



**HAL**  
open science

# Scientific Paper Recommendation Using Author's Dual Role Citation Relationship

Donglin Hu, Huifang Ma, Yuhang Liu, Xiangchun He

► **To cite this version:**

Donglin Hu, Huifang Ma, Yuhang Liu, Xiangchun He. Scientific Paper Recommendation Using Author's Dual Role Citation Relationship. 11th International Conference on Intelligent Information Processing (IIP), Jul 2020, Hangzhou, China. pp.121-132, 10.1007/978-3-030-46931-3\_12. hal-03456971

**HAL Id: hal-03456971**

**<https://inria.hal.science/hal-03456971>**

Submitted on 30 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Scientific Paper Recommendation Using Author's Dual Role Citation Relationship

Donglin Hu<sup>1</sup>, Huifang Ma<sup>1,2</sup>, Yuhang Liu<sup>1</sup>, and Xiangchun He<sup>1</sup>

<sup>1</sup>Northwest Normal University, Lanzhou Gansu 730070, China

<sup>2</sup>Guangxi Key Lab of Multi-source Information Mining & Security, Guangxi Normal University, Guilin 541004, China  
mahuifang@yeah.net

**Abstract.** Vector representations learning (also known as embeddings) for users (items) are at the core of modern recommendation systems. Existing works usually map users and items to low-dimensional space to predict user preferences for items and describe pre-existing features (such as ID) of users (or items) to obtain the embedding of the user (or item). However, we argue that such methods neglect the dual role of users, side information of users and items (e.g., dual citation relationship of authors, authoritativeness of authors and papers) when recommendation is performed for scientific paper. As such, the resulting representations may be insufficient to predict optimal author citations.

In this paper, we contribute a new model named scientific paper recommendation using Author's Dual Role Citation Relationship (ADRCR) to capture authors' citation relationship. Our model incorporates the reference relation between author and author, the citation relationship between author and paper, and the authoritativeness of authors and papers into a unified framework. In particular, our model predicts author citation relationship in each specific class. Experiments on the DBLP dataset demonstrate that ADRCR outperforms state-of-the-art recommendation methods. Further analysis shows that modeling the author's dual role is particularly helpful for providing recommendation for sparse users that have very few interactions.

**Keywords:** Matrix Factorization, Dual Role, Citation Relationship, Authoritativeness, Clustering.

## 1 Introduction

With the continuous development of information technology, scientific social networks have become the fastest and most suitable way for researchers to communicate with each other. However, as a growing numbers of scientific papers are shared in scientific social networks, which makes it difficult for researchers to locate the papers they are interested in from a large number of scientific papers. Therefore, how to recommend scientific papers of interest to researchers in social networks has become a hot research topic. Essentially, recommender system [1] provides suggestions of items that may

interest to users. At present, scientific paper recommendation methods can be divided into two categories: content-based recommendation [2] and collaborative filtering [3].

In the content-based recommendation method, researchers usually apply the title, abstract and keywords of scientific papers to generate the recommendation. Collaborative Filtering is a technique widely adapted in recommendation systems which assesses items (in this case papers) according to the former users-items interactions [4]. In this recommendation method, the researchers obtain the prediction score for the scientific papers by the researchers' scoring information on the scientific papers, and finally get a recommended list of scientific papers. Hybrid recommendation method usually combines the content-based recommendation method with the collaborative filtering recommendation method. It usually generates better recommendation results than that of recommendation methods only using one strategy, but it still has the disadvantages of data sparsity and low recommendation accuracy. To effectively improve the accuracy of scientific papers recommendation, we propose a new model named scientific paper recommendation using Author's Dual Role Citation Relationship (ADRCR), which incorporates information on both authors and papers. The contributions are summarized as follows.

- We emphasize the importance of clustering scientific papers by topic to improve the performance of recommendation.
- We develop a novel ADRCR method to learn authors' preferences for dual role by building models of explicit interactions (e.g., citation and reference) and implicit connections.
- We perform experiments on the DBLP dataset and demonstrate that the ADRCR method can improve the accuracy of scientific papers more effectively than other baseline methods.

