ACRec: A Co-authorship based Random Walk Model for Academic Collaboration Recommendation

Jing Li, Feng Xia, Wei Wang, Zhen Chen, Nana Yaw Asabere, Huizhen Jiang
School of Software, Dalian University of Technology, Dalian 116620, China
f.xia@acm.org

ABSTRACT

Recent academic procedures have depicted that work involving scientific research tends to be more prolific through collaboration and cooperation among researchers and research groups. On the other hand, discovering new collaborators who are smart enough to conduct joint-research work is accompanied with both difficulties and opportunities. One notable difficulty as well as opportunity is the big scholarly data. In this paper, we satisfy the demand of collaboration recommendation through co-authorship in an academic network. We propose a random walk model using three academic metrics as basics for recommending new collaborations. Each metric is studied through mutual paper co-authoring information and serves to compute the link importance such that a random walker is more likely to visit the valuable nodes. Our experiments on DBLP dataset show that our approach can improve the precision, recall rate and coverage rate of recommendation, compared with other state-of-the-art approaches.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval - *Information filtering*

Keywords

Big scholarly data, collaboration recommendation, link importance, random walk model

1. INTRODUCTION

We are witnessing an epoch-making proliferation of information available on the Internet. Meanwhile, information overload prevents users from acquiring relevant information. To tackle this problem, the emergence of recommendation systems and techniques is now playing a critical part of the contemporary internet world by bringing people closer to the resources they really need.

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2579034.

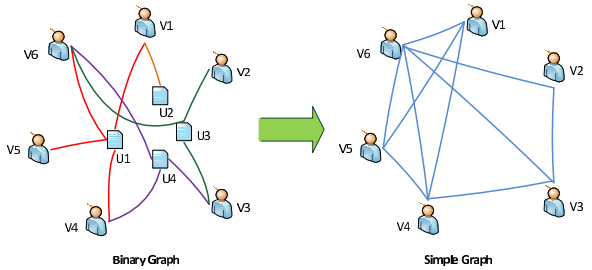


Figure 1: Extraction from a Binary Graph to a Simple Graph.

In the context of academic social networks, researchers approach big scholarly data problems such as how to obtain information pertaining to a valuable collaborator. Previous studies have confirmed that researchers or research groups with well connected cooperation networks tend to be more prolific as well as productive scholars are more inclined to cooperation [8]. Thus, it is imperative and vital for researchers to get acquainted with new valuable collaborators in academic social networks [4]. To satisfy these kinds of demands, for example making friends in academic social networks, many methods have been proposed to suggest potential new links, including link prediction and link recommendation [14, 15].

A social network (SN) can be described in an abstract way, as a graph of nodes (such as users or groups) that have certain relationships, for example, friendship and co-authorship. In traditional SN, link prediction and link recommendation methods have made significant improvements of link creations. A feature in SN called "People You May Know" has proved to be of merit in recommending friends based on a FOF (friend of friends) method [17]. Besides, typical SN systems usually recommend friends that the users already know offline based on social relationship [2]. However, in the academic context, social relationship has a different meaning considering some academic backgrounds (e.g., research interests, co-authoring information, and academic reputation). Therefore, recommending researchers in academic social networks (based on big scholarly data) is an increasingly important topic. Our focus in this paper is on Academic Social Networks (ASN) where social links are formed by certain academic ties. For instance, two researchers from the same research institute will be connected for the social tie of same affiliation. Within ASN, a co-authorship social network is an extraordinary social network due to the

academic property of co-authorship, which is a simple graph evolving from the author-paper binary graph (as shown in Fig.1). In the research world, suggesting new links can motivate researchers to build new collaborative relationships when they consider writing a new paper, then help them acquire papers of high quality. Since co-authorship social networks are of intrinsic collaboration related values, we present an academic random walk model for collaboration recommendation accordingly.

In summary, we make the following contributions in this paper: 1) To deal with scientific collaborator recommendation in the context of big scholarly data, we develop a model based on random walk with restart that learns how to bias a random walk on the network so that it can visit the potential collaborators with more probability than the others. 2) In order to improve the recommendation quality and accuracy, we propose to define the link importance by exploiting three specific academic network metrics including coauthor order, collaboration time points (i.e. the latest collaboration time) and frequency of collaboration (i.e. collaboration times). 3) We conduct extensive experiments on DBLP dataset to evaluate the performance of the proposed solution in various scenarios as compared against the basic model of RWR. Promising results are presented and analyzed.

