Predicting future influence of papers, researchers, and venues in a dynamic academic network

Fang Zhang[1] and Shengli Wu[1,2,*]
1. School of Computer Science, Jiangsu University, China
2. School of Computing, Ulster University, UK
*E-mail: s.wu1@uslter.ac.uk

ABSTRACT

Performance evaluation and prediction of academic achievements is an essential task for scientists, research organizations, research funding bodies, and government agencies alike. Recently, heterogeneous networks have been used to evaluate or predict performance of multi-entities including papers, researchers, and venues with some success. However, only a minimum of effort has been made to predict the future influence of papers, researchers and venues. In this paper, we propose a new framework WMR-Rank for this purpose. Based on the dynamic and heterogeneous network of multiple entities, we extract seven types of relations among them. The framework supports useful features including the refined granularity of relevant entities such as authors and venues, time awareness for published papers and their citations, differentiating the contribution of multiple coauthors to the same paper, amongst others. By leveraging all seven types of relations and fusing the rich information in a mutually reinforcing style, we are able to predict future influence of papers, authors and venues more precisely. Using the ACL dataset, our experimental results demonstrate that the proposed approach considerably outperforms state-of-the art competitors.

Keywords

Academic influence prediction, Dynamic academic network, Paper citation, Mutual reinforcement

1. Introduction

At the beginning of modern science, the main function of papers lies in "communicating and demonstrating". On the one hand, researchers communicate with academic peers through writing and publishing papers; on the other hand, papers have always been the most powerful means in the reward system of science. With the gradual institutionalization of scientific research activities and an increasing number of papers being published every year, the evaluation and prediction of the impact of papers and authors have become increasingly prominent. The results can be beneficial to government agencies, research funding bodies, research organizations, individual researchers, and other parties in many different ways. For instance, such results are useful for a government agency to set scientific development strategies and policies, or for a research funding body to evaluate research proposals and applicants. Such results may also be helpful for individual researchers and research organizations. Identifying high-quality new papers quickly in a related field is often a demanding and challenging task to researchers, especially to junior researchers or to those just moving into a new research area. Accurately predicting the potential prestige of a venue may facilitate research organizations and individual researchers to decide where to publish their papers (Pajić, 2015; Yu, Wang, Zhang, Zhang, & Liu, 2017).

Traditionally, citation count is a simple but useful metric for evaluating scientific impact (Garfield, 1972; Katerattanakul, Han, & Hong, 2015; Nerur, Sikora, Mangalaraj, & Balijepally, 2005), upon which some more complex metrics such as JIF, h-index (Hirsch, 2005) and g-index (Egghe, 2006) have been defined. These citation count-based metrics play an important role in evaluating papers, authors and venues, though they ignore some useful information such as co-authorship and citing papers (Wang et al., 2016). To address these problems, some graph-based methods have been proposed for ranking papers (Walker, Xie, Yan, & Maslov, 2006; Yan, Ding, & Sugimoto, 2011; Yan & Ding, 2010; Zhou, Zeng, Fan, & Di, 2016), or ranking papers, authors and venues simultaneously (Jiang, Sun, Yang, Zhuge, & Yao, 2016; Liu, Huang, Wei, & Mao, 2014; Wang et al., 2016). The graph-based ranking methods are likely more effective because they consider not only citation counts but also more information from the networks such as co-authorship, prestige of the citing papers, and so on.

Although a lot of effort has been made in evaluating and ranking the current influence of papers, researchers, and venues (Ding, 2011; Jiang et al., 2016; Liu et al., 2014; Meng & Kennedy, 2013; Zhou, Orshanshiy, Zha, & Giles, 2007), only three studies have been conducted to predict the future influence of papers, researchers, or venues (Kong, Zhou, Zhang, Wang, &

Xia, 2015; Sayyadi & Getoor, 2009; Wang et al., 2016). Performance prediction is also essential but more challenging than evaluation and ranking the current influence of an entity because of the uncertainty involved in the future. Certainly, these two tasks are related to each other. In a sense, evaluating the current performance of an entity can be seen as a special case of prediction with a very short period of time in the future. However, medium to long-term prediction needs more consideration. In this piece of work, we focus on a good solution that can predict the future influence of multi-entities including papers, researchers and venues at the same time by using a heterogeneous network of papers, researchers, and venues.

First of all, many previous studies only consider the average performance of a researcher and ignore that the performance of any researcher varies over time (Bertsimas, Brynjolfsson, Reichman, & Silberholz, 2013). On the one hand, for young researchers who begin to publish papers, their influence is undoubtedly much more limited than that of established researchers. However, they may obtain more influence over time with continuous learning and accumulating publications (Lee, 2019). Some of them may become very prestigious researchers after a certain period. On the other hand, even influential researchers in a field will see their influence decline, if they publish few or no papers for a lengthy period. Similarly, the prestige of a venue also varies over time. Therefore, it is important to consider these dynamics within any model for academic influence, especially when it comes to making predictions.

The second issue is how to compare papers published at different times. Initially, graph-based ranking methods such as PageRank and HITS do not consider the freshness of web pages. Their counterparts for academic paper evaluation have the same problem (Jiang et al., 2016; Li, Liu, & Yu, 2008; Wang et al., 2016), which is referred to as the ranking bias. Under these ranking systems, old papers take advantage of new ones. One explanation is that influence values are transferred from new papers to old ones in the citation network (Wang et al., 2016; Yu et al., 2017). To mitigate this problem, a partial solution is to assign different weights to papers published at different times. The weight of a paper decreases at a given rate as it ages (Kong et al., 2015; Li et al., 2008; Sayyadi & Getoor, 2009; Walker et al., 2006; Wang et al., 2016; Yu et al., 2017). This type of method works quite well for most papers, but they fail for the most influential (Casal, 2004). For instance, the paper that proposed the PageRank algorithm (Brin & Page, 1998) is still being cited heavily 20 years after publication. Therefore, more advanced treatment is required for such papers.

Thirdly, many classical measures are based on some simplified assumptions: all the citations and venues are equally important, while all the authors contribute equally to their coauthored paper, among others. These assumptions are questionable according to some researchers (Chakraborty & Narayanam, 2016; Du & Tang, 2013; Fujimagari & Fujita, 2014; Martin, Ball, Karrer, & Newman, 2013; Stallings et al., 2013; Wang, Tong, & Zeng, 2013; Zhu, Turney, Lemire, & Vellino, 2015). For example, for a number of authors who co-write and publish a paper together, many studies allocate the full credit of the publication to each of its coauthors (Kim & Diesner, 2014; Lindsey, 1980; Perianes-Rodriguez & Waltman, 2016). That is, every coauthor of the paper is regarded as producing the whole paper individually (Kim & Diesner, 2014). If one paper with five coauthors is cited 10 times, then each of its coauthors obtains 10 citations. Hence, overall 50 citations, rather than 10 citations in total, are allocated to this paper (Waltman, 2016). Such a phenomenon, referred to as the inflationary effect, may be questionable according to some studies (Perianes-Rodriguez & Waltman, 2016; Waltman, 2016) and should be addressed.

With a heterogeneous network including papers, researchers, and venues, we have rich information about all the entities involved. A PageRank-like algorithm is a good starting point because it is able to capture the interrelationships among the entities. All the above-mentioned issues need to be addressed in a systematic manner in order to achieve good performance.

In this study, we propose a framework, WMR-Rank (Weighted Mutual Reinforcement Ranking), to predict the future influence of papers, researchers, and venues at the same time. The proposed framework is based on the heterogeneous network of papers, authors, and venues. Seven types of relations among these entities are extracted and employed to simultaneously predict all of their future influence through an iterative process with mutual reinforcement. Overall, our approach is novel in several aspects as follows:

1. Every researcher or venue per year is treated as a separate entity. In this way, we are able to capture their dynamic nature more precisely.
2. For both researchers and venues, both their current performance, and their past performance for a number of years is taken into account to predict the future performance of all related entities.
3. For a balanced treatment of old and new papers, we consider both publication age and recent citations of all the papers involved at the same time.

Among these aspects, treating each entity per year is the most important one. It lays a solid foundation for many other measures. This is crucial for performance prediction, especially for medium to long-term prediction. One simple example is to compare two authors with similar average performance in recent years. However, if one of them is on the rise while the other is opposite in recent years, then it is predictable that the former may perform better than the later in the next couple of years. Otherwise, it is difficult to differentiate them if we only compare the average performance of them in the last couple of years. Another advantage of such a treatment is to make the framework more flexible. It enables us to do many different sorts of performance evaluation and prediction of multiple entities even with different metrics in the same framework.

Experiments with the ACL (Association for Computational Linguistics Anthology Network) dataset (Radev, Muthukrishnan, Qazvinian, & Abu-Jbara, 2013) show that the measures introduced are very effective and the proposed method considerably outperforms other state-of-the-art methods.

The remainder of this paper is organized as follows: Section 2 presents related work on performance evaluation or influence prediction of multi-entities including papers, researchers, and venues, mainly by using homogeneous or heterogeneous networks. Section 3 describes the framework in this study. Section 4 presents detailed experimental processes, settings, and results. Some analysis of the experimental results is also given. Section 5 concludes the paper.

2. Related work

How to evaluate scientific papers, authors and venues objectively is an important task, which has been investigated by many researchers for long time. Citation count is widely used for evaluating papers as well as researchers and venues (Jiang et al., 2016; Wang et al., 2016), and a lot of metrics are related to it. For the evaluation of journals, impact factor (Garfield, 1972) is the most common metric, which is the yearly average number of citations to the papers published in a given journal. Other metrics such as 5-year Impact Factor (IF) (Pajić, 2015), Eigenfactor Score (Bergstrom, 2007), Source Normalized Impact per Paper (SNIP) (Moed, 2010; Waltman, Eck, Leeuwen, & Visser, 2013), SCImago Journal Rank (SJR) (González-Pereira, Guerrero-Bote, & Moya-Anegón, 2010; González-Pereira, Guerrero-Bote, & Moya-Anegón, 2012), and sub-impact factor (SIF-) (Xu, Liu, & Rousseau, 2015) have also been proposed. Meanwhile, for the evaluation of researchers, a group of metrics such as the h-index (Hirsch, 2005), g-index (Egghe, 2006), R-index (Jin, Liang, Rousseau, & Leo, 2007), success-index (Franceschin, Galetto, Maisano, & Mastrogiacomo, 2012), DS-index (Farooq, Khan, Iqbal, Munir, & Shahzad, 2017), and others have been proposed. Among these, the h-index is likely the most popular.

The PageRank and HITS algorithms were initially proposed for ranking web pages in search engines. They have been adapted for evaluating academic performance by many studies (Dunaiski, Visser, & Geldenhuys, 2016; Liu et al., 2014; Qiao, Wang, & Liang, 2012; Walker et al., 2006; Yan & Ding, 2010; Zhang & Wu, 2018; Zhou et al., 2016). Because academic networks reflect the relations among papers, researchers, and venues, in a sense they are more complex than the web in which web pages are the only concern. Certain modification is necessary for algorithms such as PageRank and HITS to serve for academic performance evaluation.

In recent years, many different types of methods have been applied to academic performance ranking and prediction. Four models including linear regression, k-nearest neighbors, support vector regression, and a tree model for classification and regression were used for paper citation prediction in (Yu, Yu, Li, & Wang, 2014). A large number of features are extracted from papers and related sources, and then stepwise regression was used to select a good regression model for impact prediction of papers with a sub-set of features. Cao, Chen, & Liu (2016) proposed a citation pattern matched approach. First a large database of paper citations was set up. For a testing paper, its citation pattern was matched to those in the database. Then the citation patterns in those matched papers are used for citation prediction of the testing paper. A neural network-based method was used for paper citation prediction in (Abrishami & Aliakbary, 2019).

In the following, we review a number of works which are very relevant to this research. All of them are PageRank-like algorithms based on academic networks, and they will be used as baselines in this study for performance comparison.

Sayyadi and Getoor (2009) proposed FutureRank to predict the future influence of papers. Information about citation, authorship, and publication time is used in the algorithm. The axXiv (hep-th) data set, which contains approximately 28,000 articles published on high energy physics from 1993 to 2003, was used for evaluation. In their experiment, the future PageRank scores after three coming years were used as the ground-truth.

Liu et al. (2014) proposed Tri-Rank, which may rank researcher, papers, and venues at the same time. In their method, information about citation, authorship, and venue is used. Tri-Rank assigns different weights to authors in different ordering positions, and it also addresses the self-citation problem. Some data taken from ACM digital library were used for evaluation.

Kong et al. (2015) proposed TAPRank to rank researchers. Information on paper citation and authorship is used. More weights are given to new papers and new citations. The DBLP data set from ArnetMiner (Tang et al., 2008) with citation information captured from ACM Digital Library was used for evaluation. Future citation counts for five years are used as the ground-truth.

Jiang et al. (2016) proposed MutualRank, which may rank papers, authors and venues at the same time. It was observed that PageRank was biased to old papers, while HITS was biased to new papers. Therefore, a more balanced solution was taken in MutualRank. The ACL data set was used for the evaluation of the proposed method.

Wang et al. (2016) proposed MRCoRank to rank future popularity of four types of entities including papers, authors, venues, and terms through mutual reinforcement. The final ranking of multi-entities is generated through leveraging constructed text features and time-weighted citations. The DBLP data set from ArnetMiner (Tang et al., 2008) was used for experiments. Future citation counts for six years are used as the ground-truth.

Especially, three works (Sayyadi & Getoor, 2009; Kong et al., 2015; Wang et al., 2016) predict the future influence of entities. However, we find that there is limitation for using them. In particular, FutureRank only predicts paper influence and TAPRank only predicts researcher influence. Although MRCoRank is able to make prediction for multiple entities, it is designed for predicting papers published in the same year, which is an obstacle for practical use. If we try to make predictions beyond the same year, as we set out to do in this paper, MRCoRank's performance significantly deteriorates.

3. The WMR-Rank framework

In this section, we will detail all the components required for the framework and then present the multi-entity ranking algorithm.

3.1 The structure and major components of the heterogeneous network

The heterogeneous network of papers, authors and venues can be represented as $G = (V, E)$, where V denotes the set of nodes and E denotes the set of edges. There are three types of nodes: paper V_P, author V_A and venue V_V. There are three types of edges: author to paper (authorship) $E_{A\text{-}P}$, paper to venue (publication) $E_{P\text{-}V}$, and paper to paper (citation) $E_{P\text{-}P}$. Note in this network, only the paper to paper citation relation is directional, while other types of relations are not.

One feature that we want to support is the comparison of researchers' performance across different years. To enable this, we treat each author or venue in a specific year as a separate entity. Venues are treated in a similar vein.

Another feature our framework supports is differentiated treatment accorded to authors with differing contributions. If a paper is a collaborative work, then we distinguish their contributions by the ordering position of the authors in the paper (further details later). To enable this, we label every paper authorship edge, by "1st author", "2nd author", and so on. Fig. 1 shows an example of such a network.

In Fig. 1, authors, papers, and venues are represented by ovals, diamonds, and rectangles, respectively. There are two venues v(2000) and v(2001), four papers p_1, p_2, p_3, and p_4. and seven authors a_1(2000), a_2(2000), a_3(2000), a_4(2000), a_5(2001), a_6(2001), and a_1(2001). v(2000) and v(2001) are the same conference in two different years 2000 and 2001, while a_1(2000) and a_1(2001) are the same author in two different years 2000 and 2001. All three types of edges $E_{A\text{-}P}$ (paper authorship), $E_{P\text{-}V}$ (paper-venue), and $E_{P\text{-}P}$ (paper citation) can be found in Fig 1. For the rest of this paper, we will omit an explicit year label to maintain conciseness, although a year should always be assumed to be implicitly associated with each of these entities. For example, letting a_1 denote an author, within our framework, a specific year will always be associated with it, although not presented explicitly.

