CocaRank: A Collaboration Caliber-based Method for Finding Academic Rising Stars

Jun Zhang, Feng Xia, Wei Wang, Xiaomei Bai, Shuo Yu, Teshome Megersa Bekele, Zhong Peng
School of Software, Dalian University of Technology
Dalian 116620, China
f.xia@acm.org

ABSTRACT

Evaluating the scientific impact of scholars has been studied by researchers from various disciplines for a long time. However, very few efforts have been devoted to evaluate the future potential of researchers based on their performance at the initial stage of scientific careers. Academic rising stars represent junior researchers who may not be very outstanding among the peers at the initial stage of their careers, but tend to become influential scholars in the future. In this paper, we propose a novel method named CocaRank, which integrates our proposed new indicator called the collaboration caliber, the typical indicator citation counts and hybrid calculation results on heterogeneous academic networks, to find academic rising stars. In addition, we investigate the appropriate time interval for the prediction of rising stars. The experimental results on real datasets demonstrate that our method can find more top ranked rising stars with higher average citation counts than other state-of-art methods.

Keywords

Rising star; Heterogeneous network; PageRank; HITS

1. INTRODUCTION

A scientist enters upon his or her research career by cooperating with other scholars and publishing academic articles. The initial stage of scientific career is filled with challenges, however, surviving this stage can make researchers become a genuine, innovative and responsible scientist. During the initial stage of scientific research, junior researchers have to face many problems such as which scholar's work should be paid more attention to? Whom they can cooperate with? How to get supports from both the governments and institutions? The scientific issues implied from the above questions are similar with problems of scientific evaluation, expert finding, cooperation prediction etc. Although many research achievements are made on the above issues, little attention has been paid on identifying and quantifying the role of scientific initiates. In this paper, we propose a novel method based on the heterogeneous academic networks to find potential academic initiates, which can be deemed as rising stars.

Currently, there is neither uniform definition of rising star nor mature evaluation criteria for rising stars. In our paper, rising stars refer to researchers who are presently not outstanding among peers or with low research profiles, but will grow into influential or authoritative scholars in the future. Finding rising stars can shed light on a lot of scientific questions, such as providing candidates for peer reviews, searching potential cooperators, offering basis for award appraisal and foundation application etc. Therefore, identifying and evaluating rising stars with a relatively fair and comprehensive method is crucial and essential.

The existing approaches for the evaluation of scientific impact are mainly classified into three categories: citation based methods, network based methods and hybrid methods. The citation counts are widely used for evaluating the influence of articles and scholars, commonly used methods of such are like JIF [1], h-index [2] and g-index [3] etc. However, these traditional methods, which consider the number of publications and its citations, have some obvious shortcomings. In order to improve the ranking positions of themselves, researchers may only care about the quantity of publications and pay little attention to the quality of papers. In addition, these traditional citation based methods are biased towards senior researchers, because the value of which aggregates along with the growth of academic career and it is unfair to young outstanding researchers.

Due to the drawbacks of traditional citation based methods, researchers have proposed a variety of modified evaluation measures. In research cases whereby the problem involving the contributions of coauthors in the same paper are different, Stallins et al. [4] proposed an axiomatic approach to assign relative credits to the coauthors of a given paper, which can better capture a researcher's scientific impact than simply relying on the total numbers of publications and citations of the authors'. In order to solve the time delay problem in h-index, Pan et al. [5] proposed an author impact factor, which can capture trends and variations of the impact of the scientific output of scholars in time. Also most of the existing methods treat all the citations equally, however, some researchers argue that citations from influential scientists or significant papers should be considered more important and weighted more [6]. It is obvious that simply applying this kind of methods for rising stars' evaluation is not appropriate, because citation counts are related to pa-

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2890524.

pers' publication time, authors' reputation, venues' quality and etc., which are all difficult for rising stars to achieve.