2. RELATED WORK

Social networks have been studied for decades in an effort to comprehend the relationships between people and detect patterns in such interactions. Recently much research work has been done on how to utilize social network information to improve recommender systems [4, 6]. For instance, Ma et al. [11] elaborated on how the incorporation of social network information is beneficial in improving recommender systems. Perugini et al. [13] suggested that recommendation has an intrinsic social element that is intended to connect people. In contrast to previous work in traditional friend recommendation field, there is also some other research work on collaboration recommandation. Lopes et al. [10] considered the researcher's publications area and the vector space model to generate collaboration recommendation in academic social network. Besides, co-authorship social networks analysis has been studied for a long time, of which there are some positive outcomes in terms of collaboration recommendation. Newman [12] studied a variety of statistical properties of scientific collaboration networks and found out that researchers in different disciplines have different numbers of collaborators on average.

Considering combining the network structure with the features of nodes [1], some methods (for example link prediction and link recommendation) based on Random Walk with Restart (RWR) have been proposed in the state-of-the-art literature. RWR provided a good way to measure how close related two nodes are in a graph [16]. In [7], Mohsen et al. proposed a random walk model which combined the trust-based and collaborative filtering approaches for recommendation. Fouss et al. [5] presented a new perspective on characterizing the similarity among elements of a graph.

However, current research either lacks combining the network structure with the features of nodes, or treats the links with equal importance in RWR, neglecting whether the relationship is strong or not. To tackle these drawbacks, we present an academic random walk model for collaboration recommendation in co-authorship social networks.

3. PROBLEM STATEMENT

The problem we define in this work involves proposing a recommender system in a co-authoring network to recommend the most valuable collaborators for a particular researcher. Given an undirected co-author graph as shown in Fig. 1, a node v denotes a scholar and other nodes compose a set of scholar candidates C that v may create a collaboration link with. We define target nodes set $T = \{t_1, \ldots, t_2, \ldots, t_k\}$ to present nodes with which v will create collaboration links in the future, while we take no-link nodes $N = \{n_1, \ldots, n_2, \ldots, n_k\}$ to present other nodes with which v does not create links. We label candidate nodes set $C = \{ci\} = TUN$. Normally, present RWR models usually overlook information about link features because they assume that links among nodes are of the same importance. This is unscientific in academic collaboration networks. Since the relationship between two researchers with only one paper co-authored long time ago can't be compared with the relationships of those who co-authored many papers together in a latest period of time. Thus Each edge (u, v) in G should have a corresponding metric vector $S(u, v)$ to describe the extent of cooperation strength of the two linked nodes (e.g., how many times u and v have co-authored, when the cooperation was last built, and the relationship labeled by author order in the co-authored paper).

In this paper, we compute the cooperation strengths by analyzing the big data set of scientific literature information which is provided by DBLP. To achieve the goal, we take three most important academic metrics to decide the edge importance. As a result, we develop a co-authorship based Random Walk model by taking those cooperation strengths as the the initial random walk transition probabilities. For each particular node, we can run an efficient and effective random walk to recommend a list of the most valuable collaborators.

4. ACADEMIC COLLABORATION RECOMMENDATION MODEL

The academic collaboration recommendation model is inspired by the truth that scholars usually desire to cooperate with people who have high academic value. Such people normally have fruitful high-quality papers, which can generally be used to represent people's academic achievements. Besides, as the RWR model has been proved to be competent for calculating the similarity of nodes in network, we use it as a basic model for the co-authorship social networks. And the three metrics we introduced into the network structure is to bias the random walk such that it will more easily traverse to the positive nodes. The detailed process of Academic RWR is described below and the corresponding pseudocode is illustrated in Algorithm 1.