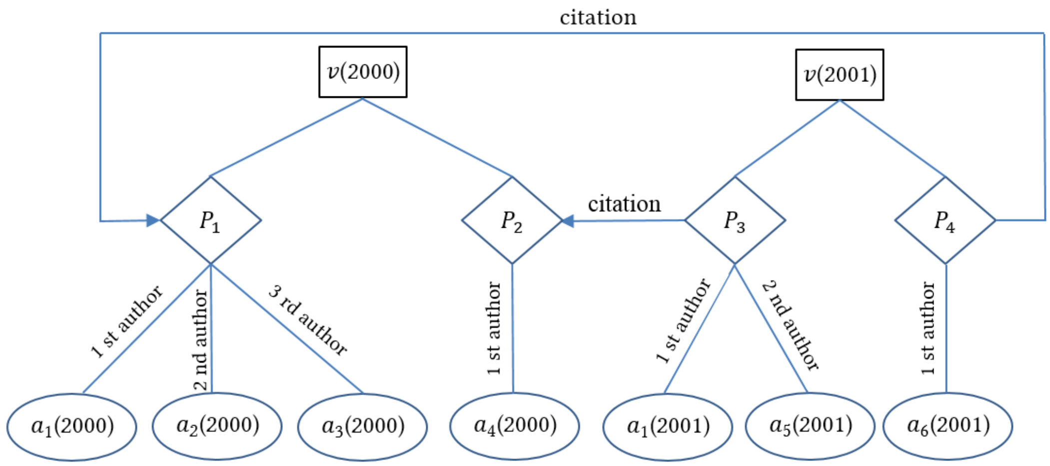


Fig. 1. An example of the heterogeneous network of authors, papers, and venues

3.1.1 Seven types of relations. Now let us discuss the seven types of relations in which we are interested. One key issue is how to measure the credit that one node may obtain from its relation with the other node directly or indirectly.

Paper citation (Type I) is one type of relation between two papers. If paper p_i is cited by p_j(denoted as $p_i \leftarrow p_j$) , then p_i obtains credit

$$W_{PP}(p_i, p_j) = \begin{cases} 1 & p_i \leftarrow p_j \\ 0 & otherwise \end{cases} \tag{1}$$

Through paper citation, we can observe an indirect relation of author citation (Type II). If there are two papers p_i and p_j, paper p_i is cited by p_j, a_m is the only author or one of the authors of p_i, and a_n is the only author or one of the authors of p_j, then a_m is cited by a_n. Much as other researchers (Abbas, 2011; Assimakis, 2010; Du & Tang, 2013; Egghe, Rousseau, & Hooydonk, 2000; Hagen, 2008; Hodge, Greenberg, & Challice, 1981; Hooydonk, 1997; Stallings et al., 2013), we take the same view that different authors may contribute differently to their collaborative paper. Under our framework, the credit that author a_m obtains from a_n through paper citation $p_i \leftarrow p_j$ is

$$W_{CA\_raw}\left(a_m, a_n, p_i, p_j\right) = \frac{1}{order(a_m, p_i) \times order(a_n, p_j)} \tag{2}$$

where $order(a, p)$ is the position of author a in paper p. It is necessary to normalize the above quantity to let the total credit that all the authors of p_i obtain from all the authors of p_j sums to 1. We have

$$W_{CA}\left(a_m, a_n, p_i, p_j\right) = \frac{W_{CA\_raw}\left(a_m, a_n, p_i, p_j\right)}{\sum_{\substack{p_i \leftarrow p_j \\ a_k \in P_A(p_i) \\ a_l \in P_A(p_j)}} W_{CA\_raw}\left(a_k, a_l, p_i, p_j\right)} \tag{3}$$

where $P_A(p)$ is the set of all the authors of paper p.

An author a_n may cite another author a_m multiple times. The total credit that a_m obtains from a_n is the summation of all the papers involved.

$$W_{CA}(a_m, a_n) = \sum_{\substack{p_i \in A_P(a_m) \\ p_j \in A_P(a_n) \\ p_i \leftarrow p_j}} W_{CA}\left(a_m, a_n, p_i, p_j\right) \tag{4}$$

where $A_P(a)$ is the set of papers written by author a.

The co-authorship relation (Type III) exists in the network if two or more author nodes connect to the same paper node. Any author obtains some credit from all other authors if they write a paper together. The credit that a_i from her co-author a_j through paper p is defined as

$$\mathrm{W}_{CoA\_raw}(a_i, a_j, p) = \frac{1}{order(a_i, p) \times order(a_j, p)} \tag{5}$$

After normalization, we have

$$\mathrm{W}_{CoA}(a_i, a_j, p) = \frac{\mathrm{W}_{CoA\_raw}(a_i, a_j, p)}{\sum_{a_k, a_l \in P_A(p)} \mathrm{W}_{COA\_raw}(a_k, a_l, p)} \tag{6}$$

Two authors may co-write more than one paper. Hence, the credit that a_i obtains from a_j over all co-authored papers is

$$\mathrm{W}_{CoA}(a_i, a_j) = \sum_{\substack{p \in A_P(a_i) \\ p \in A_P(a_j)}} \mathrm{W}_{CoA}(a_i, a_j, p) \tag{7}$$

where $A_P(a_i)$ denotes all the papers written by a_i. Author a_i is able to obtain credit from another author a_j through paper citation and co-writing papers.

$$W_{AA}(a_i, a_j) = W_{CA}(a_i, a_j) + W_{CoA}(a_i, a_j) \tag{8}$$

We may define venue citation (Type IV) in a similar way to author citation. We let

$$W_{VV}(v_i, v_j) = \frac{CitationNum(v_i, v_j)}{CitationNum(v_j)} \tag{9}$$

where v_i, v_j are two venues. $CitationNum(v_j)$ denotes the number of all the citations that venue v_j obtains, while $CitationNum(v_i, v_j)$ denotes the number of citations that v_j obtains from venue v_i.

Paper co-authorship happens very often. In this study, we let all the papers be assigned the same amount of credit. However, for a group of co-authors, their contributions to the same paper are differentiated by their ordering positions (Abbas, 2011; Assimakis, 2010; Du & Tang, 2013; Egghe et al., 2000; Hagen, 2008; Hodge et al., 1981; Hooydonk, 1997; Stallings et al., 2013). In this paper, we adopt the geometric counting approach (Egghe et al., 2000) in paper-author relation (Type V). Suppose author a_i is in the R^{th} position among all T coauthors in paper p_j, then author a_i and paper p_j obtain credit from each other

$$\mathrm{W}_{AP}(a_i, p_j) = \mathrm{W}_{PA}(p_j, a_i) = \frac{2^{T-R}}{2^T - 1} \tag{10}$$

If paper p_i is published in venue v_j (Type VI, paper-venue relation), then there is an edge between paper p_i and venue v_j, so that paper p_i and venue v_j get credit from each other. We let

$$W_{VP}(v_j, p_i) = \mathrm{W}_{PV}(p_i, v_j) = \begin{cases} \dfrac{1}{paperNum(v_j)} & p_i \in V_P(v_i) \\ 0 & otherwise \end{cases} \tag{11}$$

where $paperNum(v_j)$ is the number of papers published in venue v_j.

If author a_i publishes more than one paper in venue v_i (Type VII, author-venue), then the credit a_i obtains from v_i is the sum of credit she obtains from all the papers published in v_i. Conversely, v_i obtains an equal amount of credit from a_i.

$$\mathrm{W}_{AV}(a_i, v_j) = \mathrm{W}_{VA}(v_j, a_i) = \frac{1}{paperNum(v_j)} \sum_{\substack{p_k \in A_P(a_i) \\ p_k \in V_P(v_j)}} W_{AP}(a_i, p_k) \tag{12}$$

$1/paperNum(v_j)$ serves as the normalization coefficient. In this way, all the weights for any venue sum to one. For all above-mentioned relations, unless specified explicitly, the default is zero.

3.1.2 Time-aware weights. In order to predict the future influence of related entities precisely, it is necessary to take the time information of publications and citations into consideration (Kong et al., 2015; Sayyadi & Getoor, 2009; Walker et al., 2006; Wang et al., 2016), as recognised by some prior researchers. However, there is no consensus on how time affects influence, let alone a good representation for its effect. For example, Sayyadi and Getoor (2009) hypothesized that recently published papers might obtain more citations in the future. This view was echoed by Kong et al. (2015). However, Wang et al. (2016) had different opinions about this, claiming that not all new papers would obtain more citations than the old ones, and most papers obtained just a few citations. In this paper, we mainly consider the past performance of an entity for its prediction. In particular, we pay more attention to the more recent performance of entities. This is the third characteristic of our method.

For paper p_i, we assume that its publishing year is $T_{published}(p_i)$, the evaluation year is $T_{evaluate}$, and l latest citations to p_i are in years $t_1, t_2, \dots, t_l$. $T_{avgCitation\_l}(p_i) = (t_1 + t_2 + \cdots + t_l)/l$. Here we set $l = T_{evaluate} - T_{published}(p_i)$. If p_i does not have l citations, then we set the year in which the paper was published, $T_{published}(p_i)$ for those non-existing citations. Thus the time-aware weight of paper p_i is set to be

$$T(p_i) = \frac{e^{-\sigma_1(T_{evaluate} - T_{avgCitation\_l}(p_i))}}{T_{evaluate} - T_{published}(p_i)} \tag{13}$$

where σ_1 is a positive decaying parameter. When the difference between $T_{evaluate}$ and $T_{published}(p_i)$ is larger, $T(p_i)$ becomes smaller; on the other hand, when the difference between $T_{evaluate}$ and $T_{avgCitation}(p_i)$ is smaller, $T(p_i)$ is larger. In summary, $T(p_i)$ is biased to newer papers which are given heavier weights. However, if an old paper can still attract citations in recent years, then it will be given a relatively heavy weight. The rationales behind this are twofold. On the one hand, new papers are more likely to attract more citations than old papers in the future. On the other hand, some good old papers should not be punished overly if they are still cited by other papers recently. The above time weighting schema is a balanced policy for dealing with these two conflicting factors. In the following we treat authors and venues in the same vein.

For author a_i (an author in a specific year, to be more accurate), assume that she published a group of papers $p_1, p_2, \dots, p_m$ in this specific year, or $T_{published}(a_i)$, and the m latest citations to these papers are in years $t_1, t_2, \dots, t_m$. We define $T_{avgCitation\_m}(a_i) = (t_1 + t_2 + \cdots + t_m)/m$. If there are not enough citations, we set the earliest year, $T_{published}(a_i)$, for those non-existing citations. The time-aware weight of author a_i is set to be

$$T(a_i) = \frac{e^{-\sigma_2(T_{evaluate} - T_{avgCitation\_m}(a_i))}}{T_{evaluate} - T_{published}(a_i)} \tag{14}$$

Analogously, the time-aware weight of venue v_i can be set as

$$T(v_i) = \frac{e^{-\sigma_3(T_{evaluate} - T_{avgCitation\_n}(v_i))}}{T_{evaluate} - T_{published}(v_i)} \tag{15}$$

where $T_{published}(v_i)$ denotes the year in which venue v_i was held, and $T_{avgCitation\_n}(v_i)$ denotes the average time of the n latest citations to any of the papers in v_i, where n is the number of papers published in v_i. Again, if there are not enough citations, we take a similar treatment as to $T(p_i)$ and $T(a_i)$. With all these weights defined, we can present the multi-entity ranking algorithm, as we will do in the next section.

3.2 The multi-entity ranking algorithm

Based on the information provided by the heterogeneous network, the multi-entity ranking algorithm, WMR-Rank, scores every entity involved. After setting the initial values for all the entities, an iterative process is applied to them, and at each step every entity obtains an updated score. Note that all the entities involved affect each other. The algorithm stops when a termination condition is satisfied (e.g., a given number of iterations or a threshold for the updates). Algorithm 1 gives the details of the proposed method.

ALGORITHM 1: WMR-Rank

Procedure WMRank ($\alpha, \beta, \gamma, \sigma_1, \sigma_2, \sigma_3, \lambda, \varepsilon, \mu$)
Input: the entity sets V_P, V_A, V_V, edge sets E_{PP}, E_{PA}, E_{PV}, E_{CA}, E_{COA}, E_{AV}, E_{VV}, weight matrixes W_{PP}, W_{PA}, W_{PV}, W_{AA}, W_{AV}, W_{VV}, and parameters $\alpha, \beta, \gamma, \sigma_1, \sigma_2, \sigma_3, \lambda, \varepsilon, \mu$, t_a, t_v
Output: the scores of all the papers, authors and venues PS, AS and VS
// The loop in lines 1-3 initializes PS, AS, and VS scores for papers, authors, and venues, respectively
1 **For** every $p_i \in V_P$, $a_j \in V_A$, $v_k \in V_V$
2 $PS(p_i)^0 = \frac{1}{|V_P|}, AS(a_j)^0 = \frac{1}{|V_A|}, VS(v_k)^0 = \frac{1}{|V_V|}$ //$|V|$: the number of entities in V.
3 **End For**
4 t=-1, Δ =2 ε
5 **while** Δ > ε **do**
6 t=t+1
// Calculate the past performance of authors
7 **For** every $a_i \in V_A$, $pas(a_i)^{t+1} = \frac{1}{t_a+1}[\sum_{a_k \in pn(a_i,t_a)} AS(a_k)^t + AS(a_i)^t]$ **End For**
// Calculate the past performance of venues
8 **For** every $v_i \in V_V$, $pvs(v_i)^{t+1} = \frac{1}{t_v+1}[\sum_{v_k \in pn(v_i,t_v)} VS(v_k)^t + VS(v_i)^t]$ **End For**
/* The following loop calculates PS scores for papers by considering past performance of their authors and venues (line 10), citation (line 11), and publication time (line 12) */
9 **For** every $p_i \in V_P$
10 $temp(p_i)^{t+1} = \alpha \sum_{a_j \in P_A(p_i)} W_{PA}(p_i, a_j) \times pas(a_j)^{t+1} + (1-\alpha) \times W_{PV}(p_i, v_k) \times pvs(v_k)^{t+1}$
11 $temp(p_i)^{t+1} = \sum_{p_i \leftarrow p_j} W_{PP}(p_i, p_j) \times temp(p_j)^{t+1} + \mu \times temp(p_i)^{t+1}$
12 $PS(p_i)^{t+1} = \lambda \times T(p_i) \times temp(p_i)^{t+1} + (1-\lambda) \times \frac{1}{|V_P|}$
13 **End For**
/* The following loop calculates AS scores for authors by considering their paper and venue scores (line 15), citation and cooperation with other authors (line 16), and the time factor (line 17) */
14 **For** every $a_i \in V_A$
15 $temp(a_i)^{t+1} = \beta \sum_{p_j \in A_P(a_i)} W_{AP}(a_i, p_j) \times PS(p_j)^t + (1-\beta) \sum_{v_k \in A_V(a_i)} W_{AV}(a_i, v_k) \times pvs(v_k)^{t+1}$
16 $temp(a_i)^{t+1} = \sum_{a_j \in A_{AA}(a_i)} W_{AA}(a_j, a_i) \times temp(a_j)^{t+1} + \mu \times temp(a_i)^{t+1}$
17 $AS(a_i)^{t+1} = \lambda \times T(a_i) \times temp(a_i)^{t+1} + (1-\lambda) \times \frac{1}{|V_A|}$
18 **End For**
/* The following loop calculates VS scores for venues by considering their paper and authors' scores (line 20), and the time factor (line 21) */
19 **For** every $v_i \in V_V$
20 $temp(v_i)^{t+1} = \gamma \sum_{p_j \in V_P(v_i)} W_{VP}(v_i, p_j) \times PS(p_i)^t + (1-\gamma) \sum_{a_k \in V_A(v_i)} W_{VA}(v_i, a_k) \times pas(a_k)^{t+1}$
21 $VS(v_i)^{t+1} = \lambda \times T(v_i) \times \sum_{v_i \leftarrow v_j} W_{VV}(v_i, v_j) \times temp(v_j)^{t+1} + (1-\lambda) \times \frac{1}{|V_V|}$
22 **End For**
23 Normalize $PS(p_i)^{t+1}$, $AS(a_i)^{t+1}$, and $VS(v_i)^{t+1}$
24 Δ= $\sum_{p_i \in V_P} |PS(p_i)^{t+1} - PS(p_i)^t| + \sum_{a_i \in V_A} |AS(a_i)^{t+1} - AS(a_i)^t| + \sum_{v_i \in V_V} |VS(v_i)^{t+1} - VS(v_i)^t|$
25 **End do**

Lines 1-3 initialize all three types of entities involved. For all the entities in the same category, they are assigned the same score of 1/*n*, where *n* is the number of entities in that category. Line 7 calculates the average performance in all t_a+1 years for each of the authors. $pn(a_i, t_a)$ is a function that returns a set of t_a entities of the same author a_i in the previous t_a years. Suppose that a_i is in the year of y, we can express a_i as $a_i(y)$, so that $pn(a_i, t_a) = \{a_i(y-1), a_i(y-2), \ldots, a_i(y-t_a)\}$. Line 8 calculates the average performance of t_v+1 years for each of the venues. $pn(v_k, t_v)$ is a function that returns a set of t_v entities of the same venue v_k in the previous t_v years.