Considering the above drawbacks of citation based methods, researchers are using network based methods to evaluate the scientific impact. The major components of academic social network are articles, authors, venues and links among them. Typical algorithms include degree centrality, PageRank [7] and HITS [8]. However, these methods are designed for homogeneous networks which consist of only one type of node and relationship. Nevertheless, with the emergence of more and more entities and relationships in academic social networks, it is impossible to depict it as a homogeneous network. Hence, it is essential to evaluate scholars under appropriate network topologies. Besides, using only one kind of indicators cannot capture the scientific impact accurately, thus recent solutions are more in favor of the hybrid methods which consider both citation and network based methods together to calculate the importance of authors [9, 10].

Recent solutions for finding rising star are also mainly focus on applying hybrid methods. Li et al. [11] proposed a PubRank algorithm to identify rising stars in research communities by mining the social networks of researchers in terms of their co-authorship relationships, which considered mutual influence and static ranking of conferences or journals. However, two main problems exist, i.e. it did not consider the authors' contribution and the dynamic ranking of venues when evaluating rising stars. Therefore, Daud et al. [12] proposed an algorithm named StarRank, which made some improvements of PubRank and took the above mentioned shortcomings into consideration.

Although the studied methods can solve the problem of finding rising stars, there still exist flaws on the current solutions. One important factor that needs to be considered in rising star selection is the time when researcher started his or her own career. It is unfair to put junior researchers together with senior researchers and use the same evaluation standards, in which apparently senior scholars have more advantages. The current methods seem like ignoring this vital fact, and they generally choose a stochastic time interval for rising star's evaluation without considering that various researchers may go through different stage of their careers. As a consequence, it is essential to propose a new method that considers this crucial fact.

Another important fact that also needs to be taken into account is the social relationships that rising stars build through their careers. It is well acknowledged that rising stars can benefit a lot by cooperating with senior researchers. The current methods tend to use the mutual influence among coauthors of the same paper or the mutual reinforcement among papers, authors and venues to calculate the scientific impact [13]. However, beyond the mutual influence, there still exists one problem which has been ignored by most of the current works. Some researchers may prefer to cooperate with a static group of scholars or insist on working alone, while others may prone to collaborate with scholars from different institutions, countries, disciplines and etc. For rising stars, it is important for them to establish relationships with various researchers which can not only help them to broaden their horizons, but also improve their academic capacities. Therefore, the ability to collaborate with other scholars also needs to be considered in the evaluation of rising stars.

In this paper, by considering the facts mentioned above, we propose a novel rising stars' evaluation method, which combines the above two factors and measurements of node's influence based on heterogeneous academic networks together to evaluate the scientific impact of rising stars. The contribution of this paper can be summarized as follows:

- We propose a novel indicator named the collaboration caliber, which can capture rising stars' capacity of cooperating with various scholars.
- We use the publishing time of the first paper as the start line of rising star's evaluation instead of randomly choosing a stochastic time interval, aiming to avoid the unfair situations for rising stars in comparison to senior scholars.
- In order to find out the appropriate time interval for the evaluation of rising stars, we select several time intervals at the beginning stage of rising stars for comparison. According to the results on real datasets, we find that the growth trend of the initial 3 to 5 years is the fastest, while the future achievements of rising stars are closely related to the performances from 5 to 7 years.
- We evaluate the influence of rising stars with our proposed method CocaRank, and the experiments on real datasets demonstrate that our CocaRank outperforms other state-of-art methods.

The rest of this paper is organized as follows: in the next section, we present our novel method in detail. The evaluation results and analyses on real datasets are provided in section 3. Lastly, we expound the conclusion and future work in the section 4.

2. METHOD

Since using single indicator cannot capture the scientific impact of rising stars accurately, we propose a novel hybrid method, which combines the proposed statistical indicators and the topological structures together to calculate the influence of rising stars. Unlike previous studies which utilize the constant time interval of 5 or 6 years and the start line is randomly chosen without considering different scholars' diverse research periods, we choose 4 different time intervals. In addition, the start line is set at the time when researchers published their own papers for the first time.