## 2 Related Work

We review existing work on common latent space approach for recommendation and methods based on user feature matrix shared representation, which are related to our work, together with the emphasize of differences from ADRCR.

### 2.1 Common Latent Space Approach for Recommendation

The matrix decomposition method is a factual method of collaborative filtering using explicit feedback. The basic idea is to embed users and items into shared potential spaces [5]. By merging reference relationships into matrix factorization techniques, several methods have been considered from the perspective of users and items, such as item-based methods [6], user-based methods [7] and combinations of these two methods [8]. The hybrid model [8] associates the user with the paper through the label information of the paper to build a user model and a paper model. This recommendation model effectively alleviates the cold start problem. The previously proposed cross-domain model does not consider the two-way potential relationship between users and items, nor does it explicitly model user information and project characteristics. The

[9]method extracts multiple user preferences in the domain while retaining the relationship between users in different potential spaces to provide recommendations in each domain. [10] This method is based on the perspective of deep learning, taking users and items characteristics as the original input, using the proposed model to learn the potential factors of users and items, and then combining the obtained potential factors to make fast and accurate predictions. However, these methods only consider the single role of the user.

## 2.2 Method Based on User Feature Matrix Shared Representation

There are many recommendation models based on the shared representation of user feature matrices. User feature matrix shared representation refers to the simultaneous decomposition of the rating matrix and the social relationship matrix. The recommendation model assumes that the user feature matrix is hidden in both the rating information and the social information. Several remarkable works in the field [11,12] take user rating and social information into account, but they do not consider additional information about users and items. In fact, ratings and reviews are complementary and can be viewed as two different aspects of users and items. Therefore, [13]merging the scoring model and text reviews can effectively learn more accurate representations of users and recommended items. [14]This method considers the correlation between users. It use three independent autoencoders to learn user functions with roles of rater, truster and trustee, respectively. The method [12] is most similar to our method, but it only reflects the information of the user, not the information of the item. Our proposed method considers not only the user and item perspectives, but also the different roles assigned to the user.

## 3 Matrix Decomposition: A Model-based Method

### 3.1 Notation and Problem Statement

**Notation.** We first introduce some frequent notations utilized in the following sections. We use bold capital with subscripted letters to represent column vectors(e.g,  $\mathbf{M}_p$ ), and apply bold capital letters and subscripts with transpose superscript T to represent row vectors(e.g,  $(\mathbf{L}_a)^T$ ), respectively. We indicate all matrices by bold upper case letters (e.g.,  $\mathbf{Q}$ ), and  $q_{ap}$  denotes the entry of matrix  $\mathbf{Q}$  corresponding to the row  $a$  and column  $p$ . We denote a predicted value, by having a  $\hat{\cdot}$  over it (e.g.,  $\hat{q}$ ).

**Problem Statement.** Given a recommendation system with  $n$  authors and  $m$  papers,  $\mathbf{Q}=[q_{ap}]_{n \times m}$  represents the author-paper citation matrix, where  $q_{ap}$  is the number of times that the author  $a$  cites the paper  $p$ . Authors and papers are usually mapped to the low-dimensional feature space. After the  $\mathbf{Q}$  decomposition, the author  $a$  vector  $\mathbf{L}_a$  of  $k$ -dimensional and the paper  $p$  vector  $\mathbf{M}_p$  are obtained, respectively. Finally, we learn the feature matrices  $\mathbf{L}$  and  $\mathbf{M}$  by minimizing the sum of squares loss function:

$$\min_{\mathbf{L}, \mathbf{M}} \frac{1}{2} \sum_{(a,p) \in \mathcal{E}} ((\mathbf{L}_a)^T \mathbf{M}_p - q_{ap})^2 + \frac{\lambda_1}{2} \|\mathbf{L}\|_F^2 + \frac{\lambda_2}{2} \|\mathbf{M}\|_F^2 \quad (1)$$