The main part of the algorithm is included in a while loop. Inside the loop (lines 5-27), the scores for all the nodes involved are updated. All papers' new scores are calculated in lines 10-12. Three factors are considered: authors (line 10), venues (line

10), and citations (line 11). All authors' new scores are calculated in lines 15-17. Four factors are considered: published papers (line 15), the venues in which the papers are published (line 15), author citations (line 16), and coauthors to the published papers (line 16). All venues' new scores are calculated in line 21. Three factors are considered: published papers (line 20), authors (line 20), and venue citations (line 21). For all three types of nodes, the time factor is considered: line 12 for papers, line 17 for authors, and line 21 for venues.

A boost is given to papers whose publication years are close to the evaluation year. It is especially necessary for those without any citations, since they will not able to obtain any score otherwise. In line 11, a parameter μ is used to adjust this factor. Any paper published within the last five years, even not cited by any papers, will still be able to obtain some credit. Likewise, authors are treated in the same way (line 16).

4. Experiment

4.1 Dataset

In this experiment, we use the ACL dataset[1] (Radev et al., 2013), which is constructed from the papers published in natural language processing venues (including journals, conferences and workshops from 1965 to 2014).

Before carrying out the experiments, we pre-process the dataset as follows. First, we remove those papers that neither cite any other papers, nor receive any citations, because it is difficult for us to evaluate their influence, since they are detached completely. Second, Manual disambiguation of author and venue entities is done to let the data set more usable. Third, all of the joint conferences are considered to have dual identity. For example, COLING-ACL'2006 is a joint conference of COLING and ACL, therefore both COLING'2004 (which is held every 2 years) and ACL'2005 are equally treated as its previous conferences, and COLING-ACL'2006 is regarded as a previous conference by both COLING and ACL in later years. Third, apart from regular papers, many conferences publish short papers, student papers, demos, posters, tutorials, etc. On average, the regular papers are of a higher quality than non-regular papers. Therefore, we split the proceedings of such a conference into two, and let all regular papers remain in the main conference, while putting all non-regular papers into its companion (Jiang et al., 2016). Finally, for those papers with more than 5 authors, we only retain the first five authors and ignore the rest. In this dataset, 1175 papers (approximately 5.1%) have more than five authors. After pre-processing, the data set includes 19891 papers, 15379 authors and 372 venues without considering time, or 34917 authors (per year) and 638 venues (per year).

As an additional measure, author name disambiguation is carried out for a subset of the papers in the collection. They are papers published between 1985 and 2004 (which is referred to as the modelling partition, Section 4.2 gives more details about it). In this subset, there are 6767 papers and 5311 different author names. For each of the names, we extract affiliation information for all the papers by the given name. We assume that if affiliation is the same for all the papers by a certain name, then the author is uniquely identified. Under all these names, 2891 has only one publication; the affiliation part is the same organization for other 1294. Therefore, no further action is needed for these 2891+1294 names and we need to check the remaining 1126 names: for each of them, two or more organizations appear in the affiliation part of different papers. Because for the same author, she is likely to continue her research on the same or similar research topics and/or cooperate with her former colleagues even she moves to a new place. Thus 785 names can be identified by considering these two factors. For the remaining 341 names, we use some extra resources. Both web sites www.researchgate.net and www.arminer.cn keep publication information for a large number of authors. We refer to these two sites to identify papers by the same author. In some cases, we also look at some personal web pages to find useful information. Finally, we find that 294 names are used by one author and 26 names are used by two authors. However, for the last 21 names, after all above-mentioned measures it is still unclear if these names are used by one or more authors. Among these 21 names, 20 are associated with two institutions and one is with three institutions. If we treat each of these names with a different affiliation as a separate author, then we have 22 more authors. For all the results presented later in this paper, instead of using 5311 authors, 48 more authors are added for all the experiments. However, because the scale of the problem is small: the number of authors added is less than 1% of the total number of the authors (48/5311); and the number of papers from these additional authors is 51, also less than 1% of the total number of papers in the whole partition (51/6767). Therefore, its effect on prediction results is very small and can be ignored.

[1] See http://clair.eecs.umich.edu/aan/index.php.

4.2 Setting and evaluation metrics

In order to evaluate the future influence of all the papers, venues, and authors involved, we split the dataset into two partitions: the first for modelling and the second for evaluation. The modelling partition contains all the papers published between 1985 and 2004, and the evaluation partition contains all the papers published between 2005 and 2014. We extract the seven relations based on the modelling partition and run Algorithm 1 to set up the ranking model for all the papers, authors and venues based on the modelling partition, and then evaluate the prediction accuracy based on the evaluation partition.

For parameter setting of WMR-Rank, we let α=0.70, β=0.60, γ=0.10, λ=0.85, σ_1=0.5, σ_2=1.0, σ_3=1.1. Four different values, 0.8, 0.6, 0.4, and 0.2 are assigned to μ, for papers published in 2003 and 2004, 2002, 2001, and 2000, respectively. More detailed discussion about these parameters is given in Sections 4.5 and 4.6.

Currently, there is no universally accepted metric to assess the performance or influence of a paper or an author, though citation count is widely used (Kong et al., 2015; Sayyadi & Getoor, 2009; Wang et al., 2016; Zhang & Wu, 2018; Zhou et al., 2016). In this article, we adopt this metric to evaluate the future influence of papers. One problem of using citation count directly for authors is the inflationary effect. The second problem is that it allocates equal scores to all the authors involved. In order to avoid such problems, we use WCC (Weighted Citation Count) instead of citation count to measure the influence of authors. For author a_i, her influence score is defined as

$$WCC(a_i) = \sum_{p_j \in A_P(a_i)} W_{AP}\left(a_i, p_j\right) \times CC(p_j) \tag{16}$$

where $CC(p_j)$ is the citation count of paper p_j. See Equation 10 in Section 3 for the definition of $W_{AP}\left(a_i, p_j\right)$. The final score that a_i obtains, $WCC(a_i)$, is the sum of scores of all the papers that a_i writes.

For venues, we use MIF (Modified Impact Factor) to measure their influence (Yan & Lee, 2007). For a given venue *v*, its MIF(v) is defined as

$$MIF(v) = \frac{CitationNum(v)}{|v|} \tag{17}$$

where $CitationNum(v)$ is the total number of citations that all the papers published in venue *v* obtain, and $|v|$ is the total number of papers in venue *v*. Therefore, $MIF(v)$ is the average number of citations that every paper in v obtains.

For any given method to predict future influence, we apply it to the modelling partition to obtain scores for all the papers, venues, and authors involved, before observing their performance score in the evaluation partition. Thus, prediction accuracy can be calculated by comparing how consistent the two groups of rankings or scores are. Spearman's rank correlation coefficient (r_s) or Pearson correlation coefficient (ρ) are commonly used for this purpose (Sayyadi & Getoor, 2009; Zhou et al., 2016).

Given two sets of scores $X = \{x_{\mathrm{i}} | i = 1,2,\cdots,n\}$ and $Y = \{y_i | i = 1,2,\cdots,n\}$, the Pearson correlation coefficient between *X* and *Y* is defined by

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x_\iota}) \times (y_i - \bar{y_\iota})}{\sqrt{\sum_{x_i \in X}(x_i - \bar{x})^2}\sqrt{\sum_{y_i \in Y}(y_i - \bar{y})^2}} \tag{18}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of *X* and *Y*.

For the above-mentioned sets of scores *X* and *Y*, we may also calculate their Spearman's correlation coefficient, which is defined as the Pearson correlation coefficient of r_X and r_Y, where r_X and r_Y represent the ranked variables of *X* and *Y*, respectively. Although these two measures are related to each other, they are not the same. Spearman's correlation assesses monotonic relationships, while Pearson's correlation assesses linear relationships.

4.3 Methods for comparison

The ranking algorithms used for comparison are as follows:

1. **FutureRank (FR)**. It is a method of predicting the future citations of papers by using the authorship network and the publication time of papers (Sayyadi & Getoor, 2009).
2. **Citation Count (CC)**. Citation count is widely used to assess the influence of papers, authors and venues because it is single-valued and easy to interpret (Yan & Ding, 2010; Zhu et al., 2015). In our experiment, we predict the future influence of papers, authors and venues based on their current citation count.
3. **MIF**. This metric is used to evaluate the performance of venues (Yan & Lee, 2007). We may predict the future influence of venues based on their performance in the modelling partition using MIF, just as we use CC for authors and papers.
4. **MutualRank (MR)**. A state-of-the-art method ranking papers, authors and venues simultaneously by integrating mutual reinforcement relationships in heterogeneous networks (Jiang et al., 2016).
5. **MRCoRank (MRCR)**. A state-of-the-art method to rank the future influence of papers, authors and venues simultaneously (Wang et al., 2016).
6. **PageRank (PR)**. A famous algorithm that is originally designed to rank web pages, which has then been widely applied in many other applications including academic assessment. For example, some studies (Ding, Yan, Frazho, & Caverlee, 2009; Zhou et al., 2016) rank papers and authors based on the PageRank algorithm.
7. **TAPRank (TAP)**. Based on the scores of papers obtained by time-weighted PageRank, TAPRank calculates the scores of authors by using the author-paper network (Kong et al., 2015).
8. **Tri-Rank (TR)**. Like MutualRank, Tri-Rank also ranks papers, authors and venues simultaneously in heterogeneous networks (Liu et al., 2014).
9. **WMR-Rank (WMR)**. The proposed method in this paper (see Algorithm 1).

Note that the network presented in Section 3 only works for the proposed method WMR-Rank. For all the baseline methods involved, separate networks are set up for them with proper parameter definitions. For paper and author performance prediction and evaluation, we follow the exact procedures and settings in their original papers. However, some changes are required for venue performance prediction. WMR-Rank makes prediction of all the venues on a yearly basis, but those methods do not. To let them comparable with WMR-Rank, we treat each venue per year as a separate entity in their networks. See Appendix for details of all baseline methods involved except Citation Count and MIF.

4.4 Prediction accuracy of the proposed method

In this section, we present the evaluation results of the proposed algorithm, along with a group of state-of-the-art baseline methods.

4.4.1 Paper Influence Prediction. We first study the paper prediction performance of the proposed algorithm. For those papers in the modelling partition, we calculate scores for all the papers involved by a given prediction method. Then we find their incremental citation counts by using the evaluation partition and generate a ranked list of those papers accordingly. These two rankings are compared to see how accurate the prediction is. The time span of the evaluation partition is 10 years: from 2005 to 2014. We consider the incremental citation counts for 1 (2005), 2 (2005 and 2006),…, to up to 10 years (2005-2014). Fig. 2 and Fig. 3 show the prediction performance measured by the Spearman's ranking coefficient and Pearson correlation coefficient, respectively.

From Fig. 2 and Fig. 3, one can see that the proposed method WMR-Rank (WMR) constantly outperforms all other methods involved remarkably. FutureRank is in the second place when the Pearson correlation coefficient is used; and it is in the third place when Spearman's rank correlation coefficient is used. MRCoRank is in the second place when measured by Spearman's coefficient, but it becomes the worst when measured by the Pearson correlation coefficient. Three other methods MutualRank, Tri-Rank, and PageRank do not perform well.

In Figure 2, it is quite surprising that all the curves are slightly increasing with future years. The major reason is that the distribution of citation counts is extremely skewed. In future year one, less than 1% of the papers are cited 10 times or more, less than 6% of the papers are cited between three and nine times, over 15% of the papers are cited once or twice, while over 78% of the papers are not cited at all. Even after 10 years, over 52% of the papers are not cited, while three papers are cited over 700 times. Such a phenomenon can explain why the figure of Spearman's rank coefficient shows some unusual tendency due to too many ties in the ranking, while the figure of the Pearson correlation coefficient looks more reasonable. This may also explain why all seven methods except MRCR are better when the Pearson correlation coefficient is used for evaluation.

Next, we check how many of the top-10 cited papers between 2005 and 2009 are predicted by all the methods involved. For each of these methods, we use one column to present its prediction results. The top-10 predicted papers are presented in Table 1 with their ranking positions of citation counts between 2005 and 2009 inclusive. From Table 1 one can see that WMR-Rank is the best performer. It correctly predicts seven of top 10 papers. FutureRank, MutualRank, Tri-Rank, and PageRank successfully predict three of them; while MRCoRank fails to predict any top-10 papers.

4.4.2 Author influence prediction. In this piece of work, we predict the future influence of all the authors on a yearly basis. This is different from all previous work. In order to enable WMR-Rank comparisons against the baseline methods, we sum up all the yearly scores to obtain the total score for each author, which is then used to predict future performance. Rather than using the raw citation count only, we consider weighted citation count to be likely a more reasonable metric for the evaluation of authors (see Equation 16 in Section 4.2 for its definition). Therefore, both citation count and weighted citation count are used as metrics for the prediction of author influence. Fig. 4 shows the prediction accuracy of all the methods involved.

From Fig. 4 one can see that WMR-Rank is significantly better than all five baseline methods involved both when the Pearson correlation coefficient and when Spearman's rank correlation coefficient are used. Table 2, showing the predictions generated by all the methods for the 10 most influential authors, again sees WMR-Rank outperforming the other candidates. WMR-Rank successfully identifies half of the top 10 authors, CC identifies three, TR identifies two, both TAP and MR identify one, while MRCR identifies none of them.