Generally there are three steps in our proposed method: a) We first calculate the value of our proposed indicator named the collaboration caliber for each scholar; b) We then compute the PageRank score for each paper in citation network, which is used in calculating the HITS scores of authors and journals in author-paper network and paper-journal network respectively; c) Finally, we combine the values of the above two steps and compute the final score for each scholar according to our proposed CocaRank method. In this section, we will present our method in detail.

2.1 The Collaboration Caliber

The individual's academic level not only relies on personal scientific capacity but also relates to their social interactions, which imply the cooperation relationship between senior researchers and junior researchers. Researchers tend to cooperate with influential scholars, since they can benefit more than the improvement of academic ability, but also correct attitudes towards their scientific career. Apart from this mutual influence among researchers which has been studied by a lot of related works, the impact of cooperation among scholars with diverse backgrounds still remains to be explored.

With the advancement of science and technologies, researches on various issues are more and more depending on the integration of multiple disciplines instead of counting on the knowledge of single area as previous. The collaboration of diverse disciplines promotes both the quantity and quality of the scientific achievements, which also credits to the more and more frequent cooperation among researchers. Therefore, for an individual researcher, the ability of collaborating with other people is becoming an important component of their whole scientific careers. In this paper, for the first time and to the best of our knowledge, we propose a novel indicator named the Collaboration Caliber (Coca) to capture the scholar' ability of cooperation with other people. The specific method is illustrated as follows.

To calculate the value of Coca, we first introduce the concept of entropy. In physics, entropy is used to reveal the orderliness of a system, the larger value of it indicates the system is more chaotic, which means the system is worse. However, the founder of theory of communication C. E. Shannon [14] redefined the understanding of entropy. In theory of communication, entropy is used to quantify the usefulness of information. On the contrary to the definition in physics, the larger value of entropy suggests the richer information content. The intrinsic significance of entropy is similar to the measurement of collaboration we are using in this paper, and we apply the standard equation to calculate the value of scholar's entropy. The Coca value of rising star is computed according to the following equation.

$$Entropy(a_i) = -\sum_{i=1}^{r} W_i \log_2 (W_i) \tag{1}$$

$$Coca(a_i) = \sum_{t=1}^{n} Entropy(a_i) \tag{2}$$

where a_i represents an author, W_i is the possibility of the words in all the institutions' information of a_i's cooperators, and r is the total number of the words, which are the names of all the collaborators' institutions. $Coca(a_i)$ is the sum of author's entropy according to specific time intervals, where n refers to the time intervals as we set.

2.2 Calculations on Heterogeneous Academic Networks

In previous rising star's studies, academic networks are depicted as homogenous networks, which consist of only one kind of node and one type of link. However, it is a truth universally acknowledged that more than one type of nodes and relationships exist in the real academic networks. Typical real academic networks contain several kinds of nodes such as authors, venues, papers and etc., and a variety of links like cooperation, citation, publication and etc. As a consequence, it is improper to evaluate the influence of rising stars under homogeneous networks. In this paper, we apply our combination algorithms to calculate the influence of rising stars based on heterogeneous academic social networks.

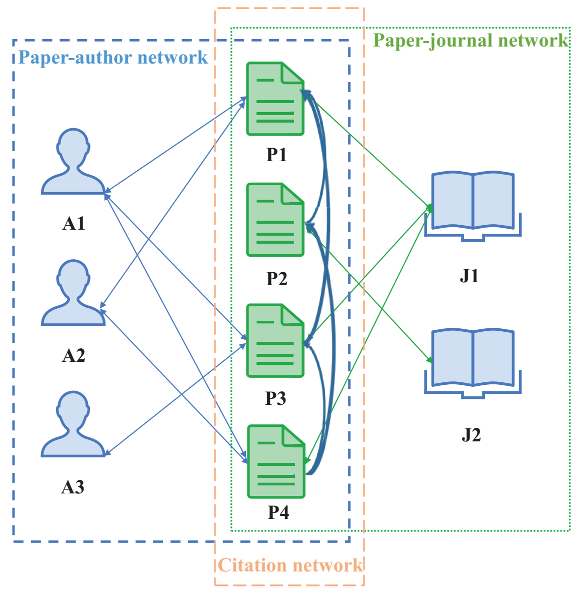

Figure 1: Illustration of heterogeneous academic networks.