where,  $\gamma$  is the set of observable (author, paper) pairs in  $\mathbf{Q}$ ,  $\|\cdot\|_F$  Frobenius paradigm; Regularization terms  $\|\mathbf{L}\|_F^2$  and  $\|\mathbf{M}\|_F^2$  are used to avoid over-fitting. The stochastic gradient descent algorithm is used to solve the local optimal solution of the function defined by equation (1), and the product of  $\mathbf{L}$  and  $\mathbf{M}$  is adopted to approximate the citation matrix  $\mathbf{Q}$ . For the missing items  $\hat{q}_{ap}$  in the citation matrix  $\mathbf{Q}$ , we apply the inner product of  $\mathbf{L}_a$  and  $\mathbf{M}_p$  to predict:

$$\hat{q}_{ap} = (\mathbf{L}_a)^T \mathbf{M}_p \quad (2)$$

### 3.2 Matrix Decomposition in Scientific Research Reference Network

Let  $G=(A, E, \mathbf{T}, \mathbf{W})$  denote a directed social reference relation network with  $n$  nodes, where  $A$  represents a set of authors and  $E$  represents the edge set.  $\mathbf{T}=[t_{ae}]_{n \times n}$  denotes the transfer matrix of influence propagation, and  $t_{ae}$  indicates the propagation probability from author  $a$  to author  $e$ ; If there is an edge from  $e$  to  $a$  in the social citation network (i.e.,  $e$  trusts  $a$ ), then  $t_{ae}>0$ , and otherwise,  $t_{ae}=0$ . The structure of  $G$  is described using the reference relation asymmetric matrix  $\mathbf{W}=[w_{ea}]_{n \times n}$  between authors, and  $w_{ea}$  expresses the strength of the reference relation between author  $e$  and author  $a$ , that is, the weight of the edge. Due to dissymmetrical property of citation, we map each author  $a$  of reference network as two distinct latent feature vectors, depicted by reference-specific feature vector  $\mathbf{L}_a$  and referenced-specific feature vector  $\mathbf{U}_a$ , respectively.  $\mathbf{L}_a$  and  $\mathbf{U}_a$  characterize the behaviors of ‘to reference others’ and ‘to be referenced by others’, respectively. After giving two vectors, the strength  $w_{ea}$  of the reference relationship is modeled as the inner product of  $\mathbf{L}_a$  and  $\mathbf{U}_e$ , and the feature matrix  $\mathbf{L} \in \mathbb{R}^{k \times n}$  and  $\mathbf{U} \in \mathbb{R}^{k \times n}$  can be learned by minimizing the following objective function:

$$\min_{\mathbf{L}, \mathbf{U}} \frac{1}{2} \sum_{(e,a) \in \delta} ((\mathbf{L}_a)^T \mathbf{U}_e - w_{ea})^2 + \frac{\lambda_1}{2} \|\mathbf{L}\|_F^2 + \frac{\lambda_2}{2} \|\mathbf{U}\|_F^2 \quad (3)$$

Where  $\delta$  is an observable (author, author) pair sets in  $\mathbf{W}$ , the specific calculations of  $t_{ae}$  and  $w_{ea}$  in the  $\mathbf{T}$  and  $\mathbf{W}$  matrices will be introduced below. The superscript  $c$  mentioned below represents a specific class.

Please note that both objective function (1) and objective function (3) use the idea of matrix factorization. The difference is that the objective function (1) learns the author's citations to the paper, while the objective function (3) learns authors' and authors' citations. The two objective functions are fused into the final objective function, reflecting the main purpose of this article, that is, the author's dual-citation role.

## 4 The Propose Method

### 4.1 Basic Framework

We now represent the proposed ADRCR model, the framework of which is illustrated in Fig.1.

