process of the score to the upstream nodes. The dotted line between node 3 and node 4 means that there is no score distribution between them, as node 3 does not satisfy the 10-year time limit

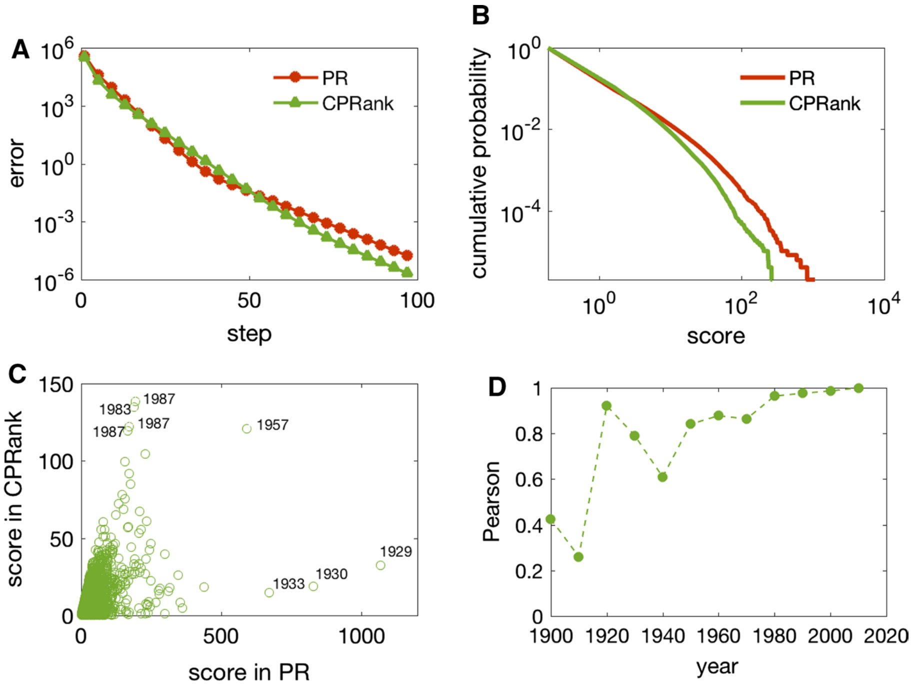

Fig. 2 (Color online) **a** The dependence of the total quantity error on iteration step t in CPRank. **b** The cumulative distribution of the final PageRank and CPRank scores. **c** The scatter plot of the scores in CPRank and PageRank. **d** The Pearson correlation coefficient of the CPRank and PageRank scores. Note that the year represents a time span of the past 10 years rather than a specific year

of the method is the distribution of the final scores. In Fig. 2b, we probe the distributions of the final CPRank and PageRank scores. We can find that the distributions of CPRank and PageRank scores both follow fat-tailed forms. The cumulative distributions seem to be substantially different in their tails, with the CPRank score being less broadly distributed than PageRank. It indicates that the influence of some papers is suppressed by CPRank method. We further depict the scatter plot of papers' PageRank and CPRank scores in Fig. 2c. We observe that some papers' scores are quiet high in PageRank but relatively low in CPRank. Meanwhile, several papers that have remarkable scores in CPRank gain comparatively low scores in PageRank. To better understand this result, we have identified the publication year of these articles. We discover that papers with outstanding scores in PageRank but poor scores in CPRank are old papers. On the contrary, those with large scores in CPRank but small scores in PageRank are new papers. This phenomenon indicates to some extend that the CPRank method effectively removes the cumulated advantage of old papers, which can be confirmed in the following analysis. Besides, we study the similarity between the CPRank and PageRank methods. We investigate the Pearson correlation coefficient between the scores of these two different methods in Fig. 2d. We take 10 years as a time span and calculate the correlation coefficient between the scores of papers in different methods from different periods. The points in Fig. 2d represent the coefficient of articles published in the past 10 years. For example, the corresponding point in 1910 represents the coefficient of papers published from 1901 to 1910. Because the articles in our dataset were first published in 1893, so the point in 1900 represents the coefficient of papers published from 1893 to 1900. One can see in Fig. 2d that the Pearson correlation coefficient of the CPRank and PageRank methods for papers published before 1910 stays at a low value (0.425 before 1900 and 0.261 before 1910 approximately). Then we can see that the coefficient witnesses a significantly upward tendency in 1920 and remains at relatively high values for the rest of the time.