To apply the combination algorithms on the heterogeneous networks, we first need to construct it. It consists of three sub-networks, which are citation network, paper-author network and paper-journal network as shown in Figure 1:

- Citation network. It composes of one type of entity, i.e. paper, and one type of link which is the citation relationship between papers.
- Paper-author network. It contains two types of entities, i.e. papers and the corresponding authors, and two kinds of links which are the citation relationship between papers, and writing relationship between papers and authors.
- Paper-journal network. It also has two types of nodes, i.e. papers and journals. Two types of links are included, which are citation relationship between papers, and publication relationship between papers and journals.

2.2.1 Calculation of Paper Influence

The citation network can be denoted as a graph represented by $G_p = (V_p, R_p)$, where V_p is the set of papers, and R_p is the set of citation relationship between papers. G_p is a directed network, where a directed edge exists if one paper cites another. Under this network topology, we measure the influence of papers based on the classical PageRank algorithm.

PageRank algorithm is designed to rank the importance of websites according to search engines' results originally. Its basic mechanism is that nodes would have high rankings if the nodes pointing to it have high scores, or influential nodes are more likely to be pointed to compare with other nodes. Currently, PageRank algorithm is widely used to rank not only the importance of websites, but also extended to calculate the importance of scholars, papers and etc. The following equation is used to calculate the PageRank score of each article:

$$PR(p_i) = \frac{1-d}{N} + d\sum_{j=1}^{m} \frac{PR(p_j)}{L(p_j)} \tag{3}$$

where p_i represents the node under consideration, N is the total number of the nodes in the network, p_j is the node that links to p_i and $L(p_j)$ is the number of outbound links on p_j. $PR(p_i)$ is the visiting probability of node p_i, which interprets the importance of p_i, and $PR(p_j)$ can be defined corresponding as $PR(p_i)$. d is the damping factor, which is the probability, at any step, the node p_i is visited following the links pointing to it. Various studies have investigated the impact of different damping factors, and it is generally assumed that the damping factor is set as 0.85 in our paper. At every step of PageRank iteration process, we update the PageRank score of each paper according to Eq. 3. At last, the iteration process stops until the scores of all the nodes converge to a stable state, and we get the final PageRank scores of all papers.

2.2.2 Calculation of Author Influence

In this part, we construct the undirected paper-author network to measure the importance of authors. It can be represented by graph $Ga = (V_a, R_a)$, where V_a is the set of papers and the corresponding authors, and R_a is the set of citation relationship between papers, and writing relationship between paper and author. The following calculation of authors' influence is applied on the above paper-author network and using HITS algorithm.

The main purpose of HITS algorithm is also calculate and rank the importance of entities under different network topologies. The HITS algorithm assigns two scores for each node: its authority and hub value. The former estimates the value of the node's content, and the latter computes the value of its links to other nodes. For nodes known as hubs, it serve as large directories that are not actually authoritative in its own content, but can lead users directly to other authoritative nodes. In other words, a good hub represents a node that points to many other nodes, and a good authority represents a node that is linked by many different hubs. The following equations are used to measure the authority score of each author:

$$auth(a_k) = \sum_{i=1}^{s} hub(l_i) \tag{4}$$

$$hub(a_k) = \sum_{i=1}^{v} auth(p_i) \tag{5}$$

where a_k is the node, s is the total number of nodes that link to a_k in the network, l_i is nodes that link to it and $auth(a_k)$ represents the influence score of it. v is the number of nodes that a_k points to, p_i indicates the node that a_k points to. In the initial step, if a_k is an author, we set the initial value of $auth(a_k)$ and $hub(a_k)$ as 1, else the values of $auth(a_k)$ and $hub(a_k)$ are set equal to the PageRank score of a_k.