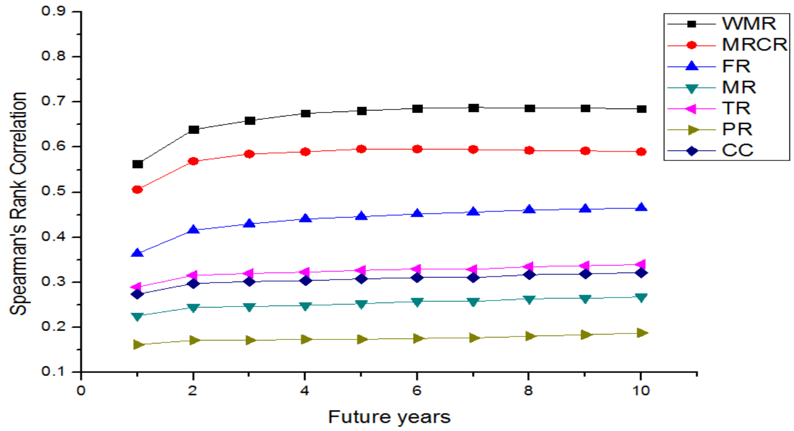


Fig. 2 Accuracy of seven paper performance prediction methods (measured by Spearman's rank correlation coefficient and prediction for 1, 2,…, 10 years ahead)

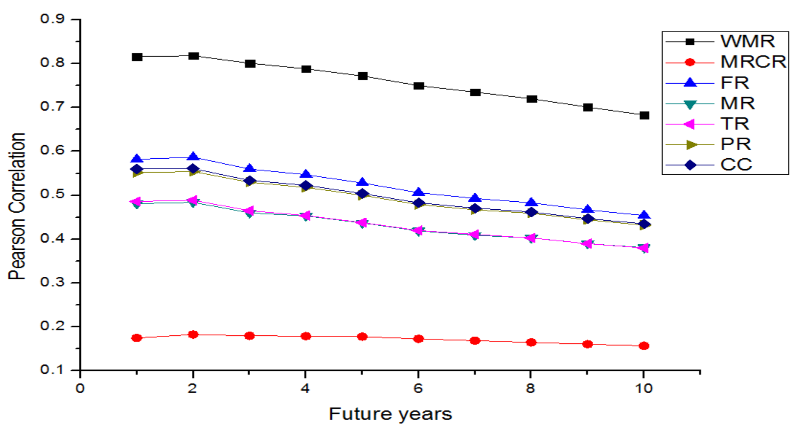
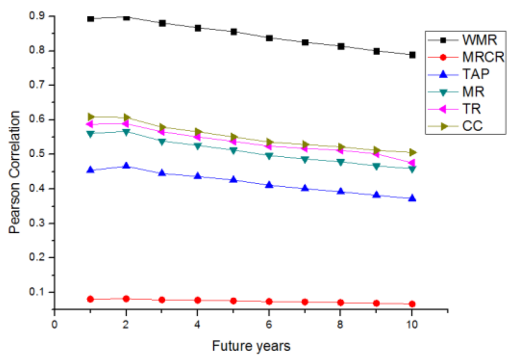


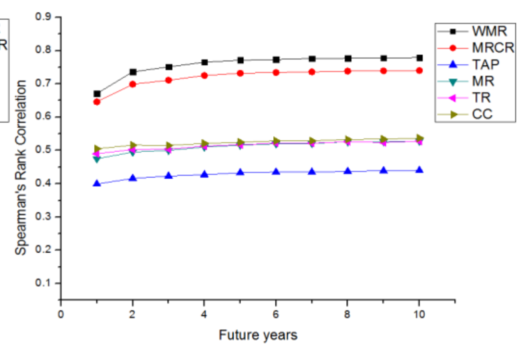
Fig. 3 Accuracy of seven paper performance prediction methods (measured by the Pearson correlation coefficient and prediction for 1, 2,…, 10 years ahead)

Table 1. Top 10 papers and their predictability by WMR-Rank and other baseline methods; for a given method, the number shows the ranking position of the paper by its citation count between 2005 and 2009

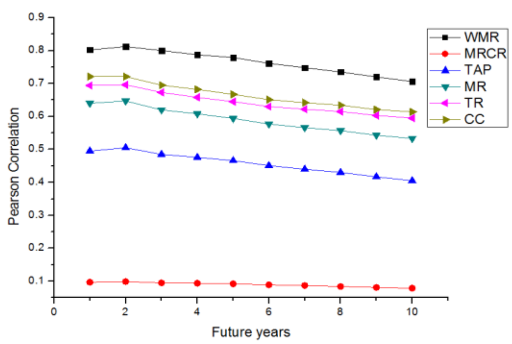
Rank	AAN ID	WMR	MRCR	MR	TR	FR	PR
1	J93-2004	**1**	216	**2**	**1**	**1**	**1**
2	P02-1040	**7**	214	49	61	55	57
3	P03-1021	13	303	753	384	324	384
4	J03-1002	**8**	215	211	162	104	163
5	N03-1017	16	381	643	357	215	348
6	J93-2003	**3**	211	**4**	**3**	**3**	**4**
7	A00-2018	**2**	217	24	20	19	20
8	J96-1002	**6**	224	**8**	**5**	**5**	**7**
9	J02-3001	**5**	222	31	37	34	36
10	P98-2127	44	221	63	44	43	42



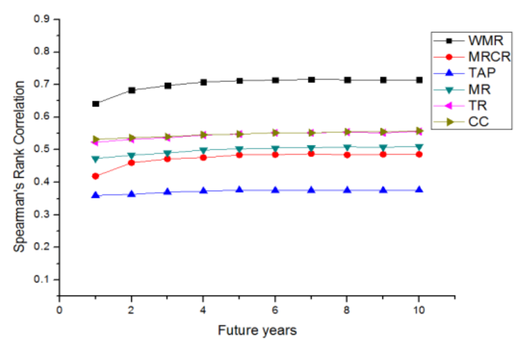
(a) Pearson correlation coefficient on WCC



(b) Spearman's rank coefficient on WCC



(c) Pearson correlation coefficient on CC

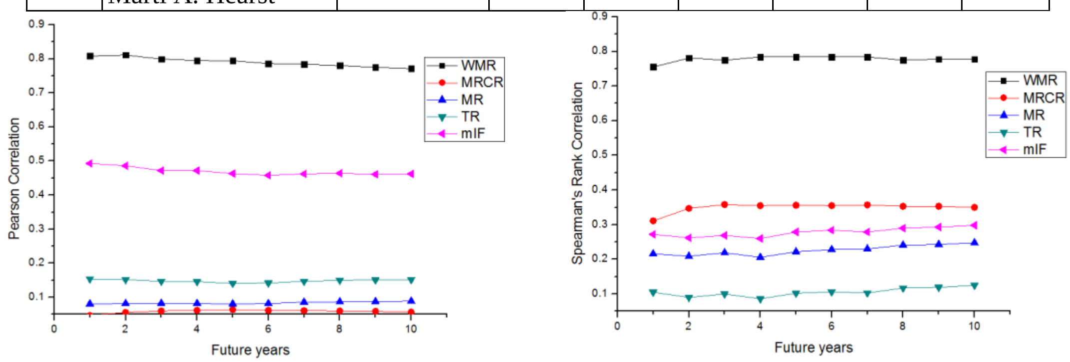


(d) Spearman's rank coefficient on CC

Fig. 4. Accuracy of six author performance prediction methods for 1, 2,…, 10 years ahead

Table 2. Top-10 authors and their predictability by WMR and other baseline methods; for a given method, the number shows the ranking position of the author by his/her weighted citation count between 2005 and 2009

Rank	Author’s name	Number of publications	WMR	MRCR	TAP	MR	TR	CC
1	Franz Josef Och	23	**2**	112	60	44	19	22
2	Michael Collins	15	**1**	381	107	20	11	**10**
3	Hermann Ney	37	**7**	39	**9**	21	14	18
4	Eugene Charniak	21	**4**	1031	27	30	16	27
5	Dekang Lin	20	19	172	22	58	85	74
6	Philipp Koehn	6	70	1301	671	514	356	454
7	Mitchell P. Marcus	13	24	2301	146	11	**8**	**6**
8	Daniel Gildea	13	**3**	85	154	99	118	53
9	Peter F. Brown	11	17	2138	495	**6**	**3**	**3**
10	Marti A. Hearst	13	32	1093	117	57	70	56



(a) Pearson correlation coefficient (b) Spearman’s rank coefficient

Fig. 5. Accuracy of five venue performance prediction methods for 1, 2,…, 10 years ahead

Table 3. Top-10 venues by WMR-Rank and other baseline methods (MIF for five years)

Rank	Venue	WMR	MRCR	MR	TR	MIF
1	CL, 1993	**2**	84	**8**	**1**	**1**
2	CL, 2003	15	36	111	141	73
3	HLT@NAACL, 2003	**3**	34	76	107	63
4	ACL, 2002	**1**	22	42	58	34
5	CL, 2004	41	**9**	143	183	176
6	HLT@NAACL, 2004	**6**	**6**	93	170	146
7	ACL, 2003	**7**	18	55	109	85
8	CL, 2002	**4**	51	86	93	41
9	CL, 1996	**10**	90	48	38	**4**
10	CL, 1998	**9**	76	57	49	**10**

Table 4. Comparison of seven feature-based variants with WMR-Rank for paper prediction (Pearson's coefficient on CC)

Future years	WMR	TW	TP	PW	T	W	P	NN
1	**0.815**	0.783	0.772	0.508	0.765	0.493	0.330	0.207
2	**0.817**	0.781	0.775	0.510	0.758	0.494	0.333	0.210
3	**0.801**	0.760	0.766	0.486	0.741	0.470	0.315	0.198
4	**0.788**	0.746	0.755	0.474	0.727	0.459	0.307	0.193
5	**0.772**	0.728	0.745	0.457	0.712	0.442	0.295	0.185
6	**0.750**	0.705	0.729	0.437	0.690	0.423	0.281	0.175
7	**0.735**	0.689	0.716	0.427	0.674	0.413	0.275	0.172
8	**0.721**	0.675	0.703	0.719	0.660	0.405	0.271	0.171
9	**0.701**	0.655	0.686	0.405	0.641	0.392	0.262	0.165
10	**0.683**	0.638	0.669	0.394	0.624	0.381	0.256	0.162

Table 5. Comparison of seven feature-based variants with WMR-Rank for author evaluation (Pearson's coefficient on MCC)

Future years	WMR	TW	TP	PW	T	W	P	NN
1	**0.894**	0.877	0.799	0.641	0.799	0.634	0.449	0.354
2	**0.898**	0.881	0.800	0.642	0.800	0.635	0.446	0.351
3	**0.881**	0.861	0.790	0.611	0.784	0.603	0.420	0.329
4	**0.867**	0.847	0.779	0.594	0.771	0.587	0.408	0.319
5	**0.856**	0.835	0.769	0.581	0.759	0.573	0.395	0.307
6	**0.838**	0.816	0.757	0.563	0.745	0.555	0.383	0.297
7	**0.825**	0.802	0.748	0.553	0.735	0.545	0.378	0.293
8	**0.814**	0.790	0.738	0.544	0.725	0.537	0.374	0.290
9	**0.800**	0.776	0.727	0.532	0.712	0.524	0.366	0.284
10	**0.789**	0.765	0.718	0.525	0.703	0.517	0.362	0.281

Table 6. Comparison of seven feature-based variants with WMR-Rank for venue evaluation (Pearson's coefficient on MIF)

Future years	WMR	TW	TP	PW	T	W	P	NN
1	**0.826**	0.769	0.647	0.158	0.760	0.094	0.060	-0.054
2	**0.830**	0.774	0.657	0.157	0.773	0.094	0.058	-0.052
3	**0.818**	0.762	0.657	0.154	0.766	0.092	0.053	-0.052
4	**0.815**	0.760	0.652	0.152	0.768	0.090	0.055	-0.051
5	**0.813**	0.759	0.657	0.148	0.768	0.088	0.053	-0.050
6	**0.804**	0.749	0.647	0.149	0.758	0.089	0.053	-0.048
7	**0.801**	0.746	0.639	0.154	0.750	0.094	0.057	-0.043
8	**0.797**	0.742	0.633	0.157	0.745	0.097	0.061	-0.038
9	**0.792**	0.736	0.626	0.158	0.738	0.099	0.063	-0.036
10	**0.788**	0.732	0.617	0.160	0.731	0.100	0.066	-0.032

4.4.3 Venue influence prediction. Venue influence is evaluated by using MIF (Modified Impact Factor, see Equation 17 in Section 4.2 for its definition). Fig. 5 shows the predicted influence of venues for one to ten years ahead. WMR-Rank performs much better than all the others, both in terms of the Pearson correlation coefficient and Spearman's rank correlation coefficient. Table 3 shows the results of the predictions made by the methods for the 10 most influential venues. Eight of them are predicted successfully by WMR-Rank, and one to three venues are predicted successfully by four other methods.

4.4.4 Evaluation of the impact of three components in WMR-Rank. In WMR-Rank, time-aware information, past performance of authors and venues and weighting are three major components. Two of them, time-aware information and past performance, are implemented by defining entities of authors and venues on a yearly basis. In order to investigate how these components contribute to the final success of the algorithm, we compare WMR-Rank with some of its variants in which one or more components are disabled. Seven variants of WMR-Rank are defined as follows. P denotes past performance, T denotes time-aware information, while W denotes the weighting schema for papers, authors, and venues. For example, WMR-T means that time-aware information is supported while the two others are not, WMR-PW means that both past time and weighing are supported while time-aware information is not. WMR-NN has all three components cancelled.

1. **WMR-PW**. A variant of WMR-Rank that does not consider time-aware information.
2. **WMR-TP**. A variant of WMR-Rank that sets all weights equally.
3. **WMR-TW**. A variant of WMR-Rank that does not consider the past performance of authors and venues.
4. **WMR-T**. A variant of WMR-Rank that does not consider the past performance of authors and venues, and meanwhile sets all weights equally.
5. **WMR-W**. A variant of WMR-Rank that does not consider the past performance of authors and venues, as well as the time-aware information.
6. **WMR-P**. A variant of WMR-Rank that does not consider the time-aware information, and meanwhile sets all weights equally.
7. **WMR-NN**. A variant of WMR-Rank. It sets all weights equally, and considers neither time-aware information nor the past performance of authors and venues.

The results are shown in Tables 4, 5, and 6 for papers, authors, and venues, respectively. Pearson's coefficients are used as the metric for comparison. From these tables, one can see that unsurprisingly, WMR-Rank performs better than all of its seven variants in prediction of papers, authors and venues. In all the cases, the three components are effective for performance prediction, either used separately or in combination. The more components are involved, the better prediction performance we are able to achieve. However, the capabilities of these three components to affect the final prediction are different. Among them, time-awareness has the largest impact, weighting is in the middle, while the past performance of authors and venues has the smallest impact. Besides, the impact of time-awareness is far larger than that of the two others.

4.4.5 Alternative paper citation weights for WMR-Rank. Some studies claim that citations are not equally important (Chakraborty & Narayanam, 2016; Wan & Liu, 2014; Wang et al., 2013; Zhu et al., 2015). To further study the effect of different weighting schemes of paper citations $W_{PP}(p_i, p_j)$,, we test three more methods: WMR-TI, which considers time interval between publication and citation (Yan & Ding, 2010; Yu et al., 2017; Zhang & Wu, 2018); WMR-CC, which considers similarity between the citing paper and the cited paper (Amsler, 1972); and WMR-CA, which considers the age of the citing paper (Wang et al., 2016).