The structure of Academic RWR is depicted in Fig. 2. The whole collaboration recommendation work can be called as ACRec. Firstly, we extract a collaboration graph based on volumes of co-authored paper information. Then, as part of our major contributions, we derive three academic metrics from large scholarly data and then take them to compute each edge strength (link importance), which is going to be utilized to guide a co-authorship based random walk model (also called as Academic RWR). When Academic RWR ends, we can generate a top N recommendation list. Below we will detail how to derive the three co-authorship

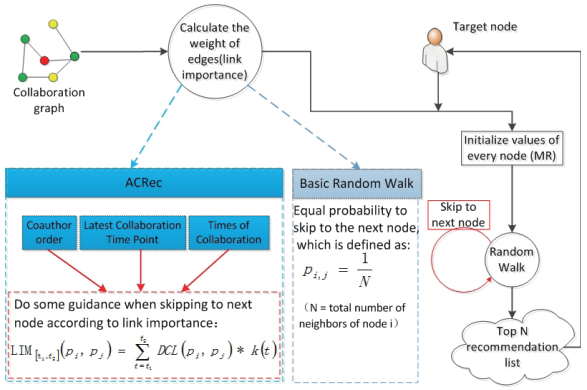


Figure 2: The Structure of ACRec.

Algorithm 1 ACRec(R, a, MaxIteration, MinDelta)

$S \leftarrow$ ComputeTransferMatrix()
$MR_0 \leftarrow R$
$Q \leftarrow R$
for $k \leftarrow 0$ to $MaxIteration - 1$ **do**
  $diff \leftarrow 0$
  for $i \leftarrow 0$ to $len(Q) - 1$ **do**
    $MR_{k_i} = \alpha \sum_{j=0}^{len(Q)} S_{i,j} MR_j + (1-\alpha) Q_i$
    $diff \leftarrow diff + (MR_k - MR_{k-1})$
  end for
  if $diff < MinDelta$ **then**
    break
  end if
end for
$Predictions \leftarrow predictions(MR)$
return $Predictions$

metrics those are co-author order, latest collaboration time point, and times of collaboration respectively. As well, we will present Academic RWR in detail by taking into account the three academic metrics.

4.1 Co-author Order

In most cases, there is a list of co-authors for one paper. Normally, their contributions to the paper differ from each other. For example, the first and the second authors usually make more contributions than the rest authors. In such cases, the cooperation relationship between the first two authors is competently strong. Moreover, the co-author order can reflect cooperation relationship strength. As a general rule, the contribution value is inversely proportional to the co-author order, and the weight of relationship is contributed by the relevant two nodes. Therefore, we propose a measure of the link importance based on the coauthor order: DCL (distance in coauthor list).

Consider two nodes p_i, p_j in a co-author list. Assume that $j > 1$, and there are more than one author of a paper. We calculate $DCL(p_i, p_j)$ as follows.

$$DCL(p_i, p_j) = \begin{cases} \frac{1}{i} + \frac{1}{j} & j \leq 3 \\ \frac{1}{i} + \frac{2}{j} & j > 3, i \leq 3 \\ \frac{2}{i} + \frac{2}{j} & i > 3 \end{cases} \quad (1)$$

According to this definition, it is clear that the DCL value between the first and the second authors is 1.5, which is the maximum. The relationship between the first two authors is the closest, while the relationship between the first author and the rest of the authors is relatively weak.

4.2 Latest Collaboration Time Point

Previous studies considered social networks to be static. However, academic social networks are time-varying, where the links among scholars change over time. For instance, scholars might be more willing to collaborate with who they co-authored a paper last month, as compared to the coauthors that they cooperated ten years ago. Hence, we measure the link dynamics using $LIM_t(p_i, p_j)$ (i.e. Link Importance):

$$LIM_t(p_i, p_j) = DCL(p_i, p_j) * k(t) \quad (2)$$

where $k(t)$ is a monotonically increasing function over time. We can measure the impact of different coauthoring time points by adjusting the parameter $k(t)$. Here, we define $k(t)$ as:

$$k(t) = \frac{t_i - t_0}{t_c - t_0} \quad (3)$$

where t_i is the link formation time (in year here), t_c is the current time (i.e. 2013 in this paper) and t_0 is the first link formation time.

4.3 Times of Collaboration

In academic social networks, if two authors coauthor a paper, there will be a link between them. Furthermore, these two authors may collaborate many times. However, no previous study has taken into consideration the times of collaboration. Here we measure the impact of different times of coauthoring as follows:

$$\begin{aligned} LIM_{[t_1,t_2]}(p_i, p_j) &= \sum_{t=t_1}^{t_2} LIM_t(p_i, p_j) \\ &= \sum_{t=t_1}^{t_2} DCL(p_i, p_j) * k(t) \end{aligned} \quad (4)$$

In (4), during time period (t_1, t_2), if there are n links between p_i and p_j, we will calculate the sum of each link importance.