2.2.3 Calculation of Journal Influence

In the following part, we construct the undirected paper-journal network to measure the influence of journals. It can be represented a graph $Gj = (V_j, R_j)$, which V_j is the set of papers and the corresponding journals, and R_j is the set of citation relationship between papers, and publication relationship between paper and journal. As we mentioned above, the HITS algorithm is also applied to calculate the importance of journals under paper-journal network. The following equations are used to calculate the authority score of each journal:

$$auth(j_k) = \sum_{i=1}^{u} hub(l_i) \tag{6}$$

$$hub(j_k) = \sum_{i=1}^{x} auth(p_i) \tag{7}$$

where j_k represents a node, u is the total number of nodes that link to it in the network, l_i is the node that point to it and $auth(j_k)$ represents the influence score of it. x is the number of nodes that j_k point to, and p_i is the node that j_k point to. In the initial step, if j_k is a paper, we set the initial value of $auth(j_k)$ and $hub(j_k)$ equal to the value of $auth(j_k)$ we get last step, else the values of $auth(j_k)$ and $hub(j_k)$ are set equal to 1.

2.2.4 Calculation of CocaRank

After finishing the above two parts' calculation, we then propose our CocaRank method. In our method, it contains three main parts, which are author's citation counts, CC value and the total combination results on the importance of papers, journals and authors under the above mentioned three sub-networks. The following composite equation is used to calculate the final score of authors:

$$\begin{aligned} CocaRank(a_i) = & auth(a_i)\{\sum_{p=1}^{n} ord(a_i)PR(p_i)auth(j_k)\} \\ & Cita(a_i)Coca(a_i) \end{aligned} \tag{8}$$

where $CocaRank(a_i)$ is the final score of author a_i, p is the number of total papers written by author a_i, and $Cita(a_i)$ is the total citation counts of it. $ord(a_i)$ means the author's sequence in a paper, representing the author's contribution in a paper, and set as $1/n$ for simplicity, where n is the sequence of the author in a paper. $PR(p_i)$ is the paper's PagaRank score in citation network, $auth(a_i)$ is the HITS score of author a_i in paper-author network and $auth(j_k)$ is the respectively journal's HITS score in paper-journal network.

3. EXPERIMENTS AND RESULTS

In this section, we apply our CocaRank method on the real datasets, and evaluate its performance. For the ranking of rising stars, there is no ground truth of it currently. Therefore, we adopt the scholar's future citation counts as the ground truth to validate whether these rising stars have achieved their expectations. The higher citation counts of a scholar, the more outstanding he or she is. We then list the top 10 researchers in each comparison methods, which we will introduce in this section. Finally, we compare the average citation counts, and calculate the Spearman Correlation Coefficient between the citation counts and each rank list in different time intervals, which are set as the initial 3, 5, 7 and 10 years of researchers' scientific career.

Table 1: Top 10 Rising Stars by CocaRank

Ranking-T1	CitCounts	Ranking-T2	CitCounts	Ranking-T3	CitCounts	Ranking-T4	CitCounts
W. C. Wester. III	7984	W. C. Wester. III	7984	K. Honscheid	10093	W. C. Wester. III	7984
L. Demortier	10450	L. Demortier	10450	W. C. Wester. III	7984	K. Honscheid	10093
D. Benjamin	9599	Y. Rozen	1833	L. Demortier	10450	L. Demortier	10450
D. Zhang	1474	K. Honscheid	10093	L. Zhang	8889	L. Zhang	8889
L. Z. Wang	1059	M. Garcia-Sciveres	5933	Y. Rozen	1833	M. Garcia-Sciveres	5933
L. I. Glazman	2896	D. Benjamin	9599	M. Garcia-Sciveres	5933	Y. Rozen	1833
J. Xu	859	D. Zhang	1474	D. Benjamin	9599	G. Introzzi	7229
A. Navin	584	M. Selen	5412	S. Kopp	5063	E. Kajfasz	5670
M. Jiang	80	J. K. Nelson	2260	E. Kajfasz	5670	D. Benjamin	9599
S. Muto	105	H. Tajima	2399	M. Selen	5412	S. Kopp	5063