To study the effect of different weighting schemes of paper citation $W_{PP}(p_i, p_j)$, we compare the one that WMR-Rank uses (see Equation 1) with five other options: WMR-TI (TI for Time Interval), WMR-CC (CC stands for Coupling and Citation), WMR-Cit (Cit stands for Citation), WMR-Cou (Cou stands for Coupling), and WMR-CA (CA stands for Citing paper's Age). See below for more details about them.

1. **WMR-TI**. A variant of WMR-Rank that considers the time interval between publication and citation. It is considered that a paper is likely to be influential if the paper is cited shortly after its publication. Therefore, for each citation, we use the time interval between the cited paper and the citing paper to decide the weight of the citation. Moreover, self-citations are not accorded as much value as citation by others. Let p_i be a paper and $T_{published}(p_i)$ be the year in which p_i is published, and $F_A(p_i)$ be the first author of p_i, then the weight of the citation $W_{PP}(p_i, p_j)$ is defined as:

$$W_{PP}(p_i, p_j) = \begin{cases} e^{-\tau(T_{published}(p_i) - T_{published}(p_j))}, & if \;\; F_A(p_i) \neq F_A(p_j) \\ 0.5 * e^{-\tau(T_{published}(p_i) - T_{published}(p_j))}, & if \;\; F_A(p_i) = F_A(p_j) \end{cases} \tag{19}$$

2. **WMR-CC**. A variant of WMR-Rank that considers bibliographic coupling and co-citation. It is considered that the two papers are more similar if they have more bibliographic couplings or co-citations. Greater weights are given to papers that are more similar. Let $I(p_i)$ be the set of in-link neighbors of p_i, $O(p_i)$ be the set of out-link neighbors of p_i, the weights of co-citation and coupling are balanced by a parameter (we set it to 0.5, as in (Amsler, 1972)), then the weight we assign to p_i is

$$W_{PP}(p_i, p_j) = 0.5 \times \frac{|I(p_i) \cap I(p_i)|}{|I(p_i) \cup I(p_i)|} + 0.5 \times \frac{|O(p_i) \cap O(p_i)|}{|O(p_i) \cup O(p_i)|} \tag{20}$$

3. **WMR-Cit**. A variant of WMR-Rank that only considers bibliographic co-citation. It is considered that the two papers are more similar if they have more bibliographic co-citation. Let $I(p_i)$ be the set of in-link neighbors of p_i, the weight we assign to p_i is

$$W_{PP}(p_i, p_j) = \frac{|I(p_i) \cap I(p_i)|}{|I(p_i) \cup I(p_i)|} \tag{21}$$

4. **WMR-Cou**. A variant of WMR-Rank that only considers bibliographic coupling. It is considered that the two papers are more similar if they have more bibliographic couplings. Let $O(p_i)$ be the set of out-link neighbors of p_i, the weight we assign to p_i is

$$W_{PP}(p_i, p_j) = \frac{|O(p_i) \cap O(p_i)|}{|O(p_i) \cup O(p_i)|} \tag{22}$$

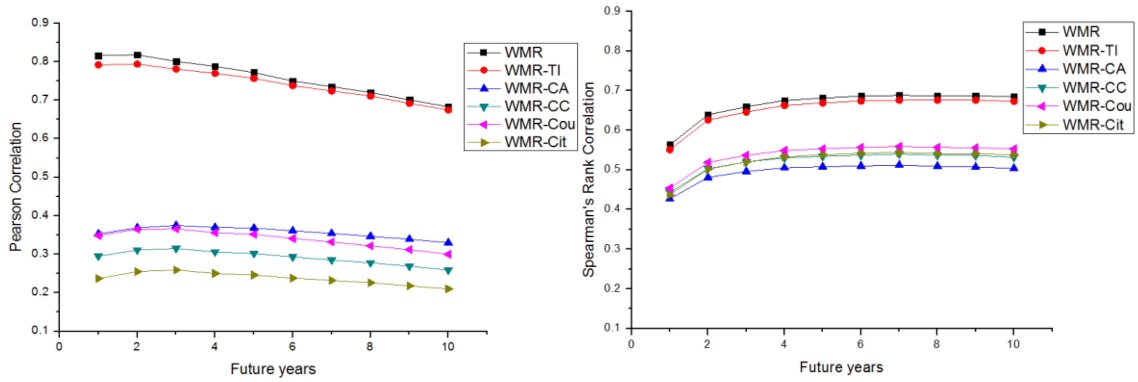
5. **WMR-CA**. It is a variant of WMR-Rank that considers the age of the citing paper. It is considered that papers obtain citation recently are more likely to attract new citations. Let $T_{evaluate}$ be the evaluation year, $T_{published}(p_i)$ be the year of citing paper p_i, then the weight of citation is set as

$$W_{PP}(p_i, p_j) = e^{-\delta(T_{evaluate} - T_{published}(p_i))} \tag{23}$$

   where δ is set to 2 (Wang et al., 2016).

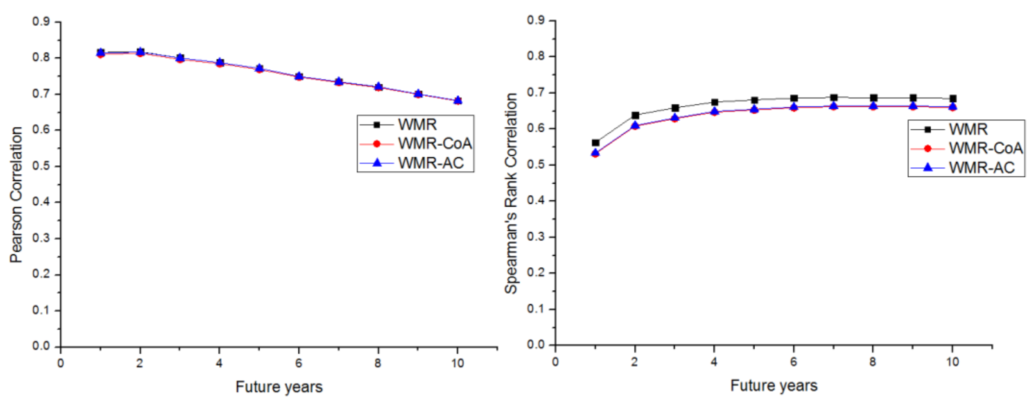
Experimental results are shown in Fig. 6. One can see that WMR-Rank outperforms the others. WMR-TI is close to WMR-Rank in performance, but the other four are much worse. This demonstrates that using simple binary weights for paper citation is a very good option for WMR-Rank.

4.4.6 Impact of author citation and co-authorship on WMR-Rank. In WMR-Rank, a researcher may obtain credit from another researcher through either paper citation or co-authorship. Their impact on paper or author ranking is merged. To have a clear view of the impact of co-authorship and author citation separately on WMR-Rank, we introduce two more variants of WMR-Rank. WMR-CoA is the one that considers co-authorship but not author citation, or $W_{AA}(a_i, a_j) = W_{CoA}(a_i, a_j)$; and WMR-AC is the one that considers author citation but not co-authorship, or $W_{AA}(a_i, a_j) = W_{CA}(a_i, a_j)$. Fig. 7 shows the performance of them for paper and author performance prediction. From Fig. 7 (a) and (b), one can see that both WMR-CoA and WMR-AC are very close to WMR-Rank for paper performance prediction. However, their impacts on author performance prediction are different. From Fig. 7 (c) and (d), one can see that WMR-AC and WMR-Rank are close in performance, while WMR-CoA is not as good as WMR-AC. Therefore, we may conclude that considering either co-authorship or author citation or both do not affect much on paper performance prediction, but author citation is a more important aspect than co-authorship for author performance prediction.

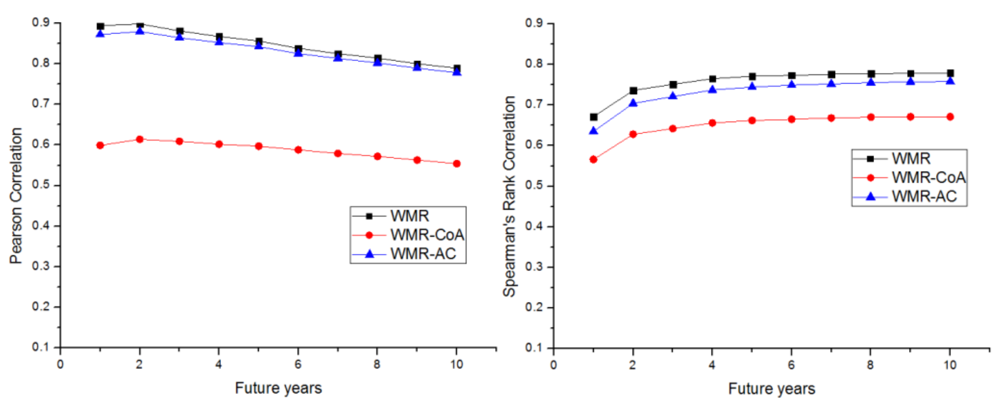


(a) Pearson correlation coefficient (b) Spearman's rank coefficient

Fig. 6 Paper prediction performance of five variants of WMR-Rank with different weighting schemas of W_{PP}



(b) Pearson correlation coefficient (paper) (b) Spearman's rank coefficient (paper)



(c) Pearson correlation coefficient (author) (d) Spearman's rank coefficient (author)

Fig. 7 Comparison of WMR-Rank with its variants WMR-CoA and WMR-AC

4.5 Parameter setting and convergence rate of WMR-Rank

There are nine parameters in WMR-Rank: α, β, γ, σ_1, σ_2, σ_3, λ, ε and μ, Among these parameters, the function of λ is similar to the damping factor in PageRank, and a typical value of 0.85 is used. Another parameter μ is used to increase the scores of papers published in recent years. A step function is a good option for this purpose. In this study, we let μ =0.80 for papers published in year 2003 or 2004; μ =0.60, 0.40, and 0.20, for papers published in year 2002, 2001, and 2000, respectively; μ =0 for papers published before 2000. Note that the modelling partition contains all the papers published between 1985 and 2004.

Now let us look at six remaining parameters. Based on their definitions and roles, α, β, and γ are in the range of 0 and 1, while σ_1, σ_2 and σ_3 should take positive numbers. To set up these parameters so as to obtain good performance, we initialize them to some reasonable values: let α=0.50, β=0.50, γ=0.50, λ=0.85 as PageRank, $\varepsilon = 1\mathrm{e}-6$ as in (Jiang et al., 2016), σ_1=1.0, σ_2=1.0, and σ_3=1.0. Then we let one of them vary to see its effect, as Fig. 8 shows the results when Pearson correlation coefficient is used for performance evaluation.

From Fig. 8 (a), one can see that paper prediction performance is quite stable when α is in the range of 0.10 and 0.90. The best performance is achieved when α=0.70. In a similar way, we get β=0.60, γ=0.10, σ_1=0.50, σ_2=1.00, and σ_3=1.10 by observing Fig. 8 (b), (c), and (d). The Spearman's rank coefficient is also evaluated for the same purpose. We have α=0.30, β=0.50, γ=0.10, σ_1=0.30, σ_2=1.00, and σ_3=1.00. It is understandable that these two weighting schemes are not the same. Both of them are reasonably good weighting schemes for WMR-Rank, but none of them is highly optimized because the hyper-parameter tuning we carried out for them is modest.
Line 10 of Algorithm 1 calculates the score of a given paper, taking into account two factors: the authors who write it and the venue in which the paper is published. α and (1-α) are used to adjust the relative weights of these two components. A larger α value does not necessarily mean that one factor is more important than the other one because these two components are not directly comparable — α partially serves as a normalization measure. We may have the same conclusion for other parameters β, γ, σ_1, σ_2, and σ_3.

To test the convergence rate of our algorithm WMR-Rank, we vary the threshold ε to control the number of iterations. For convenience, the difference between two consecutive iterations (see line 25 in Algorithm 1) is rewritten here

$$\Delta = \sum_{p_i \in V_P} |PS(p_i)^{t+1} - PS(p_i)^t| + \sum_{a_i \in V_A} |AS(a_i)^{t+1} - AS(a_i)^t| + \sum_{v_i \in V_V} |VS(v_i)^{t+1} - VS(v_i)^t| \quad (24)$$

If $\Delta \leq \varepsilon$, then the algorithm stops.

We compare the convergence rate of WMR-Rank with that of three other algorithms MutualRank, Tri-Rank, and MRCoRank, all of which rank papers, authors and venues simultaneously.

Fig. 9 shows the convergence rates of these four algorithms. We may observe that Tri-Rank is the fastest to converge, WMR-Rank is in the second place, while MutualRank and MRCoRank converge more slowly. WMR-Rank achieved convergence after 12 iterations, while Tri-Rank, MutualRank, and MRCoRank converged after 10, 43, and 52 iterations, respectively.

4.6 Some statistics of the data set and their implications

There are nine parameters in WMR-Rank: α, β, γ, σ_1, σ_2, σ_3, λ, ε and μ, Among these parameters, the function of λ is similar to the damping factor in PageRank, and a typical value of 0.85 is used. Another parameter μ is used to increase the scores of papers published in recent years. A step function is a good option for this purpose. In this study, we let μ =0.80 for papers published in year 2003 or 2004; μ =0.60, 0.40, and 0.20, for papers published in year 2002, 2001, and 2000, respectively; μ =0 for papers published before 2000. Note that the modelling partition contains all the papers published between 1985 and 2004.

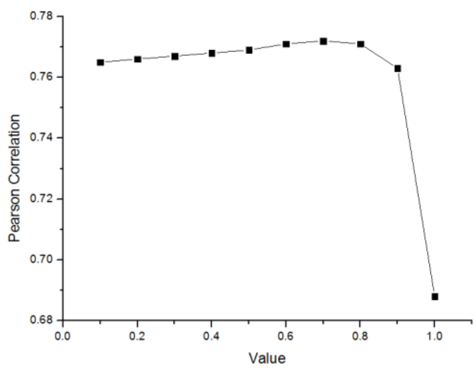
After pro-processing, the ACL data set includes 19,891 papers, 15,379 authors, and 372 venues. All the papers are cited 124,710 times and on average each is cited 6.27 times. The top 5% of authors (based on the number of publications) produced 12,548 papers in total, which is more than 60% of the 19,891 papers in the whole data set. That is to say, the top one-twentieth of researchers are involved in 60% of papers. If we consider the top 10% of researchers, they contribute to nearly three quarters of all publications. In all, the authors attracted 869,624 citations. Of these, the top 10% authors attracted

438,429 citations in total, which is almost half of all the citations. Not surprisingly, there is a considerable difference between the very top authors and the average author if we consider the number of publications. The difference is even greater if we consider the number of citations they attract.

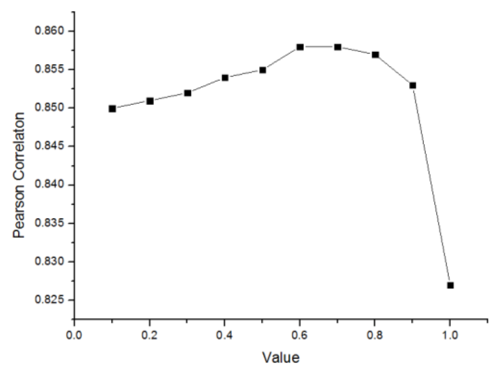
For all 6767 papers and 5310 authors involved in the modelling partition, we measure their future influence in the evaluation partition. On average, each paper is cited 5.54 times and each author gets 7.06 in WCC. Fig. 10 shows the non-increasing influence curves of papers/authors from most influential to least. We may observe that both curves are extremely skewed.