The effectiveness of the CPRank method

It is generally accepted that a good ranking algorithm should be effective in identifying the truly influential nodes. Owing to the detailed information of the raw data, we can have access to the title of each paper, which enables us to determine which papers are recognized by prizes. Here we focus on a list of physics papers of remarkable significance called Milestone Letters which have been carefully selected by the editors of the APS for "having made long-lived contributions to physics, or by announcing significant discoveries, or by initiating new areas of research". Mariani et al. (2016) have taken use of these papers in their research to validate different algorithms. Similarly, we choose these papers as a benchmark set to test the performance of different algorithms in identifying the high quality papers. If an algorithm is able to give these articles higher scores, we consider the algorithm is better. Here, we picked 87 Milestone Letters from the year 1958–2002 as our benchmark articles. We respectively compute the scores of the Milestone Letters that are assigned by different algorithms, and compare the mean rank of these papers. Here we consider five metrics: PageRank (*PR*), CPRank, CiteRank, rescaled PageRank (PR_{rs}) and citation rank (c_{10}). It is noteworthy that the mean rank is easy to calculate but the result might be dominated by some lowly ranked articles. In order to avoid this problem, we first take the logarithm of the ranks and then calculate the mean value. The result is shown in Fig. 3. One can see in Fig. 3 that the CPRank algorithm can compete and even outperform other methods in identifying milestone papers, as these milestone papers have a smaller average ranking in the CPRank algorithm.

In order to better understand the reason behind the above results, we plot scatter diagrams of CPRank and other methods' rankings containing all articles in Fig. 4 and mark the dots representing Milestone Letters with red. Points locating above the diagonal suggests that PageRank outperforms another method. On the contrary, points locating below the diagonal reveals that another method outperforms PageRank. As shown in Fig. 4, one can see a clear difference between the CPRank and the other existing ranking methods that aim to identify recent influential papers. In Fig. 4a, it can be seen that most of the points

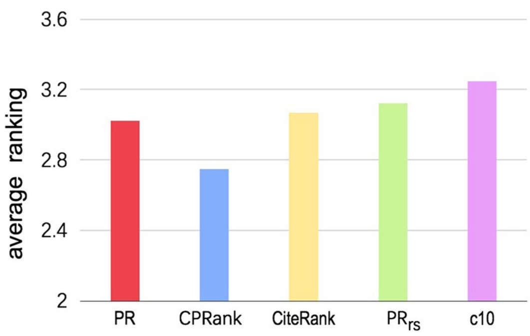

Fig. 3 (Color online) The average ranking of Milestone Letters in different methods. Note that the average ranking here is the average value of the logarithms of original rankings