4.4 The Academic RWR

We have already introduced three co-authorship metrics and detailed how to compute link importance above. Now we can take account of using them to conduct a random walk in the co-authorship social network for collaboration recommendation. We define the link importance between the nodes p_i and p_j as $w_{i,j}$. Then we can acquire the transfer matrix $S = ComputeTransferMatrix()$ based on the link importance, as described in algorithm 1. To be more specific we use P_i as the current node and P_j as the next node. S is the set of probabilities for each P_i in G skipping to next node P_j. This can be described in equation (5).

$$S_{i,j} = \frac{W_{i,j}}{\sum_{P_k \in N(P_i)} W_{i,k}} \quad (5)$$

while $N(P_i)$ is the set of neighbors of P_i.

Our random walk starts with initializing the rank score vector $MR^{(0)}$ as well as the restart probability vector q as (0,...,1,...,0). Set target node P_i as 1 while others 0. The random walk iterates the traversal starting with node P_i

until it stops walking and assigns each candidate node P_k a stable probability MR_k. Thus we get the rank score vector MR. Then sort the nodes by the corresponding rank score. For a node P_i, the initializing $MR^{(0)}$ can be described as:

$$MR(p_i) = \frac{1-\alpha}{N} + \alpha \sum_{p_j \in M(p_i)} \frac{MR(p_j)}{L(p_j)} \quad (6)$$

where MR represents the rank score vector, $MR(p)$ is the rank score of node p, which is the quantized importance of node p to the target node. $M(p_i)$ is the set of nodes incident to node p_i, with $L(p_j)$ being the number of all the neighbors of node p_j. α denotes the probability of the walker continuing walking to the next neighbor.

Equation (6) represents only the step to get the rank score of a node. As for each node in the whole graph, the iterative process is defined by (7), which is also a personalized random walk model.

$$MR^{(t+1)} = \alpha \mathbf{S} MR^{(t)} + (1-\alpha)q \quad (7)$$

In (7), $MR^{(t)}$ represents the rank score vector at step t, and q is the row vector, and its form is $(0, \ldots, 1, \ldots, 0)$. In fact, at the beginning, $MR^{(0)} = q$, and the rank score of target node is 1, while others' are 0. $\mathbf{S}$ is the transfer matrix, representing the probability for each node to skip to other nodes, as the definition above.

Now we can recommend nodes in the TOP N of the list MR to target nodes. Of course, we can take the nodes out from the TOP N list, which have been in its co-author list before recommending.

5. EVALUATION AND ANALYSIS

We conducted various experiments using data from DBLP [9], a computer science bibliography website hosted at University Trier. The datasets we extracted are all in the field of data mining involving 34 journals and 49 conferences altogether. We divided the dataset into two parts: the data before year 2011 as a training set, and others as a testing set.

We designed different experiments to compare ACRec with the basic model of RWR in terms of the aforementioned metrics (i.e. precision, recall rate and coverage rate). For each experiment, we examine separate aspects, including the influence of different parameters, the effect of the three metrics we count on, and the performance on the best settings.

5.1 Influence of Various Parameters

In this section, we present the exploration on the influence of different parameters, including range of target nodes' degree, damping coefficient and the number of recommended nodes. When the effect of a parameter is under exploitation, the other parameters are assigned with the default values. After the experiment, we can attain the best values of them for the tests afterwards.

5.1.1 Target Nodes' Degree

In academic social network, the degree of strong node is larger than that of weak node. To evaluate the influence of the target nodes' degree to the experiment, we defined four ranges of degree according to the features of our dataset, and the ranges are shown in Fig. 3.

As shown in Fig. 3, the target node's degree has an obvious influence on the metrics. For a practical meaning, it is different to recommend coauthors to those who have different number of collaborators. As for precision in Fig. 3(a), the larger the target node's degree, the better the model performance. Besides, we can see ACRec relatively has higher precision than RWR. At the range of 0 to 10, ACRec performs similarly to RWR. But when the target node's degree gets larger than 30, the precision can be as high as 18.1%, much more than RWR. Thus we can conclude that, ACRec has the higher precision for strong node, but not obvious for weak node.