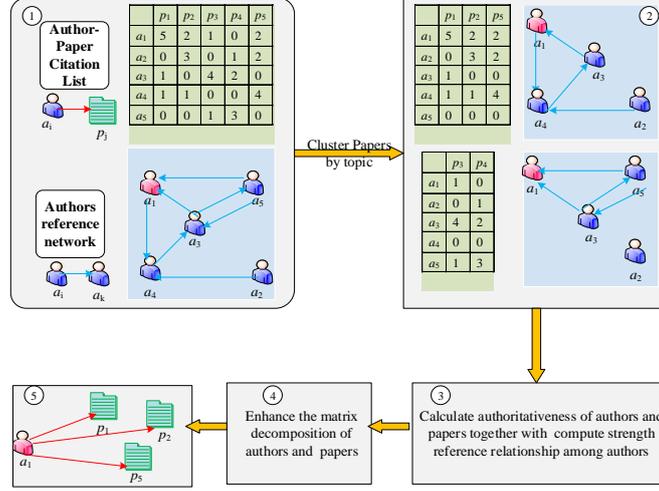


Fig. 1. The framework of ADRCR

There are five components in the framework: (1) Dividing author-paper citation matrix and reference network between authors; (2) Clustering papers by topic and deriving social reference networks for authors (excluding authors who have not cited any papers) in a specific class; (3) Calculating the strength of citation relationships between authors and the authority of authors and papers; (4) Enhancing the matrix decomposition for authors and papers; and (5) Predicting the author's citation to the paper.

## 4.2 Citation Relationship Strength and Authoritative Calculation

**Citation Relationship Strength Calculation.** In a specific scientific social reference relation network, if author  $e$  and author  $a$  have a reference relationship in class  $c$ , the number of times that author  $e$  cites author  $a$ 's paper will be used to measure the strength of the reference relation between them, that is, the weight on the edge is  $x_{ea}^c$  [15].

The author's research interests may also change over time. According to the number of papers published in the six research fields on the DBLP dataset, we construct the attribute vector of each author  $a$ , and the attribute value of each dimension corresponds to the author  $a$ . The number of papers published in these fields can be adopted to calculate the similarity of author's research interest using cosine similarity, which is defined as follows:

$$s_{ea}^c = \begin{cases} 1 & , a = e \\ \frac{\mathbf{b}_a \cdot \mathbf{b}_e}{|\mathbf{b}_a| \times |\mathbf{b}_e|} & , a \neq e \wedge (a, e) \in E \\ 0 & , others \end{cases} \quad (4)$$

Finally, the edge weights constructed by authors  $a$  and  $e$  are defined as follows:

$$w_{ea}^c = \frac{1}{1 + \lg(x_{ea}^c \times s_{ea}^c)} \quad (5)$$

*Influence calculation.* In the scientific social reference network, the structure of the node reflects the author's authority to a certain extent. Authors with higher authority can provide valuable reference information to other authors; Low authority authors are willing to refer to the suggestions of authoritative authors. In this paper, the influence propagation algorithm is used to calculate the authoritative value of the author in each class[16]. Let  $f_{a \rightarrow e}^c$  denote the influence of author  $a$  on author  $e$  in class  $c$  as follows:

$$f_{a \rightarrow e}^c = \frac{1}{1 + \lambda} \left( \sum_{k \in N_e^c} t_{ke}^c f_{a \rightarrow k}^c + v_{a \rightarrow e}^c \right) \quad (6)$$

$$s.t. \quad f_{a \rightarrow a}^c = \alpha_a^c, \quad e = a \quad \text{and} \quad \alpha_a^c > 0 \quad (7)$$

Where  $N_e^c = \{e_1, e_2, \dots, e_n\}$  denotes a set of trusted friends of author  $e$  in class  $c$  (i.e., author  $e$  refers to a class of other authors).  $t_{ke}^c$  represents the propagation probability from author  $k$  to author  $e$ , and  $\sum_{k=1}^n t_{ke}^c \leq 1$ . The parameter  $\lambda$  is the damping coefficient, and  $\alpha_a^c \in [0, 1]$  is the domain knowledge of the author  $a$ . In particular,  $f_a^c$  and  $v_{a,a}^c$  are set as default in reference [16].

*Centrality calculation.* We utilize three centralities to evaluate the status of nodes in the network. The central node can be regarded important because it has a favorable and influential position in the network. When a given node has more neighbors, it will occupy an important position. For a given node  $a$ , the final status of the node in the network is obtained by calculating the average of the three centralities[17], which is defined as follows:

$$AC_a^c = \frac{\sum_{k=1}^z C_k^c}{z} \quad (8)$$

where  $z$  is the number of centralities, and  $C_k^c$  is the centrality value of node  $a$  measured according to the centrality. In our example, the value of  $k$  is 3, since we adopt three centralities.