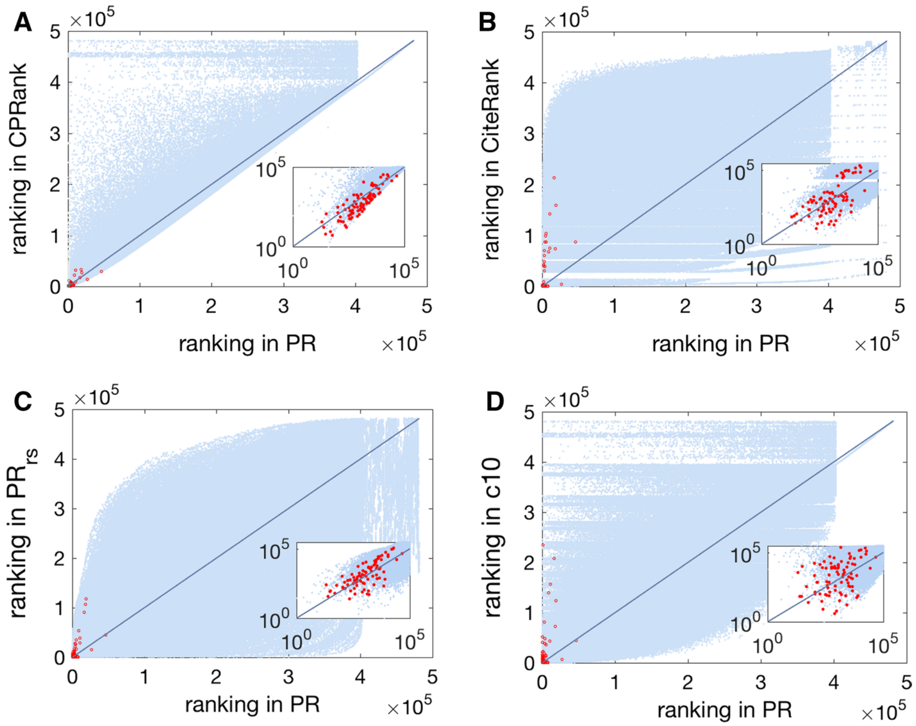

Fig. 4 (Color online) The scatter plot rankings of all papers in the PageRank method with the CPRank, CiteRank, rescaled-PageRank, c10 metrics, respectively. The red points represent milestone papers

locate above the diagonal line and there are only a small number of points below the diagonal. Meanwhile, the points below the diagonal are close to the line and there are no points locating in the lower right corner. We infer that this is because that the CPRank method makes the relatively new articles rank higher than in PageRank by suppressing the rank of some old articles which rank better in PageRank instead of enhancing the rank of articles which rank bad in PageRank. In the inset, the advantage of CPRank towards other existing methods can be observed as most of the red dots (milestone papers) locate under the diagonal line, indicating improvement of their rank by the CPRank method. The specific number of these milestone papers is 64 with the total number of 87. Interestingly, we discover that 23 milestone papers which rank worse in CPRank than in PageRank are all published before 1970.

Besides, we study the ranking of milestone articles in all articles published in the same decade. The result is shown in Fig. 5a. As depicted in the figure, the PageRank algorithm performs better for the relatively older papers published in 1961–1970. But in the next three decades, the CPRank algorithm performs better. The results indicate again that CPRank can weaken the time bias for old articles in the PageRank algorithm, and increase the opportunity of obtaining a good ranking. We then compare the ranking of milestone papers published in different periods with CPRank and other metrics as compared to all papers in the dataset in Fig. 5b. The result is somewhat different with Fig. 5a. Specifically, one can see that in Fig. 5b the CPRank method underperforms not only for the oldest

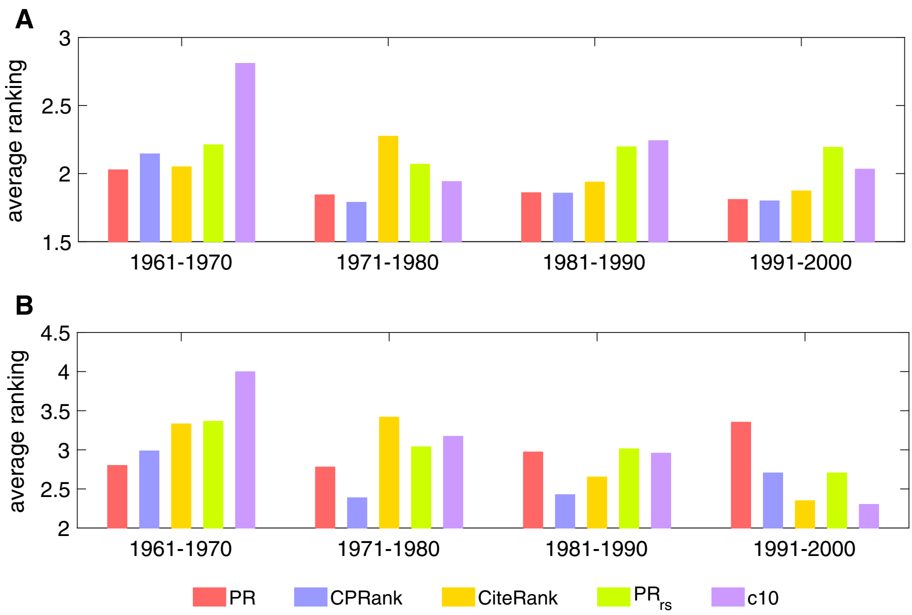

Fig. 5 (Color online) **a** The average ranking of Milestone Letters among all papers in the same period. **b** The average ranking of of Milestone Letters published in different periods as compared to all papers in the dataset. Note that the average ranking is the average value of the logarithms of original rankings