Over 80% of the papers are cited 5 times or less, while less than 20% of papers are cited 6 times or more. Only 671 authors obtain a WCC of 7.06 or above, while the remaining 4639 authors do not reach the average. This suggests that only a small percentage of top papers or authors need to be considered when we try to make predictions for future influence. If we rank papers by the number of citations they obtain, then many of the less influential papers will be cited very few times, resulting in a lot of ties. Under such circumstances, Spearman's rank correlation coefficient or Kendall's τ rank coefficient may not be very good measures for us to compare the similarity of two rankings; while Pearson linear correlation coefficient is likely to be more reliable.

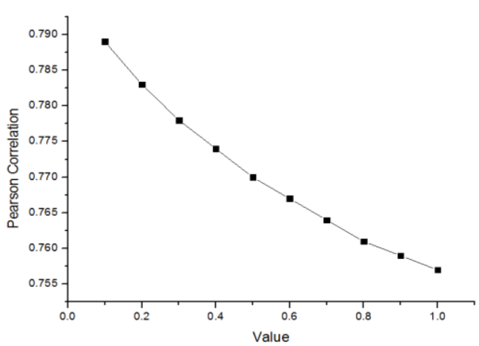
On the other hand, the performance of each researcher or venue may very over time. Fig. 11 shows the performance of five authors and four venues over time. From Fig. 11(a) and (b) we may observe that the performance of each author/venue may vary considerably from one year to next. Therefore, it is adequate to treat each author and venue per year as a separate entity. For a given author, her recent performance, rather than her average performance over long time, is a more accurate estimator of her future influence. Time-aware weights are another measure for us to focus on most recent citations.



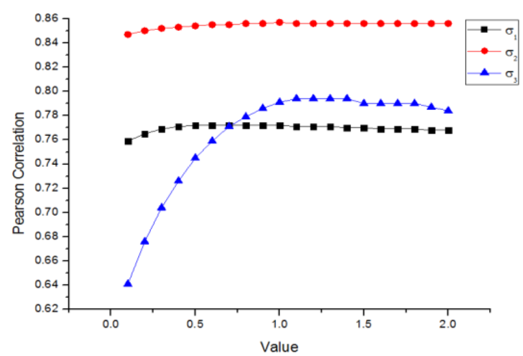
(a) Effect of α on papers



(b) Effect of β on authors



(c) Effect of γ on venues



(d) Effect of σ_1, σ_2 and σ_3 on papers, authors and venues, respectively

Fig. 8 Effect of different parameter values on ranking performance

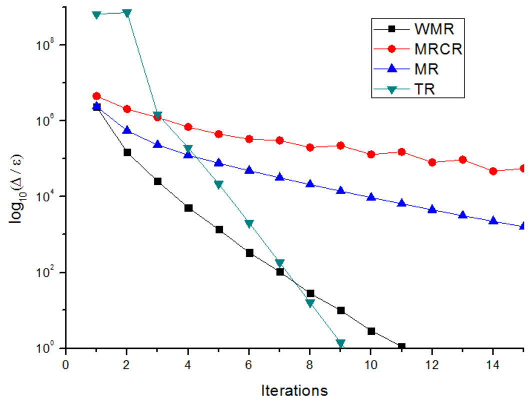
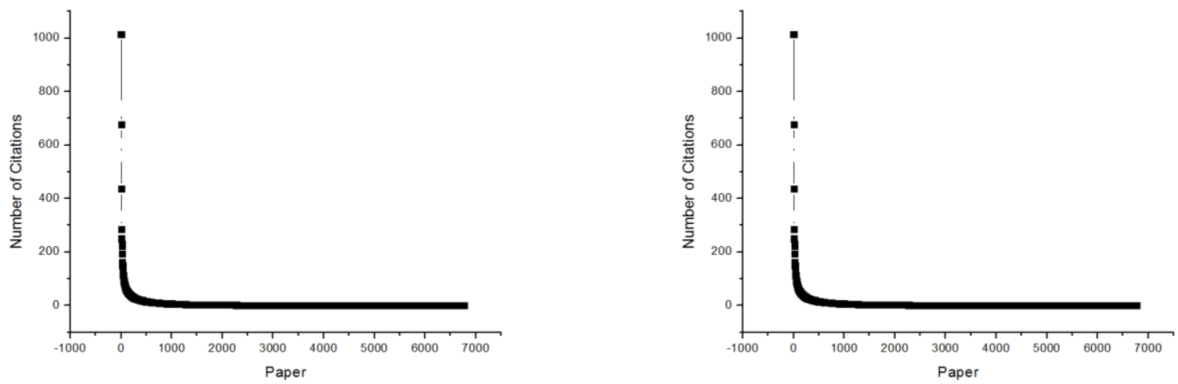


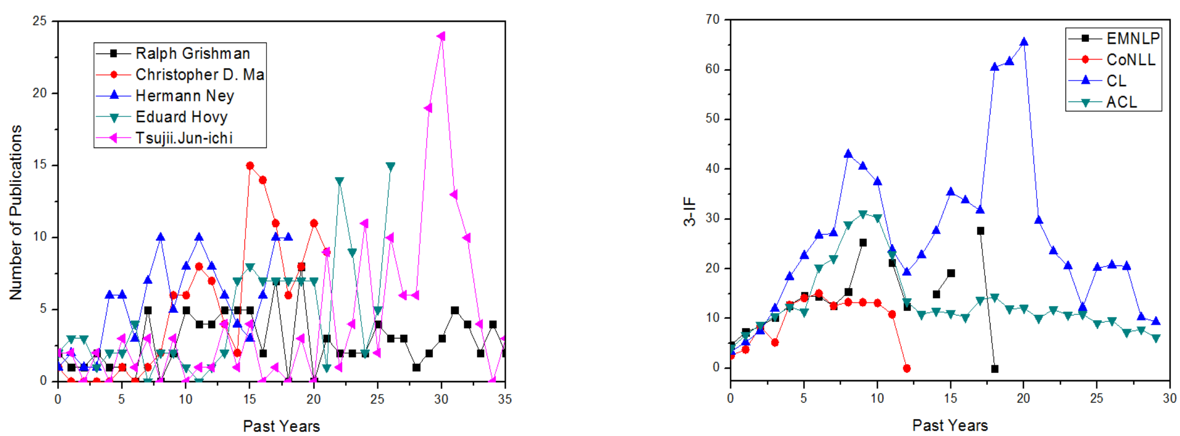
Fig. 9 Convergence Rate of four Algorithms



(a) Future influence of papers

(b) Future influence of authors

Fig. 10 Future influence of all papers/authors in the modelling partition in the next 10 years



(a) Variation of five authors' performance over 35 years

(b) Variation of four venues' performance over 30 year

Fig. 11 Performance variation of authors and venues over a number of years

5. Conclusions

Predicting the future influence of papers, authors and venues is an essential task for scientists, research organizations, research funding bodies, and government agencies alike. In this paper, we have proposed a ranking algorithm, WMR-Rank, which is able to predict the future influence of papers, authors and venues at the same time. Several measures have been introduced to make the prediction more accurate. They include refined granularity of concerned entities such as papers and venues, time awareness for published papers and their citations, differentiating the contribution of multiple coauthors to the same paper, and others. Experiments have been conducted on a real world dataset ACL and results show that our approach is much more effective than state-of-the-art approaches.

As our future work, we would consider paper citation in depth. Traditional citation-based analysis treats all citations equally, as we do in this paper. However, due to various reasons, some citations are deep, while some others are superficial. Recently, it has been found that misconduct such as coercive citations and padded citations also exists in many cases (Fong & Wilhite, 2017). One possible approach to tackle these problems is to divide citations into different levels based on the relevancy between the cited paper and the citing paper. Then more sophisticated metrics may be defined for the evaluation of entities such as papers, authors, and venues. A closely related issue is how to decide the relevancy of any two papers effectively and efficiently. Manual categorization is too costly and information extraction and deep learning techniques may be helpful to automate this process.

References

Abbas, A. M. (2011). Weighted indices for evaluating the quality of research with multiple authorship. *Scientometrics*, 88(1):107–131.

Abrishami, A., & Aliakbary, S. (2019) Predicting citation counts based on deep neural network learning techniques. *Journal of Informetrics*, 13(2): 485-499.

Amsler, R. (1972). Application of citation-based automatic classification. Technical report, University of Texas at Austin, Austin.

Assimakis, N. (2010). A new author's productivity index: p-index. *Scientometrics*, 85(2), 415–427.

Bergstrom, C. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *College and Research Libraries News*, 68(5), 314–316.

Bertsimas, D., Brynjolfsson, E., Reichman, S., & Silberholz, J. (2013). Network analysis for predicting academic impact. *International conference on information systems*, 1-11.

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.

Cao, X., Chen, Y., & Liu, K. J. R. (2016). A data analytic approach to quantifying scientific impact. *Journal of Informetrics*, 10(2): 471-484

Casal, G. B. (2004). Assessing the quality of articles and scientific journals: proposal for weighted impact factor and a quality index. *Psychology in Spain*, 8(1), 60–76.

Chakraborty, T., & Narayanam, R. (2016). All Fingers are not equal: intensity of references in scientific articles. *Conference on Empirical Methods in Natural Language Processing*, 1348-1358.

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.

Du, J., & Tang, X. (2013). Potential of harmonic counts for encouraging ethical co-authorship practices. *Scientometrics*, 96(1), 277–295.

Dunaiski, M., Visser, W., & Geldenhuys, J. (2016). Evaluating paper and author ranking algorithms using impact and contribution awards. *Journal of Informetrics*, 10(2), 392–407.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131-152.

Egghe, L., Rousseau, R., & Hooydonk, G. V. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science and Technology*, 51(2), 145–157.

Farooq, M., Khan, H. U., Iqbal, S., Munir, E. U., & Shahzad, A. (2017). DS-index: Ranking authors distinctively in an Academic network. *IEEE Access*, 5, 19588-19596.

Franceschin, F., Galetto, M., Maisano, D., & Mastrogiacomo, L. (2012). The success-index for evaluating an individual's research output. *Scientometrics*, 92(3), 621-641.

Fong, E. A., & Wilhite, A. W. (2017). Authorship and citation manipulation in academic research. *PLOS ONE*, 12(12), e0187394.

Fujimagari, H., & Fujita, K. (2014). Detecting research fronts using neural network model for weighted citation network analysis. *Journal of Information Processing*. 23(6), 753–758.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060), 471-479.

González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379–391.

González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals scientific prestige: The SJR2 indicator. *Journal of Informetrics*, 6(4), 674–688.

Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLOS ONE*, 3(12), e4021.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science of the United States of America*, 102(46), 16569–16572.

Hodge, S. E., Greenberg, D. A., & Challice, C. E. (1981). Publication credit. *Science*, 213(4511), 950–952.

Hooydonk, G. V. (1997). Fractional counting of multiauthored publications: Consequences for the impact of authors. *Journal of the American Society for Information Science and Technology*, 48(10), 944–945.

Jiang, X., Sun, X., Yang, Z., Zhuge, H., & Yao, J. (2016). Exploiting heterogeneous scientific literature networks to combat ranking bias: Evidence from the computational linguistics area. *Journal of the Association for Information Science and Technology*, 67(7), 1679–1702.

Jin, B., Liang, L., Rousseau, R., & Leo, E. (2007). The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin*, 52(6), 855-863.

Katerattanakul, P., Han, B., & Hong, S. (2015). Objective quality ranking of computing journals. *Communications of the ACM*, 46(10), 111-114.

Kim, J., & Diesner, J. (2014). A network-based approach to coauthorship credit allocation. *Scientometrics*, 101(1), 587–602.

Kong, X., Zhou, J., Zhang, J., Wang, W., & Xia, F. (2015). TAPrank: A time-aware author ranking method in heterogeneous networks. *IEEE International Conference on Smart City*/*SocialCom*/*SustainCom* (*SmartCity*), 242–246.

Lee, D. H. (2019). Predicting the research performance of early career scientists. *Scientometrics*, 121(3), 1481-1504.

Li, X., Liu, B., & Yu, P. (2008). Time sensitive ranking with application to publication search. *International Conference on Data Mining*, 893–898.

Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. *Social Studies of Science*, 10(2), 145-162.

Liu, Z., Huang, H., Wei, X., & Mao, X. (2014). Tri-Rank: An authority ranking framework in heterogeneous academic networks by mutual reinforce. *IEEE International conference on Tools with Artificial Intelligence*, 493-500.

Martin, T., Ball, B., Karrer, B., & Newman, M. E. J. (2013). Coauthorship and citation in scientific publishing. *Physical Review E*, 91(1), 012825

Meng, Q., & Kennedy, P. J. (2013). Discovering influential authors in heterogeneous academic networks by a co-ranking method. *ACM International Conference on Information & Knowledge Management*, 1029–1036.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277.

Nerur, S., Sikora, R., Mangalaraj, G., & Balijepally, V. G. (2005). Assessing the relative influence of journals in a citation network. *Communications of the ACM*, 48(11), 71-74.

Pajić, D. (2015). On the stability of citation-based journal rankings. *Journal of Informetrics*, 9(4), 990-1006.

Perianes-Rodriguez, A., Waltman, L., & Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, 10(4), 1178–1195.

Qiao, H., Wang, Y., & Liang, Y. (2012). A value evaluation method for papers based on improved pagerank algorithm. *International Conference on Computer Science and Network Technology*, 1029–1036.

Radev, D. R., Muthukrishnan, P., Qazvinian, V., & Abu-Jbara, A. (2013). The ACL Anthology Network Corpus. *Language Resources and Evaluation*, 47(4), 919-944.

Sayyadi, H., & Getoor, L. (2009). FutureRank: Ranking scientific articles by predicting their future pagerank. *SIAM International Conference on Data Mining*, 533–544.

Stallings, J., Vance, E., Yang, J., Vannier, M., Liang, J., Pang, L., et al. (2013). Determining scientific impact using a collaboration index. *Proceedings of the National Academy of Science of the United States of America*, 110(24), 9680–9685.

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). ArnetMiner: extraction and mining of academic social networks. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 990-998.

Walker, D., Xie, H., Yan, K., & Maslov, S. (2006). Ranking scientific publications using a simple model of network traffic. *Journal of Statistical Mechanics: Theory and Experiment*, 6(6), P06010–P06015.

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391.

Waltman, L., Eck, N. J., Leeuwen, T. N., & Visser, M. S. (2013). Some modifications to the snip journal impact indicator. *Journal of Informetrics*, 7(2), 272–285.

Wan, X., & Liu, F. (2014). Are all literature citations equally important? Automatic citation strength estimation and its applications. *Journal of the American Society for Information Science and Technology*, 65(9), 1929–1938.

Wang, S., Xie, S., Zhang, X., Li, Z., Yu, P., & He, Y. (2016). Coranking the future influence of multiobjects in bibliographic network through mutual reinforcement. *ACM Transactions on Intelligent Systems and Technology*, 7(4), 64:1–64:28.

Wang, Y., Tong, Y., & Zeng, M. (2013). Ranking scientific articles by exploiting citations, authors, journals, and time information. *AAAI Conference on Artificial Intelligent*, 14–18.

Xu, F., Liu, & W., Rousseau, R. (2015). Introducing sub-impact factor (SIF-) sequences and an aggregated SIF-indicator for journal ranking. *Scientometrics*, 102(2), 1577–1593.