Table 2: Top 10 Rising Stars by StarRank

Ranking-T1	CitCounts	Ranking-T2	CitCounts	Ranking-T3	CitCounts	Ranking-T4	CitCounts
T. Abbott	206	T. Abbott	206	T. Abbott	206	J. Zhang	11577
R. Bellotti	287	R. Bellwied	4111	J. Zhang	11577	P. Fallon	2228
X. Zhao	1919	P. Fallon	2228	P. Fallon	2228	T. Abbott	206
D. M. Cullen	485	R. Binder	875	D. Benjamin	9599	D. Benjamin	9599
R. Bellwied	4111	D. Benjamin	9599	R. Bellotti	287	Z. V. Vardeny	1099
Richard T. Scalettar	994	R. Bellotti	287	R. Bellwied	4111	Bellotti	287
R. Binder	875	S. J. Yennello	1319	L. I. Glazman	2896	R. Bellwied	4111
H. Namatame	1362	Z. V. Vardeny	1099	R. Binder	875	Andrew R. Liddle	2880
L. I. Glazman	2896	H. Namatame	1362	L. I. Mazin	4413	L. I. Mazin	4413
S. J. Yennello	1099	S. V. Greene	3446	Z. V. Vardeny	1099	Francesco Sciortino	1066

3.1 Dataset

The sub dataset used for our experiments is acquired from the American Physical Society (APS) datasets. It contains the detailed information of each article from 12 physical journals, which includes article's DOI, title, authors, the date of publication, venues, authors' affiliations and its citation relationships. To conduct our experiments, we first pre-process the metadata. In order to evaluate and validate the scientific impact of rising stars, we choose scholars that the publication time of their first paper is at the same year, and their academic career's lengths up to validation time are the same. Therefore, we choose researchers who published their first articles in 1993, and their academic career is not ended until 2013.

3.2 Comparison Methods

In order to evaluate the performance of our proposed method, we compare the performances of our methods' different variants, and also evaluate the results of StarRank, which is chosen as the state-of-art method for comparison. The details of the above methods are as follows:

- **CocaRank**. This is our proposed method, which integrates citation counts, our proposed indicator CC's value and the combination results under heterogeneous academic networks.
- **CocaRank-PaHit**. It is the variant of our proposed method CocaRank, which only considers the combination results under heterogeneous networks to evaluate the ranking of rising stars.
- **StarRank**. This evaluation approach is introduced above, and we choose it as the state-of-art method for comparison.

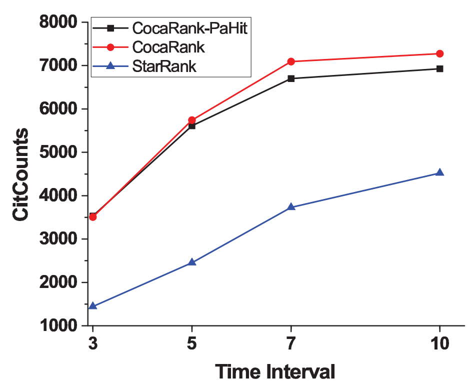

Figure 2: Comparison of future average CitCounts in different time intervals.

3.3 Results

As we assumed above, the higher citation counts indicate the more excellent of a researcher. We use CitCounts to represent the real citation counts of all the scholars up to

2013, which are chosen as our experiments' dataset. The time interval we select is 3, 5, 7, and 10 years, and we use Rank-T3, Rank-T5, Rank-T7 and Rank-T10 to present the ranking in the above time intervals respectively. We first show the top 10 researchers' ranking list by our CocaRank method together with their citation counts in the four time intervals as shown in Table 1. Comparing with the results shown in Table 2, our method can more accurately selects the top ranking scholars according to citation counts than the StarRank method.

Besides, we also infer from both Table 1 and Table 2 that the number of common members of two consecutive ranking rows is increasing with the prolonging of time intervals, however, when the time interval is extended to a comparative long time, the variation between two consecutive ranking rows is small. For instance, in Table 1, it exists 4 common members between Rank-T3 and Rank-T5, 7 common members between Rank-T5 and Rank-T7 and 9 common members between Rank-T7 and Rank-T10.