milestone papers, but also for the most recent ones. In fact, PageRank is strongly biased towards old papers, so when comparing the most recent Milestone Letters with all papers in the dataset, it has the worst performance as in Fig. 5b. And because CPRank can weaken the bias to old papers, it outperforms PageRank in Fig. 5b. However, Fig. 5a only compares papers in the same periods. When calculating the ranking of the most recent milestone papers, the old articles with advantages are not taken into account. In this way, the performance of PageRank is not weakened by its bias to old paper (as shown in Fig. 5a). However, the CiteRank and c_{10} metrics are biased towards new articles, so they have good performance on ranking the Milestone Letters comparing with all papers in the dataset (in Fig. 5b), but not for comparing papers in the same periods (in Fig. 5a).

In order to further verify whether the CPRank algorithm can modify the time bias in PageRank and make papers of different age comparable on the same scale, we used the statistical test in Radicchi et al.'s paper (2008) to assess the bias by the paper age of the rankings by CPRank with other metrics. We compute the percentage of papers in different decades that appear in the top 0.5%, 1% and 2% of the general ranking in different methods. The results are shown in Fig. 6. One can see that PageRank (PR) has strong bias towards old papers while the rescaled PageRank (PR_{rs}) has smallest bias by definition. As for CPRank, one can see that the peak locates closer to the middle. It implies that the ranking is biased towards more recent papers. But in general, it is still biased. That is to say, the CPRank algorithm can only weaken the bias of old articles but it still cannot make a perfectly fair comparison of articles from different periods. Similar to CPRank, a strong bias is also present for c_{10}. These biases are limitations of the two methods.

In the previous analysis, we have shown that the CPRank method has obviously good performance in identifying high-quality papers while we select the Milestone

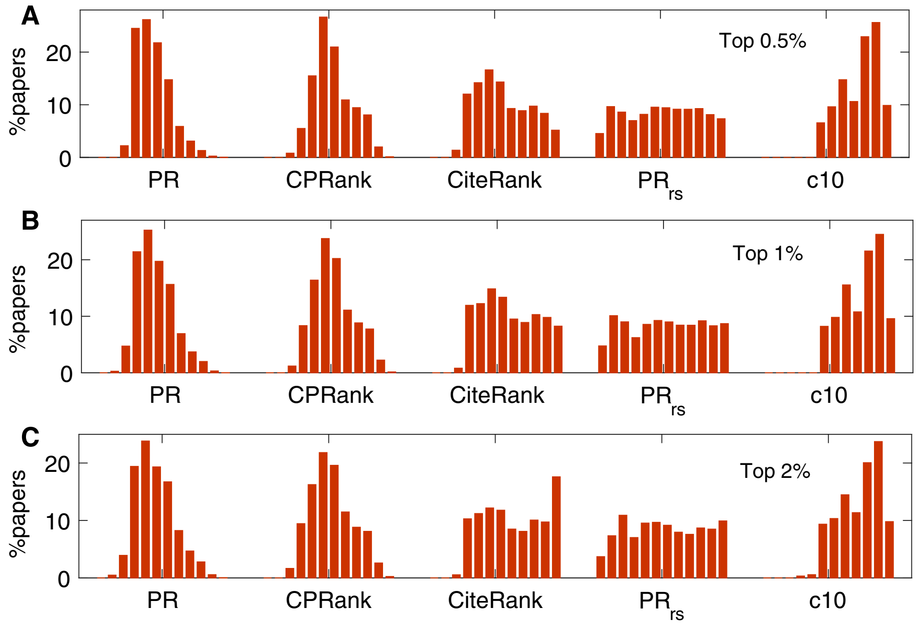

Fig. 6 (Color online) The percentage of articles in different decades in the top 0.5%, 1% and 2% of the general ranking with different ranking metrics