Yan, E., & Ding, Y. (2010). Weighted citation: An indicator of an article's prestige. *Journal of the American Society for Information Science and Technology*, 61(8), 1635–1643.

Yan, E., Ding, Y., & Sugimoto, C. R. (2011). P-Rank: An indicator measuring prestige in heterogeneous scholarly networks. *Journal of the American Society for Information Science and Technology*, 62(3), 467–477.

Yan, R., Tang, J., Liu, X., Shan, D., & Li, X. (2011) Citation count prediction: learning to estimate future citations for literature. *ACM International Conference on Information & Knowledge Management*, 1247-1252

Yan, S., & Lee, D. (2007). Toward alternative measures for ranking venues: A case of database research community. *ACM/IEEE on Joint Conference on Digital Libraries*, 235–244.

Yu, D., Wang, W., Zhang, S., Zhang, W., & Liu, R. (2017). A multiple-link, mutually reinforced journal-ranking model to measure the prestige of journals. *Scientometrics*, 111(1), 521–542.

Yu, T., Yu, G., Li, P., & Wang, L. (2014) Citation impact prediction for scientific papers using stepwise regression analysis. *Scientometrics*, 101(2): 1233-1252.

Zhang, F., & Wu, S. (2018). Ranking scientific papers and venues in heterogeneous academic networks by mutual reinforcement. *ACM/IEEE on Joint Conference on Digital Libraries*, 127-130.

Zhou, D., Orshanskiy, S. A., Zha, H., & Giles, C. L. (2007). Co-ranking authors and documents in a heterogeneous network. *IEEE International Conference on Data Mining*, 739–744.

Zhou, J., Zeng, A., Fan, Y., & Di, Z. (2016). Ranking scientific publications with similarity-preferential mechanism. *Scientometrics*, 106(2), 805–816.

Zhu, X., Turney, P., Lemire, D., & Vellino, A. (2015). Measuring academic influence: Not all citations are equal. *Information Science and Technology*, 66(2), 408–427.

Appendix

Details about six baseline methods involved in our experiment are given in this appendix. Their implementation is the same as in their original papers. Minor changes to some of the concepts keep the presentation of this paper consistent.

MRCoRank

Seven sub-networks are defined. They are paper citation, author co-author, paper-author, paper-venue, author-venue, paper-text feature, author-text networks. Their edge weights are defined as follows.

Paper citation network:

$$\mathrm{W}_{PP}(p_i, p_j) = \begin{cases} e^{-\rho(T_{evaluate} - T_{p_j})} & if\ paper\ p_j\ cites\ paper\ p_i \\ 0 & otherwise \end{cases} \tag{A1}$$

where $\rho = 2$, $T_{evaluate}$ denotes the evaluation year, T_{p_j} denotes the publication year of paper p_j.

Author co-author network:

$$\mathrm{W}_{AA}(a_i, a_j) = \sum_{\substack{p_k \epsilon A_P(a_i) \\ p_k \epsilon A_P(a_j)}} e^{-\rho(T_{evaluate} - T_{p_k})} \tag{A2}$$

where $A_P(a_i)$ denotes all the papers written by a_i.

Paper-author network:

$$\mathrm{W}_{PA}(p_j, a_i) = \mathrm{W}_{AP}(a_i, p_j) = \begin{cases} 1 & if\ author\ a_i\ writes\ paper\ p_j \\ 0 & otherwise \end{cases} \tag{A3}$$

Paper-venue network:

$$\mathrm{W}_{PV}(p_i, v_j) = W_{VP}(v_j, p_i) = e^{-\rho(T_{evaluate} - T_{p_i})} \tag{A4}$$

Author-venue network:

$$\mathrm{W}_{AV}(a_i, v_j) = \mathrm{W}_{VA}(v_j, a_i) = \sum_{\substack{p_k \in A_P(a_i) \\ p_k \in V_P(v_j)}} e^{-\rho(T_{evaluate} - T_{p_k})} \tag{A5}$$

where $V_P(v_i)$ denotes all papers published in v_i.

Paper-text feature network:

$$\mathrm{W}_{PF}(p_j, f_i) = \mathrm{W}_{FP}(f_i, p_j) = tf\_idf(f_i, p_j) \tag{A6}$$

where $tf\_idf(f_i, p_j)$ is the tf-idf weight of the feature f_i.

Author-text network:

$$\mathrm{W}_{AF}(a_i, f_j) = \mathrm{W}_{FA}(f_j, a_i) = \frac{f(f_j, a_i)}{\max\{f(f_k, a_i) : f_k \in f(a_i)\}} \cdot log \frac{|V_A|}{|V_A{}^{f_j}|} \tag{A7}$$

where $f(f_j, a_i)$ is the frequency of feature f_j used by author a_i, $\max\{f(f_k, a_i) : f_k \in f(a_i)\}$ returns the highest feature frequency of author a_i, $|V_A|$ is the number of authors, and $|V_A{}^{f_j}|$ is the number of authors using feature f_j.

Note that MRCoRank regarding each year as a time window and the degree of innovativeness of feature x_i in the time window $<t_{j-1}, t_j>$ is defined as:

$$\mathrm{E}_{x_i}^{<t_{j-1},t_j>} = \frac{|x_i^{<t_{j-1},t_j>} - \tilde{\lambda}_i|}{\tilde{\lambda}} \cdot [\sum_{s=1}^{u} (\frac{x_i^{<t_{j-1},t_j>} - x_i^{<t_{j-s-1},t_{j-s}>}}{\tilde{\lambda}_i}) \frac{1}{s}] \cdot e^{-\rho(t_j - t_0)} \tag{A8}$$

where $x_i^{<t_{j-1},t_j>}$ is the feature frequency of text feature x_i in the j_{th} time window, $\tilde{\lambda}_i$ is the estimated mean frequency of text feature x_i, $\tilde{\lambda}$ is the estimated mean frequency of all the text features, u is the number of previous windows and set to 3, ρ is the time-decaying parameter, t_0 is the time when the text feature first appears in the paper collection.

Based on the weighted networks defined, MRCoRank ranks the future influence of multi-entities by using Algorithm A1 as follows:

ALGORITHM A1: MRCoRank

Procedure MRCoRank ($\alpha_{pp}, \alpha_{pa}, \alpha_{pv}, \alpha_{aa}, \alpha_{va}, \alpha_{v}, \alpha_{f}$)

Input: the entity sets V_P, V_A, V_V, V_F, edge sets E_{PP}, E_{PA}, E_{PV}, E_{COA}, E_{AV}, E_{PF}, E_{AF}, weight matrixes W_{PP}, W_{PA}, W_{PV}, W_{AA}, W_{AV}, W_{PF}, W_{AF} and parameters $\alpha_{pp}, \alpha_{pa}, \alpha_{pv}, \alpha_{aa}, \alpha_{va}, \alpha_{v}, \alpha_{f}$, ε

Output: the vectors of all the papers, authors, venues and features $\boldsymbol{PS}, \boldsymbol{AS}$, $\boldsymbol{VS}$ and $\boldsymbol{FS}$

// Initialize $\boldsymbol{PS}, \boldsymbol{AS}$, $\boldsymbol{VS}$ and $\boldsymbol{FS}$ scores for papers, authors, venues and features, respectively

$\boldsymbol{PS}^0 = \frac{I_P}{|V_P|}, \boldsymbol{AS}^0 = \frac{I_A}{|V_A|}, \boldsymbol{VS}^0 = \frac{I_V}{|V_V|}, \boldsymbol{FS}^0 = \frac{I_F}{|V_F|}$ //$|V|$: the number of entities in V, I_P is the unit vector.

Δ =2 ε

while Δ > ε **do**

// update the paper vector **PS**, author vector **AS**, venue vector **VS** and text feature vector **FS**

$\boldsymbol{PS}^{t+1} = \alpha_{pp}(\mathrm{W}_{PP}\boldsymbol{PS}^t) + \alpha_{pa}(\mathrm{W}_{PA}\boldsymbol{AS}^t) + \alpha_{pv}(1-\alpha_{pp}-\alpha_{pa})(\mathrm{W}_{PV}\boldsymbol{VS}^t) + (1-\alpha_{pv})(1-\alpha_{pp}-\alpha_{pa})(\mathrm{W}_{PF}\boldsymbol{FS}^t)$

$\boldsymbol{AS}^{t+1} = \alpha_{aa}(\mathrm{W}_{AA}\boldsymbol{AS}^t) + \alpha_{pa}(\mathrm{W}_{AP}\boldsymbol{PS}^t) + \alpha_{av}(1-\alpha_{aa}-\alpha_{pa})(\mathrm{W}_{AV}\boldsymbol{VS}^t) + (1-\alpha_{av})(1-\alpha_{aa}-\alpha_{pa})(\mathrm{W}_{AF}\boldsymbol{FS}^t)$

$\boldsymbol{VS}^{t+1} = \alpha_{v}(\mathrm{W}_{VA}\boldsymbol{AS}^t) + (1-\alpha_{v})(\mathrm{W}_{VP}\boldsymbol{PS}^t)$

$\boldsymbol{FS}^{t+1} = [\alpha_{f}(\mathrm{W}_{FA}\boldsymbol{AS}^t) + (1-\alpha_{f})(\mathrm{W}_{FP}\boldsymbol{PS}^t)]\boldsymbol{E}^t$

Normalize $\boldsymbol{PS}^{t+1}$, $\boldsymbol{AS}^{t+1}$, $\boldsymbol{VS}^{t+1}$ and $\boldsymbol{FS}^{t+1}$

Δ= $||\boldsymbol{PS}^{t+1} - \boldsymbol{PS}^t||_1 + ||\boldsymbol{AS}^{t+1} - \boldsymbol{AS}^t||_1 + ||\boldsymbol{VS}^{t+1} - \boldsymbol{VS}^t||_1 + ||\boldsymbol{FS}^{t+1} - \boldsymbol{FS}^t||_1$

End while

According to the original paper, the best ranking is achievable for all the entities when $\alpha_{pp} = 0.6$, $\alpha_{pa} = 0.2$, $\alpha_{pv} = \alpha_{av} = 0.4$, $\alpha_{aa} = \alpha_{v} = \alpha_{f} = 0.5$, $\varepsilon = 1\mathrm{e}-6$.

MutualRank

Six sub-networks are defined and used. Their edge weights are defined as follows:

Paper citation network:

$$\mathrm{P}(p_i, p_j) = \begin{cases} 1 & \text{if paper } p_j \text{ cites paper } p_i \\ 0 & \text{otherwise} \end{cases} \tag{A9}$$

Author citation network:

$$\mathrm{A}(a_i, a_j) = \sum_{\substack{p_k \in A_P(a_i) \\ p_l \in A_P(a_j)}} \mathrm{P}(p_k, p_l) \tag{A10}$$

where $A_P(a_i)$ denotes all the papers written by a_i.

Venue citation network:

$$\mathrm{V}(v_i, v_j) = \sum_{\substack{p_k \in V_P(v_i) \\ p_l \in V_P(v_j)}} \mathrm{P}(p_k, p_l) \tag{A11}$$

where $V_P(v_i)$ denotes all the papers published in venue v_i.

Paper-author network:

$$\mathrm{PA}(p_i, a_j) = \mathrm{AP}(a_j, p_i) = \begin{cases} 1 & \text{if author } a_j \text{ writes paper } p_i \\ 0 & \text{otherwise} \end{cases} \tag{A12}$$

Paper-venue network:

$$\mathrm{PV}(p_i, v_j) = \mathrm{VP}(v_j, p_i) = \begin{cases} 1 & \text{if paper } p_i \text{ is published in venue } v_j \\ 0 & \text{otherwise} \end{cases} \tag{A13}$$

Author-venue network:

$$\mathrm{AV}(a_i, v_j) = \mathrm{VA}(v_j, a_i) = \begin{cases} 1 & \text{if author } a_i \text{ publishes papers in venue } v_j \\ 0 & \text{otherwise} \end{cases} \tag{A14}$$

Thus each network can be represented as a matrix. Rather than using the original matrix, its transition matrix is used. Let us take paper citation network as an example. Let P be the adjacency matrix corresponding to paper citation network, $\widetilde{\mathrm{P}}$ is the transition matrix by normalizing each row of P, the transition matrix $\widetilde{\mathrm{P}}$ is rewritten as $\overline{\mathrm{P}}$ in PageRank-like formalization.

$$\overline{\mathrm{P}} = \lambda\left(\widetilde{\mathrm{P}} + d_r \frac{e^T}{|V_P|}\right) + (1-\lambda)\frac{ee^T}{|V_P|} \tag{A15}$$

where $|V_P|$ is the number of papers, d_r is an $|V_P|$-dimensional vector indicating dangling nodes, e is a $|V_P|$-dimensional identity vector.

After transforming all nine adjacency matrices P, PA, PV, A, AP, AV, V, VP and VA, we obtain transition matrices $\overline{\mathrm{P}}$, $\overline{\mathrm{PA}}$, $\overline{\mathrm{PV}}$, $\overline{\mathrm{A}}$, $\overline{\mathrm{AP}}$, $\overline{\mathrm{AV}}$, $\overline{\mathrm{V}}$, $\overline{\mathrm{VP}}$ and $\overline{\mathrm{VA}}$.

Based on these networks, MutualRank ranks the influence of multi-entities by using Algorithm A2 as follows:

ALGORITHM A2: MutualRank

Procedure MutualRank $(\alpha, \beta, \gamma, \delta, \lambda, \varepsilon)$
Input: the entity sets V_P, V_A, V_V, edge sets E_{PP}, E_{PA}, E_{PV}, E_{AA}, E_{AV}, E_{VV}, transition matrixes $\overline{\mathrm{P}}$, $\overline{\mathrm{PA}}$, $\overline{\mathrm{PV}}$, $\overline{\mathrm{A}}$, $\overline{\mathrm{AP}}$, $\overline{\mathrm{AV}}$, $\overline{\mathrm{V}}$, $\overline{\mathrm{VP}}$, $\overline{\mathrm{VA}}$ and parameters $\alpha, \beta, \gamma, \delta, \lambda, \varepsilon$
Output: the authority of papers **paut**, the soundness of papers **psnd**, the importance of authors **rimp** and the prestige of venues **vprs**
// Initialize **paut**, **psnd**, **rimp** and **vprs** scores for papers, papers, authors and venues, respectively
$\mathbf{paut}^0 = \frac{I_P}{|V_P|}, \mathbf{psnd}^0 = \frac{I_P}{|V_P|}, \mathbf{rimp}^0 = \frac{I_A}{|V_A|}, \mathbf{vprs}^0 = \frac{I_V}{|V_V|}$ //$|V|$: the number of entities in V, I_P is the unit vector.
$\Delta = 2\varepsilon$
while $\Delta > \varepsilon$ **do**
// update the vectors of **paut**, **psnd**, **rimp** and **vprs**
$\mathbf{paut}^{t+1} = \alpha\,\mathbf{paut}^t + \beta(1-\alpha)\overline{\mathrm{P}}^T\,\mathbf{psnd}^t + \delta(1-\alpha)(1-\beta)\overline{\mathrm{AP}}^T\,\mathbf{rimp}^t + (1-\alpha)(1-\beta)(1-\delta)\overline{\mathrm{VP}}^T\,\mathbf{vprs}^t$
$\mathbf{psnd}^{t+1} = \alpha\,\mathbf{psnd}^t + \beta(1-\alpha)\overline{\mathrm{P}^T}^T\,\mathbf{paut}^t + \delta(1-\alpha)(1-\beta)\overline{\mathrm{VP}}^T\,\mathbf{vprs}^t + (1-\alpha)(1-\beta)(1-\delta)\overline{\mathrm{AP}}^T\,\mathbf{rimp}^t$
$\mathbf{rimp}^{t+1} = \alpha\overline{\mathrm{A}}^T\,\mathbf{rimp}^t + \beta(1-\alpha)\overline{\mathrm{VA}}^T\,\mathbf{vprs}^t + \delta(1-\alpha)(1-\beta)\overline{\mathrm{PA}}^T\,\mathbf{paut}^t + (1-\alpha)(1-\beta)(1-\delta)\overline{\mathrm{PA}}^T\,\mathbf{psnd}^t$
$\mathbf{vprs}^{t+1} = \alpha\overline{\mathrm{V}}^T\,\mathbf{vprs}^t + \beta(1-\alpha)\overline{\mathrm{AV}}^T\,\mathbf{rimp}^t + \delta(1-\alpha)(1-\beta)\overline{\mathrm{PV}}^T\,\mathbf{psnd}^t + (1-\alpha)(1-\beta)(1-\delta)\overline{\mathrm{PV}}^T\,\mathbf{paut}^t$
Normalize $\mathbf{paut}^{t+1}$, $\mathbf{psnd}^{t+1}$, $\mathbf{rimp}^{t+1}$ and $\mathbf{vprs}^{t+1}$
$\Delta = ||\mathbf{paut}^{t+1} - \mathbf{paut}^t||_1 + ||\mathbf{psnd}^{t+1} - \mathbf{psnd}^t||_1 + ||\mathbf{rimp}^{t+1} - \mathbf{rimp}^t||_1 + ||\mathbf{vprs}^{t+1} - \mathbf{vprs}^t||_1$
End while

Parameters are set as: $\alpha = 0.3$, $\beta = 0.45$, $\delta = 0.25$, $\lambda = 0.85$, $\varepsilon = 1e-6$.