If author  $a$  has a greater influence on author  $e$  and a very important status, then author  $a$  is more authoritative. We sort the influence and centrality of  $n^c$  authors, then let author  $a$  rank as  $s_a^c = f_{a \rightarrow e}^c \times AC_a^c$  and  $s_a^c \in [1, n^c]$ , where  $n^c$  denotes the number of authors in  $W^c$ . The authoritative value of author  $a$  is defined as:

$$O_a^c = \frac{1}{1 + \lg(s_a^c)} \quad (9)$$

It can be seen from the above formula that when the author's influence value ranks first in class  $c$ , then  $O_a^c = 1$ , and  $O_a^c$  decreases as the author's ranking decreasing, that is, the author's influence value ranks lower, and the corresponding authoritative value is lower. Therefore, the authoritative values of all authors can be calculated using equation (9).

**Authoritative Calculation of the Paper.** Generally, the higher the authority, the more important the paper is on the scientific research platform, and the higher the citation

rate of the paper by authors in related fields. On the contrary, the lower the authority, the lower the citation rate is. We then calculate the weight  $H_{(p)}^c$  as follows:

$$H_{(p)}^c = \beta \begin{cases} 1, & p^c \in A \\ 2/3, & p^c \in B \\ 1/3, & p^c \in C \end{cases} + (1-\beta)e^{\frac{-(t-y_p)+1}{g_p+1}} \quad (10)$$

where  $\beta$  is a parameter.  $(A, B, C)$  respectively represents the collection of  $A, B$  and  $C$  category papers recommended by the CCF Association, and  $p^c \in A$  indicates that the paper  $p$  in category  $c$  belongs to  $A$  category.  $t$  represents the current time.  $y_p$  represents the publication time of the paper  $p$ , and  $g_p$  indicates the number of times the paper  $p$  has been cited. The authoritative value  $I_p^c$  of the paper  $p$  is normalized using the Sigmoid function.

### 4.3 Recommendation Model

In the scientific social reference network, if an author is more authoritative, it means that he has more citation times to high-quality papers; If the authoritativeness of a paper is higher, it demonstrates that it will be cited by many authors in the field. Therefore, this article also uses the authority of the author and the paper to measure the author's citations of the paper, namely:

$$\min_{L, M^c} \frac{1}{2} \sum_{(a,p) \in \gamma^c} O_a^c I_p^c (g((L_a^c)^T M_p^c) - q_{ap}^c)^2 + \frac{\lambda_1}{2} \|L^c\|_F^2 + \frac{\lambda_2}{2} \|M^c\|_F^2 \quad (11)$$

$\gamma^c$  is a set of pairs (authors, papers) in  $\mathcal{Q}^c$ . In formula (11), if author  $a$  has higher authority in class  $c$  and the authority of the paper is also high, it means that author  $a$  has cited the paper multiple times. That is, the error between the predicted number of citation and the actual number of citation is small; In contrast, there is a large error between the predicted citation times and the actual citation times.

In order to more conveniently learn the parameters, the author's authorization value should be set between 0 and 1, thus the time of citation is mapped to the  $[0,1]$  interval by using the function  $f(x) = x / Q_{\max}$ , where  $Q_{\max}$  is the maximum citation time. The Logistic function  $g(x) = 1 / (1 + \exp(-x))$  fixes the inner product of the predicted citation case  $(L_a^c)^T M_p^c$  to the interval  $[0,1]$ . Therefore, the optimization objective function in the citation model is defined as follows:

$$Z = \frac{1}{2} \sum_{(a,p) \in \gamma^c} O_a^c I_p^c (g((L_a^c)^T M_p^c) - q_{ap}^c)^2 + \frac{1}{2} \lambda_c \sum_{(e,a) \in \mathfrak{E}^c} w_{ea}^c (g((U_e^c)^T L_a^c) - w_{ea}^c)^2 + \alpha(L, M, U) \quad (12)$$

$$\alpha(L, M, U) = \frac{\lambda_1}{2} \|L^c\|_F^2 + \frac{\lambda_2}{2} \|M^c\|_F^2 + \frac{\lambda_3}{2} \|U^c\|_F^2 \quad (13)$$

The above objective function can be minimized by performing the following gradient descent for all authors and papers  $M_p^c, L_a^c, U_e^c$ . Parameter  $\lambda_c$  controls the influence proportion between the times of citations and the reference relationship in the training model.