Letters as benchmark papers. In this section, we will discuss CPRank's performance on Nobel prize winning papers. Here, we picked 23 Nobel prize winning articles in Physics published from the year 1961–2000 as our benchmark articles. We compare these papers' average rankings in Fig. 7a. It can be seen that in Fig. 7a, similar to the result of milestone papers, the CPRank method performs best among these methods. Since our method has a great performance at the article level, then how does it perform at the author level? Due to the detailed information of the raw data, we can gain access to the information of authors of each paper. As a result, we are able to identify all articles written by an author. We calculate the sum of the scores of all articles for each author

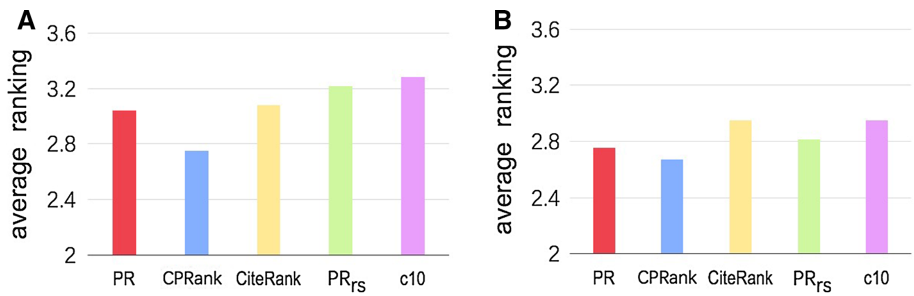

Fig. 7 (Color online) **a** The average ranking of Nobel Prize winning papers. **b** The average ranking of authors awarded Nobel Prize. (Note that the ranking values have been taken the logarithm before the calculation of the average)

and this sum represents the author's score. On the basis of these scores, we can rank the authors by sorting their scores in descending order. We selected 27 scholars who were awarded Nobel prize in physics as a benchmark set to examine the performance of different methods in identifying outstanding scholars. Something worthy to note is that these scholars are the authors of 23 Nobel prize winning papers mentioned above. The average rankings of Nobel laureates in different methods are shown in Fig. 7b. One can observe that the average ranking of authors awarded Nobel Prize in CPRank is better than in other methods.

We further explore the rankings of Nobel prize winning paper of Nobel laureates among all their personal papers in different measures. The result is shown in Fig. 8. Each arm is the result of a Nobel laureate and there are 20 circles in each arm which are top-20 papers (ranking by total citation count) of the author. Closer to the core means higher rank. The ranking positions of Nobel prize winning papers are marked by red, blue, yellow, green and purple in PageRank, CPRank, CiteRank, rescaled PageRank and citation rank respectively. One can see in Fig. 8 that in general the blue circles are closer to the core than circles in any other colors, indicating that the CPRank method outperforms other methods in identifying each author's highest quality paper.