Fig. 3(b) shows the comparison of recall rate with the changing degree. The first two columns are almost the same for recall rate, while the gaps between the two models get larger for other columns, which appears the same as precision. i.e. when the degree gets larger than 30, the corresponding recall rate of ACRec is 12.3%, much higher than that of RWR (10.4%). Hence we can claim that ACRec performs better than RWR on recall rate with varying target node's degree.

We can see the effect of target nodes' degree on the coverage rate from Fig. 3(c). The overall trend of coverage is distinct from the former metrics. The values of both models decrease respectively from 2.3 to 0.95 and 2.3 to 0.9. The result means that, for weak nodes, the neighbouring network becomes sparser with less valuable information, leading to the random walk going further. And for strong nodes, there are enough valuable nodes to be recommended in neighbouring network.

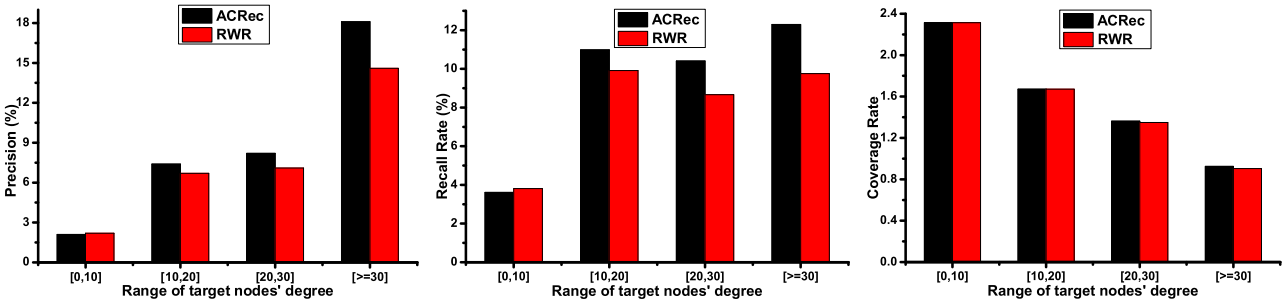
The reason counting for the phenomenon is also due to that, weak nodes are not so active than strong nodes, and there is not enough valuable information to analyze and make recommendation. The analysis above leads us to the conclusions that ACRec outperforms the RWR and it can make a better recommendation for strong nodes.

5.1.2 Damping Coefficient

In Random Walk model, there is a damping coefficient. According to the equation (3), we know that the value of damping coefficient determines the probability for random walker to jump back to the original node when randomly walking. This parameter has a realistic significance as it controls how far the MR value will be dispersed. In this section, we analyze how the damping coefficient influences the performance of the two algorithms on the three metrics.

Generally, as depicted by Fig. 4, ACRec and RWR nearly share the same trend and for the majority of tested data, but ACRec keeps recommending with higher precision, recall rate and coverage rate, compared to the basic RWR.

Fig. 4(a) shows that the precisions generally increase with the growth of damping coefficient. For ACRec, it can be as high as 18.1%, corresponding to the damping coefficient of 0.8. According to Fig. 4(b), the recall rate also comes to the highest value of 12.3% when the damping coefficient is 0.8. In Fig. 4(a) and Fig. 4(b), for ACRec, we can find that both precision and recall rate decrease when damping coefficient is over 0.8. Moreover, from the Fig. 4(c), we can see that the coverage rate generally decreases until the damping coefficient is over 0.8, and then increases rapidly. Therefore, there is a trade-off between recommendation precision and coverage. In all, if the value of coefficient is 0.8, the performance is not bad for ACRec.



(a) Precision (b) Recall Rate (c) Coverage Rate

Figure 3: Performance of ACRec and Basic RWR over Target Nodes' Degree

Table 1: Experimental Results of the Two Models

Model	Precision	Recall Rate	Coverage Rate
RWR	15.3%	10.156%	0.967
ACRec	18.1%	12.187%	0.99

5.1.3 Number of Recommended Nodes

Fig. 5 illustrates how the number of recommended nodes influence the performance of ACRec and RWR with respect to precision, recall rate and coverage rate.

Fig. 5(a) shows the trend of precision. We can easily find that the precision decreases dramatically with the number of recommended nodes increasing. The highest precision of ACRec is 16.2% when we recommend 10 nodes to a target node while the highest precision of RWR is 13.3% when we return a 10-node recommendation list.