Tri-Rank
Seven sub-networks are defined and their weights are as follows:
Paper citation network:

$$\mathrm{W}_{PP}(p_i, p_j) = \begin{cases} 1 & F_A(p_i) \neq F_A(p_j) \\ 0.8 & F_A(p_i) = F_A(p_j) \end{cases} \tag{A16}$$

where $F_A(p_i)$ is the first author of p_i.
Author citation network:

$$\mathrm{W}_{CA}(a_i, a_j) = \sum_{\substack{p_k \in A_P(a_i) \\ p_l \in A_P(a_j)}} \mathrm{W}_{PP}(p_k, p_l) \tag{A17}$$

where $A_P(a_i)$ denotes all the papers written by a_i.
Author co-author network:

$$\mathrm{W}_{COA}(a_i, a_j) = |A_P(a_i) \cap A_P(a_j)| \tag{A18}$$

Considering both author citation and co-authorship, the weight between author a_i and a_j is:

$$\mathrm{W}_{AA}(a_i, a_j) = \mathrm{W}_{CA}(a_i, a_j) + \mathrm{W}_{COA}(a_i, a_j) \tag{A19}$$

Venue citation network:

$$\mathrm{W}_{VV}(v_i, v_j) = \sum_{\substack{p_k \in V_P(v_i) \\ p_l \in V_P(v_j)}} \mathrm{W}_{PP}(p_k, p_l) \tag{A20}$$

where $V_P(v_i)$ denotes all the papers published in venue v_i.

Paper-author network:

$$\mathrm{W}_{PA}(p_i, a_j) = \mathrm{W}_{AP}(a_j, p_i) = \frac{1}{order(p_i, a_j)} \tag{A21}$$

where $order(p_i, a_j)$ is the position of author a_j in the author list of paper p_i.

Paper-venue network:

$$\mathrm{W}_{PV}(p_i, v_j) = \mathrm{W}_{VP}(v_j, p_i) = \begin{cases} 1 & \text{if paper } p_i \; is\ published\ in\ venue\ v_j \\ 0 & otherwise \end{cases} \tag{A22}$$

Author-venue network:

$$\mathrm{W}_{AV}(a_i, v_j) = \mathrm{W}_{VA}(v_j, a_i) = \sum_{\substack{p_k \in A_P(a_i) \\ p_k \in V_P(v_j)}} \mathrm{W}_{PA}(p_k, a_i) \tag{A23}$$

Based on the weighted network defined above, Tri-Rank ranks the influence of papers, authors and venue by using Algorithm A3 as follows:

ALGORITHM A3: Tri-Rank

Procedure Tri-Rank ($\alpha, \beta, \gamma, \varepsilon$)
Input: the entity sets V_P, V_A, V_V, edge sets E_{PP}, E_{PA}, E_{PV}, E_{AA}, E_{AV}, E_{VV}, weight matrixes W_{PP}, W_{PA}, W_{PV}, W_{AA}, W_{AV}, W_{VV}, and parameters $\alpha, \beta, \gamma, \varepsilon$
Output: the scores of all the papers, authors and venues $\boldsymbol{PS}, \boldsymbol{AS}$ and $\boldsymbol{VS}$
// Initialize $\boldsymbol{PS}, \boldsymbol{AS}$ and $\boldsymbol{VS}$ scores for papers, authors and venues, respectively
$\boldsymbol{PS}^0 = I_P, \boldsymbol{AS}^0 = I_P, \boldsymbol{VS}^0 = I_P$ //I_P is the unit vector.
$\boldsymbol{PS}^1 = \mathrm{W}_{PP}\boldsymbol{PS}^0$ // do one-iteration PageRank
$\Delta = 2\ \varepsilon$
while $\Delta > \varepsilon$ **do**
// update the vectors of $\boldsymbol{PS}, \boldsymbol{AS}$ and $\boldsymbol{VS}$
$\boldsymbol{AS}^{t+1} = \beta(\mathrm{W}_{AP}\boldsymbol{PS}^{t+1}) + (1-\beta)(\mathrm{W}_{AV}\boldsymbol{VS}^{t})$
Normalize $\boldsymbol{AS}^{t+1}$
$\boldsymbol{AS}^{t+2} = \mathrm{W}_{AA}\boldsymbol{AS}^{t+1}$
$\boldsymbol{VS}^{t+1} = \gamma\frac{\mathrm{W}_{VP}\boldsymbol{PS}^{t+1}}{|V_P|} + (1-\gamma)\frac{\mathrm{W}_{VA}\boldsymbol{AS}^{t+2}}{|V_A|}$
Normalize $\boldsymbol{VS}^{t+1}$
$\boldsymbol{VS}^{t+2} = \mathrm{W}_{VV}\boldsymbol{VS}^{t+1}$
$\boldsymbol{PS}^{t+2} = \alpha(\mathrm{W}_{PV}\boldsymbol{VS}^{t+2}) + (1-\alpha)(\mathrm{W}_{PA}\boldsymbol{AS}^{t+2})$
Normalize $\boldsymbol{PS}^{t+2}$
$\boldsymbol{PS}^{t+3} = \mathrm{W}_{PP}\boldsymbol{PS}^{t+2}$
$\Delta = \frac{||\boldsymbol{PS}^{t+3} - \boldsymbol{PS}^{t+1}||_1}{|V_P|} + \frac{||\boldsymbol{AS}^{t+2} - \boldsymbol{AS}^{t}||_1}{|V_A|} + \frac{||\boldsymbol{VS}^{t+2} - \boldsymbol{VS}^{t}||_1}{|V_V|}$
End while

We set $\alpha = 0.8$, $\beta = 0.7$, $\delta = 0.9$, $\varepsilon = 1\mathrm{e}-8$ as in the original paper.

FutureRank

Three sub-networks are defined as follows:

Paper citation network:

$$\mathrm{W}_{PP}(p_i, p_j) = \begin{cases} 1 & if\ paper\ p_j\ cites\ paper\ p_i \\ 0 & otherwise \end{cases} \tag{A24}$$

Paper-author network:

$$\mathrm{W}_{PA}(p_i, a_j) = \mathrm{W}_{AP}(a_j, p_i) = \begin{cases} 1 & if\ author\ a_j\ writes\ paper\ p_i \\ 0 & otherwise \end{cases} \tag{A25}$$

Time-ware factor:

$$R_{p_i}^{Time} = e^{-\rho*(T_{evaluate} - T_{p_i})} \tag{A26}$$

where $T_{evaluate}$ denotes the evaluation year, T_{p_i} is publication year of p_i.

Based on the weighted network, FutureRank ranks the future influence of papers by using Algorithm A4 as follows:

ALGORITHM A4: FutureRank

Procedure FutureRank ($\alpha, \beta, \delta, \rho, \varepsilon$)
Input: the entity sets V_P, V_A, edge sets E_{PP}, E_{PA}, weight matrixes W_{PP}, W_{PA}, and parameters $\alpha, \beta, \delta, \rho, \varepsilon$
Output: the scores of all the papers, authors $\boldsymbol{PS}$ and $\boldsymbol{AS}$
// Initialize $\boldsymbol{PS}$ and $\boldsymbol{AS}$ scores for papers and authors respectively
$\boldsymbol{PS}^0 = \frac{I_P}{|V_P|}, \boldsymbol{AS}^0 = \frac{I_A}{|V_A|}$ //$|V|$: the number of entities in V, I_P is the unit vector.
$\Delta = 2\varepsilon$
while $\Delta > \varepsilon$ **do**
// update the vectors of $\boldsymbol{PS}$ and $\boldsymbol{AS}$
$\boldsymbol{PS}^{t+1} = \alpha * \mathrm{W}_{PP}\boldsymbol{PS}^t + \beta * \mathrm{W}_{PA}\boldsymbol{AS}^t + \delta * R^{Time} + (1 - \alpha - \beta - \delta) * \frac{I_P}{|V_P|}$
$\boldsymbol{AS}^{t+1} = \mathrm{W}_{AP}\boldsymbol{PS}^{t+1}$
Normalize $\boldsymbol{PS}^{t+1}$, $\boldsymbol{AS}^{t+1}$
$\Delta = ||\boldsymbol{PS}^{t+1} - \boldsymbol{PS}^t||_1 + ||\boldsymbol{AS}^{t+1} - \boldsymbol{AS}^t||_1$
End while

$\alpha = 0.4$, $\beta = 0.1$, $\delta = 0.5$, $\rho = 0.62$, $\varepsilon = 1\mathrm{e} - 6$.

TAPRank

Three subnetworks are defined as follows:

Paper citation network:

$$\mathrm{W}_{PP}(p_i, p_j) = \begin{cases} \dfrac{e^{-\rho*(T_{evaluate} - T_{p_j})}}{L(p_j)} & if\ paper\ p_j\ cites\ paper\ p_i \\ 0 & \text{otherwise} \end{cases} \tag{A27}$$

where $T_{evaluate}$ denotes the evaluation time, T_{p_j} is publication time of p_j, $L(p_j)$ is the sum of outgoing links of p_j.

Paper-author network:

$$\mathrm{W}_{AP}(a_j, p_i) = \begin{cases} \dfrac{1}{C(p_i)} & if\ author\ a_j \text{ writes paper } p_i \\ 0 & \text{otherwise} \end{cases} \tag{A28}$$

$$\mathrm{W}_{PA}(p_i, a_j) = \begin{cases} \dfrac{1}{C(a_j)} & if\ author\ a_j \text{ writes paper } p_i \\ 0 & \text{otherwise} \end{cases} \tag{A29}$$

where $C(p_i)$ is the total number of authors of p_i, $C(a_j)$ is the total number of papers written by author a_j.

Time-ware factor:

$$R_{p_i}^{Time} = \frac{e^{-\rho*(T_{evaluate} - T_{p_i})}}{\sum_{p_k \in V_P} e^{-\rho*\left(T_{evaluate} - T_{p_k}\right)}} \tag{A30}$$

Based on the weighted network, TAPRank rank the future influence of papers by using Algorithm A5 as follows:

ALGORITHM A5: TAPRank

Procedure TAPRank $(\lambda, \rho, \varepsilon)$
Input: the entity sets V_P, V_A, edge sets E_{PP}, E_{PA}, weight matrixes W_{PP}, W_{PA}, and parameters $\lambda, \rho, \varepsilon$
Output: the scores of all the papers, authors $\boldsymbol{PS}$ and $\boldsymbol{AS}$
// Initialize $\boldsymbol{PR}$ scores for all papers
$\boldsymbol{PR}^0 = \frac{I_P}{|V_P|}$ //$|V|$: the number of entities in V, I_P is the unit vector.
// compute the time-ware PageRank value of papers
while $\Delta > \varepsilon$ **do**
    $\boldsymbol{PR}^{t+1} = \lambda * \mathrm{W}_{PP}\boldsymbol{PR}^t + (1-\lambda) * R^{Time}$
    $\Delta = ||\boldsymbol{PR}^{t+1} - \boldsymbol{PR}^t||_1$
End while
// Initialize $\boldsymbol{PS}$ and $\boldsymbol{AS}$ scores for papers and authors respectively
$\boldsymbol{PS}^0 = \boldsymbol{PR}, \boldsymbol{AS}^0 = \frac{I_A}{|V_A|}$
while $\Delta > \varepsilon$ **do**
// update the vectors of $\boldsymbol{PS}$ and $\boldsymbol{AS}$
    $\boldsymbol{AS}^{t+1} = \lambda * \mathrm{W}_{AP}\boldsymbol{PS}^t + (1-\lambda) * \frac{I_A}{|V_A|}$
    $\boldsymbol{PS}^{t+1} = \lambda * \mathrm{W}_{PA}\boldsymbol{AS}^{t+1} + (1-\lambda) * \frac{I_P}{|V_P|}$
    Normalize $\boldsymbol{PS}^{t+1}$, $\boldsymbol{AS}^{t+1}$
    $\Delta = ||\boldsymbol{PS}^{t+1} - \boldsymbol{PS}^t||_1 + ||\boldsymbol{AS}^{t+1} - \boldsymbol{AS}^t||_1$
End while

Parameter setting: $\rho = 0.3$, $\lambda = 0.85$, $\varepsilon = 1\mathrm{e}-6$.

PageRank
PageRank uses paper citation network as follows:
Paper citation network:

$$\mathrm{W}_{PP}(p_i, p_j) = \begin{cases} \dfrac{1}{L(p_j)} & \text{if paper } p_j \textit{ cites paper } p_i \\ 0 & \textit{otherwise} \end{cases} \tag{A31}$$

where $L(p_j)$ is the sum of outgoing links of p_j, or the number of papers that are cited by p_j.
PageRank ranks the influence of papers by using Algorithm A6 as follows:

ALGORITHM A6: PageRank

Procedure PageRank (λ, ε)
Input: the entity sets V_P , edge sets E_{PP}, weight matrixes W_{PP}, and parameters λ, ε
Output: the scores of all the papers $\boldsymbol{PR}$
// Initialize $\boldsymbol{PR}$ scores for papers
$\boldsymbol{PR}^0 = \frac{I_P}{|V_P|}$ //$|V|$: the number of entities in V, I_P is the unit vector.
// update the vectors of $\boldsymbol{PR}$
while $\Delta > \varepsilon$ **do**
    $\boldsymbol{PR}^{t+1} = \lambda * \mathrm{W}_{PP}\boldsymbol{PR}^t + (1-\lambda) * \frac{I_P}{|V_P|}$
    $\Delta = ||\boldsymbol{PR}^{t+1} - \boldsymbol{PR}^t||_1$
End while

$\lambda = 0.85$, $\varepsilon = 1\mathrm{e}-6$.