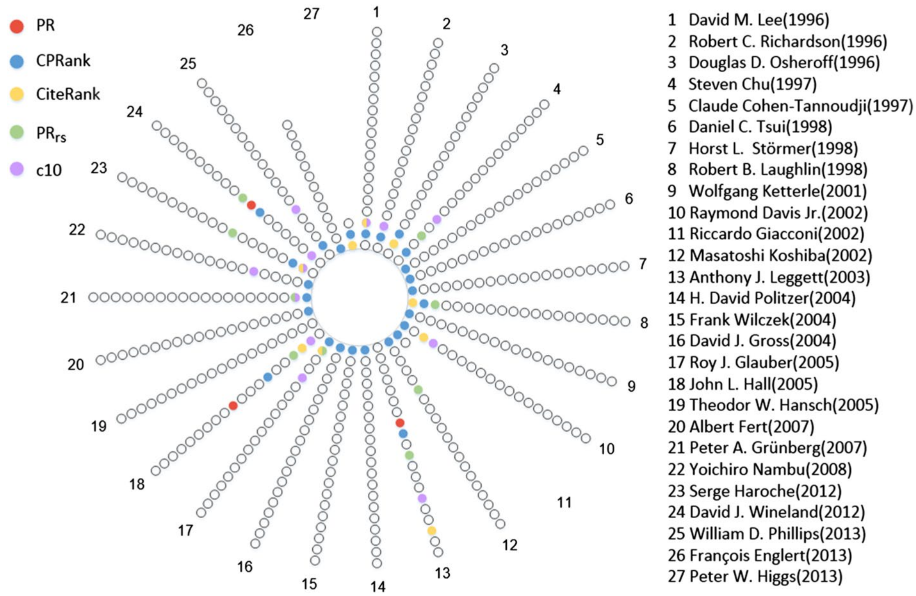

Fig. 8 (Color online) The ranks of Nobel prize winning papers with PageRank, CPRank, CiteRank, rescaled PageRank and citation count methods. Each arm is the result of a Nobel laureate. Their names and the years when they won the Nobel prize are listed on the right side of the figure. The circles from core to periphery represent the ranks of each Nobel laureate's top-20 papers (the arm with fewer than 20 circles means that the Nobel laureate published fewer than 20 papers in APS journals). The ranking positions of Nobel prize winning papers are marked by red, blue, yellow, green and purple respectively in PageRank, CPRank, CiteRank, rescaled PageRank and citation rank. The closer the circle to the core, the higher the rank is. If the ranking position obtained by the PageRank, CiteRank, rescaled PageRank and citation count methods are equal with that from the CPRank method, we also apply blue color to mark the circle. With respect to other methods except the CPRank method, the circles are divided by two colors when the ranking positions are equal in two method

Discussion

Objectively ranking the influence of scientific publications has been a challenging research focus for a long time in scientometrics. The PageRank algorithm now is widely considered to be a better metric than the number of citations because it takes into the global information of the network, but one should accept that the algorithm has time bias for old articles. In this paper, we propose a new iterative ranking algorithm which only considers the citations within 10 years after the publication of a paper to eliminate the time accumulation advantage of old papers. By validating in the APS citation network, we find that the new method can improve the effectiveness with respect to identifying the influential papers as well as the outstanding scholars.

In fact, according to the CPRank method, if a paper does not succeed in the earliest 10 years after publication, it has zero chance to be considered as an influential paper. However, some influential papers (e.g. "sleeping beauties") receive many more citations long time after their publication (Ke et al. 2015). Besides, receiving many citations long after publication might be a manifestation of long-term significance (Wasserman et al. 2015). This effect is neglected by CPRank and the existing c_{10} metric. There are some modifications that can be made to overcome it. For instance, one can extend the first 10 year citation in CPRank and c_{10} to 20 or even longer. Another possibility is to replace the first 10 years after publication with the 10 years where the paper acquires the largest number of citations.

There are many other extensions that could be made. As mentioned in the previous text, one of the limitations of the CPRank method is that it is still biased and cannot make a perfectly fair comparison of papers from different years. More effort is needed to solve this issue in the future research. Moreover, " c_{10}" is originally used to find out the highest-impact work published in which position in the sequence of papers published by a scientist during his/her career. It inspires us that maybe we can explore the timing of a scientist's most outstanding achievement with CPRank, which could offer a better understanding of the emergence of scientific excellence and might improve our ability to nurture high-impact scientists.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant Nos. 61603046, 61374175, 61573065) and the Natural Science Foundation of Beijing (Grant No. L160008).

References

Amsterdamska, O., & Leydesdorff, L. (1989). Citations: Indicators of significance? *Scientometrics*, *15*, 449–471.

Bollen, J., Rodriquez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, *69*(3), 669–687.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1–7), 107–117.

Chen, P., Xie, H., Maslov, S., & Redner, S. (2007). Finding scientific gems with google's pagerank algorithm. *Journal of Informetrics*, *1*(1), 8–15.

Ding, Y. (2011). Applying weighted PageRank to author citation networks. *Journal of the American Society for Information Science and Technology*, *62*(2), 236–245.

Ding, Y., Yan, E., Frazho, A., & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, *60*(11), 2229–2243.

Fiala, D. (2012a). Time-aware PageRank for bibliographic networks. *Journal of Informetrics*, *6*(3), 370–388.

Fiala, D. (2012b). Bibliometric analysis of CiteSeer data for countries. *Information Processing and Management*, *48*(2), 242–253.