Table 3: Comparison of Spearman Correlation Coefficient

Time Interval	CocaRank	CocaRank-PaHit	StarRank
3	0.76391	0.59248	-0.05716
5	0.84361	0.71729	-0.04962
7	0.76541	0.73836	0.34887
10	0.75789	0.69925	0.08722

To further verify our inference obtained above, we then compute the top 10 rising stars' average citation counts by applying our proposed method and the above mentioned comparison methods in different time intervals. As shown in Figure 2, the average CitCounts by our proposed CocaRank is the highest among the three methods in the four time intervals. It is also observed that the growth rate of average citation counts is the highest between 3 and 5 years, then decreases as the time interval prolonging from 7 to 10 years. In other words, the first 3 to 5 years is a very crucial stage in scholar's whole scientific career, which researchers grow rapidly in this period. While the increase rate from 5 to 7 years is not so drastic comparing with the rate in 3 to 5 years, but the researchers' future influence is closely related to the performances of this period.

In addition, we also calculate the Spearman Correlation Coefficient to measure the correlation between the CitCounts and the above methods for comparison. The value of Spearman Correlation Coefficient varies from -1 to 1 with correlation ranging from the most negative to the most positive. As shown in Table 3, the value of CocaRank is the highest among the three methods, and clearly indicate that our method has made an remarkable improvement compare to other methods.

4. CONCLUSION AND FUTURE WORK

In this paper, we propose a new approach to find rising stars. The proposed CocaRank method integrates both the merits from statistical area and the topological features of academic network, which combines the value of Coca, citation counts and the importance calculation results on heterogeneous academic networks. In addition, we explore the effects of different time intervals on the evaluation of rising stars, and it is interesting to notice that the growth trend of citation counts in the initial 3 to 5 years are the fastest, then decreases a little from 5 to 7 years and becomes flat from 7 to 10 years. The results on APS datasets show that our approach achieves a more appropriate performance than other methods in the selection of top ranking rising stars.

In future work, we will test the performance of our method on more datasets and examine its reliability. Furthermore, more types of nodes and relationships in heterogeneous academic networks would be considered, as well as find more essential indicators that influence the ranking of rising stars.

5. REFERENCES

[1] E. Garfield. The history and meaning of the journal impact factor. *Jama*, 295(1):90–93, 2006.

[2] J. E. Hirsch. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.*, 102(46):16569–16572, 2005.

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

[4] J. Stallings, E. Vance, J. Yang, M. W. Vannier, J. Liang, L. Pang, L. Dai, I. Ye, and G. Wang. Determining scientific impact using a collaboration index. *Proc. Natl. Acad. Sci. U.S.A.*, 110(24):9680–9685, 2013.

[5] R. K. Pan and S. Fortunato. Author impact factor: tracking the dynamics of individual scientific impact. *Sci. Rep.*, 4, 2014.

[6] M. Valenzuela, V. Ha, and O. Etzioni. Identifying meaningful citations. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[7] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: bringing order to the web. 1999.

[8] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5):604–632, 1999.

[9] Y. Wang, Y. Tong, and M. Zeng. Ranking scientific articles by exploiting citations, authors, journals, and time information. In *Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

[10] M. Nykl, K. Ježek, D. Fiala, and M. Dostal. Pagerank variants in the evaluation of citation networks. *J. Informetr.*, 8(3):683–692, 2014.

[11] X. Li, C. S. Foo, K. L. Tew, and S. K. Searching for rising stars in bibliography networks. In *Database System Advanced Applications*, pages 288–292. Springer, 2009.

[12] A. Daud, R. Abbasi, and F. Muhammad. Finding rising stars in social networks. In *Database System for Advanced Applications*, pages 13–24. Springer, 2013.

[13] Z. Liu, H. Huang, X. Wei, and X. Mao. Tri-rank: An authority ranking framework in heterogeneous academic networks by mutual reinforce. In *26th IEEE International Conference on Tools with Artificial Intelligence*, pages 493–500. IEEE, 2014.

[14] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.