As for the performance of recall rate, Fig. 5(b) shows that the recall rate increases gradually, which is opposite to that of precision. Fig. 5(c) also depicts clearly that precision is almost inverse to coverage. Additionally, it is shown by the figure that ACRec performs a little better than basic RWR, but not obvious.

In summary, the consideration of academic social metrics (i.e. co-author order, latest collaboration time point and times of collaboration) helps ACRec recommend more precisely with higher recall rate, in a wider scope in a co-author network, at least not worse than the benchmark model. Besides, the parameters we take into account affect the performance in diverse manners and we have found their best values for ACRec.

5.2 Experiments in the Best Settings

After determining the three variables above, we carried out experiment to exploit how good ACRec performs on earth compared with RWR. In this section, we test 500 times for each model, and keep the target nodes same. The parameters are assigned the best values based on the former experiment, i.e. target node's degree: > 30, damping coefficient: 0.8, and number of recommended nodes: 10. And we set the time segment of testing data and training data as 2011, set the iteration times as 25;

The result of the exploitation is shown in Table 1. It's obvious that both precision and recall rate of ACRec are higher than those of RWR, as the precision of ACRec can come to 18.1% as well as the recall rate is 12.187. Besides, for the coverage rate, ACRec performs a little better than RWR, which is 0.99 in this case. In a word, we can claim that all the three factors we proposed performs well, and ACRec model is more effective than RWR.

In the meantime, some drawbacks are shown when we research others' work, the precision and recall rate of ACRec are lower than some other recommender method, for example, [3], in which the recall rate is 95.18%. Nevertheless, it does not mean that our work performs badly. It is due to the scale of data set. Our data set is huger and more pervasive than that of [3], where only 629 researchers from 45 Brazilian institutions are considered.

6. CONCLUSIONS

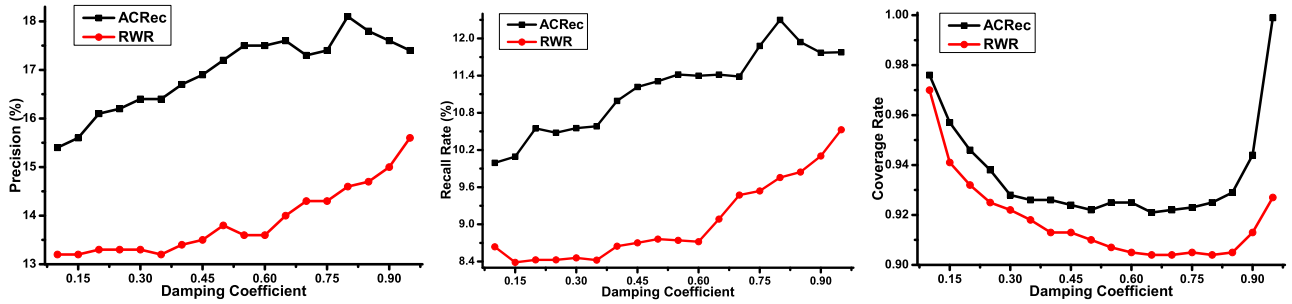
In this paper, we focus on how to find scholars' collaboration based on coauthor networks (i.e. big scholarly data) which is rarely studied in the literature. To this end, we have proposed a new model named ACRec, by injecting three academic metrics into RWR and the metrics are coauthor order, latest collaboration time point and collaboration times, constituting the weight of link importance between two authors for recommendation. We conducted extensive experiments on the DBLP data set to examine the performance of Academic RWR with respect to various aspects, including e.g. varying parameters and impact of the metrics. The experiment results show that our proposed approach performs better than RWR.

Nonetheless, there is still room for future study in this direction. We only count on three academic metrics while many other features exist, such as citation relationship. Besides, there are more reasons for two scholars with no collaboration before to cooperate. For example, they might attend the same meeting and get acquainted to each other by chance, or they are from the same institution. The relationship among coauthors of a paper is far more complicated than what we have imagined. More experiments on the entire DBLP data set may be conducted.