Fiala, D., & Tutoky, G. (2017). PageRank-based prediction of award-winning researchers and the impact of citations. *Journal of Informetrics*, *11*, 1044–1068.

Franceschet, M. (2010). The difference between popularity and prestige in the sciences and in the social sciences: A bibliometric analysis. *Journal of Informetrics*, *4*, 55–63.
Frey, B. S., & Rost, K. (2010). Do rankings reflect research quality? *Journal of Applied Economics*, *13*, 1–38.
Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *Science*, *122*, 108–111.
Gonzalez-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegon, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, *4*, 379–391.
Ke, Q., Ferrara, E., Radicchi, F., & Flammini, A. (2015). Defining and identifying sleeping beauties in science. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(24), 7426.
Ma, N., Guan, J., & Zhao, Y. (2008). Bringing PageRank to the citation analysis. *Information Processing & Management*, *44*(2), 800–810.
Mariani, M. S., Medo, M., & Zhang, Y.-C. (2015). Ranking nodes in growing networks: When pagerank fails. *Scientific Reports*, *5*, 16181.
Mariani, M. S., Medo, M., & Zhang, Y.-C. (2016). Identification of milestone papers through time-balanced network centrality. *Journal of Informetrics*, *10*(4), 1207–1223.
Maslov, S., & Redner, S. (2008). Promise and pitfalls of extending google's PageRank algorithm to citation networks. *The Journal of Neuroscience*, *28*(44), 11103–11105.
Nykl, M., Jezek, K., Fiala, D., & Dostal, M. (2014). PageRank variants in the evaluation of citation networks. *Journal of Informetrics*, *8*(3), 683–692.
Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17268–17272.
Radicchi, F., Fortunato, S., Markines, B., & Vespignani, A. (2009). Diffusion of scientific credits and the ranking of scientists. *Physical Review E*, *80*(5), 056103.
Rendner, S. (2004). Citation statistics from more than a century of physical review. *Physics Today*, *58*, 49.
Sidiropoulos, A., & Manolopoulos, Y. (2006). Generalized comparison of graph-based ranking algorithms for publications and authors. *The Journal of Systems and Software*, *79*, 1679–1700.
Sinatra, R., Wang, D., Deville, P., Song, C., & Barabsi, A. L. (2016). Quantifying the evolution of individual scientific impact. *Science*, *354*(6312), aaf5239–aaf5239.
Sorzano, C. O. S., Vargas, J., Caffarena-Fernndez, G., & Iriarte, A. (2014). Comparing scientific performance among equals. *Scientometrics*, *101*, 1731–1745.
Su, C., Pan, Y., Zhen, Y., Ma, Z., Yuan, J., Guo, H., et al. (2011). PrestigeRank: A new evaluation method for papers and journals. *Journal of Informetrics*, *5*(1), 1–13.
Walker, D., Xie, H., Yan, K. K., & Maslov, S. (2007). Ranking scientific publications using a model of network traffic. *Journal of Statistical Mechanics*, *06*, P06010.
Wang, D., Song, C., & Barabasi, A. L. (2013). Quantifying long-term scientific impact. *Science*, *342*(6154), 127–132.
Wasserman, M., Zeng, X. H. T., & Amaral, Lu A. Nunes. (2015). Cross-evaluation of metrics to estimate the significance of creative works. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5), 1281–6.
Yan, E. (2014). Topic-based PageRank: Toward a topic-level scientific evaluation. *Scientometrics*, *100*(2), 407–437.
Yan, E., & Ding, Y. (2011). Discovering author impact: A PageRank perspective. *Information Processing and Management*, *47*(1), 125–134.
Yao, L. Y., Wei, T., Zeng, A., Fan, Y., & Di, Z. R. (2014). Ranking scientific publications: The effect of nonlinearity. *Scientific Reports*, *4*, 6663.
Zeng, A., Shen, Z. S., Zhou, J. L., Wu, J. S., Fan, Y., Wang, Y. G., et al. (2017). The science of science: From the perspective of complex systems. *Physics Reports*, *714–715*, 1–73.
Zhou, J. L., Zeng, A., Fan, Y., & Di, Z. R. (2016). Ranking scientific publications with similarity-preferential mechanism. *Scientometrics*, *106*, 805–816.