7. REFERENCES

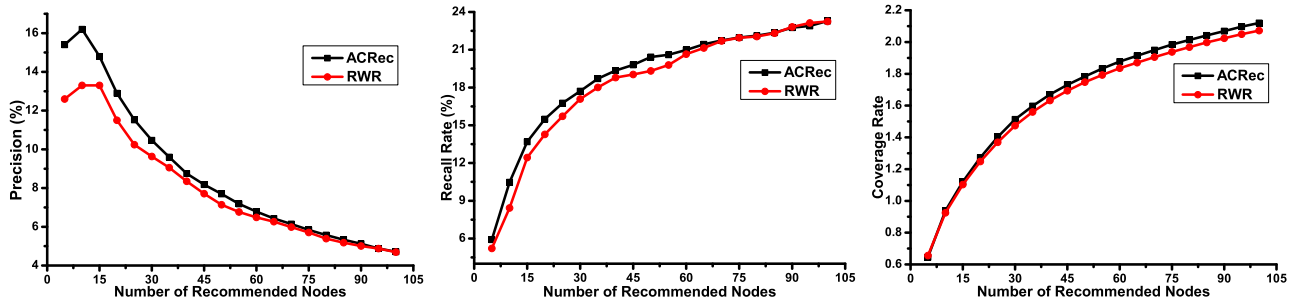
[1] L. Backstrom and J. Leskovec. Supervised random walks: Predicting and recommending links in social networks. In *Proc. of ACM WSDM'11*, pages 635–644, Hong Kong, China, 2011.

[2] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.



(a) Precision (b) Recall Rate (c) Coverage Rate

Figure 4: Performance of ACRec and Basic RWR over Damping Coefficient



(a) Precision (b) Recall Rate (c) Coverage Rate

Figure 5: Performance of ACRec and Basic RWR over Number of Recommended Nodes

[3] M. A. Brandão, M. M. Moro, G. R. Lopes, and J. P. Oliveira. Using link semantics to recommend collaborations in academic social networks. In *Proc. of WWW'13 Companion*, pages 833–840, 2013.

[4] H.-H. Chen, L. Gou, X. Zhang, and C. L. Giles. Collabseer: A search engine for collaboration discovery. In *Proc. of ACM/IEEE JCDL'11*, pages 231–240, Ottawa, Ontario, Canada, 2011.

[5] F. Fouss, A. Pirotte, J.-M. Renders, and M. Saerens. Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 19(3):355–369, 2007.

[6] J. Freyne, S. Berkovsky, E. M. Daly, and W. Geyer. Social networking feeds: Recommending items of interest. In *Proc. of ACM RecSys'10*, pages 277–280.

[7] M. Jamali and M. Ester. Trustwalker: A random walk model for combining trust-based and item-based recommendation. In *Proc. of ACM KDD'09*, pages 397–406, Paris, France, 2009.

[8] S. Lee and B. Bozeman. The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5):673–702, 2005.

[9] M. Ley. Dblp: some lessons learned. *Proceedings of the VLDB Endowment*, 2(2):1493–1500, 2009.

[10] G. Lopes, M. M. Moro, L. K. Wives, and J. P. M. de Oliveira. Collaboration recommendation on academic social networks. In *Advances in Conceptual Modeling ĺC Applications and Challenges*, volume 6413 of *Lecture Notes in Computer Science*, pages 190–199. Springer Berlin Heidelberg, 2010.

[11] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King. Recommender systems with social regularization. In *Proc. of ACM WSDM'11*, pages 287–296, Hong Kong, China, 2011.

[12] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. *Physical Review E*, 64(1):016131, Jun 2001.

[13] S. Perugini, M. A. Gonçalves, and E. A. Fox. Recommender systems research: A connection-centric survey. *Journal of Intelligent Information Systems*, 23(2):107–143, 2004.

[14] J. Tang, S. Wu, J. Sun, and H. Su. Cross-domain collaboration recommendation. In *Proc. of ACM KDD'12*, pages 1285–1293, Beijing, China, 2012. ACM.

[15] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: extraction and mining of academic social networks. In *Proc. of ACM KDD'08*, pages 990–998, Las Vegas, Nevada, USA, 2008. ACM.

[16] H. Tong, C. Faloutsos, and J.-Y. Pan. Random walk with restart: Fast solutions and applications. *Knowledge and Information Systems*, 14(3):327–346, 2008.

[17] A. Töscher, M. Jahrer, and R. Legenstein. Improved neighborhood-based algorithms for large-scale recommender systems. In *Proc. of NETFLIX'08*, pages 4:1–4:6, Las Vegas, Nevada, 2008. ACM.