## 5 Experiments

All experiments are performed on a computer with an Intel (R) Core (TM) i5-6402P CPU, 2.80 GHz, and 8 GB RAM. The Operation System is Windows 10 Professional, using MATLAB2017 data processing. We perform experiments on DBLP dataset to answer the following research questions:

- **RQ1:** How does the performance of our proposed ADRCR compare with state-of-the-art recommended methods designed to learn from a large number of recommended scientific papers?
- **RQ2:** Can ADRCR help to solve data sparsity problem?

In what follows, we will first describe the experimental settings and then answer the above two questions.

### 5.1 Experimental Settings

In order to verify the effectiveness of the proposed method, the experiment is designed and verified.

**Dataset.** In order to verify the effectiveness of the proposed method, 8,301 papers published by journalists in the field of data mining (DM) in journals (DMKD, TKDE) or conferences (KDD, ICDM, SDM) are selected.

**Evaluation Metrics.** two of the most classic evaluation metrics [18] are used in our experiments: mean absolute error (MAE) and root mean square error (RMSE). We perform 10-fold cross-validation. In each fold, 20% of the data set is randomly selected as the test set, and the remaining 80% is used as the training set.

**Comparison Methods and Parameter Setting.** To compare and evaluate the performance of our proposed methods, we chose the following three representative methods as competitors. We set the best parameters according to the corresponding references or based on our experiments.

- IBCF [6] and UBCF [7]: These methods are chosen to consider only unilateral item or user information.
- TrustMF [12]: It considers the dual role of users and the social trust network between the same users to improve the performance of the recommendation system, but it does not think over the information of the item.

The parameters  $\beta, \lambda_c, \lambda, \lambda_1, \lambda_2, \lambda_3$  and the dimension  $k$  of the implicit feature vector will affect the recommended performance. By performing many experiments, it is found that the performance of this method is the best when the parameters  $\beta=1/2, \lambda_c=1, \lambda=0.176, \lambda_1=\lambda_2=\lambda_3=0.001$ , and  $k=10$ .

## 5.2 Experimental Results and Analysis

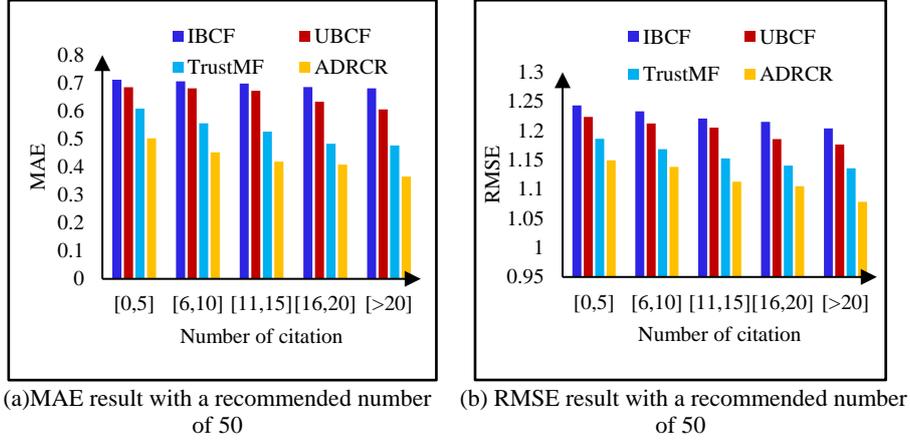
**Performance Analysis(RQ1).** We first compare our method with the collaborative filtering recommendation performance of state-of-the-art methods. Then, we study performance when the recommended number  $N$  is set to [10, 20, 50, 80]. Please note that for a author, our evaluation measure will rank all the unobserved papers in the training set. In this case, a smaller value of  $N$  will make the result more unstable. Therefore, we report relatively large results. The experimental results are recorded in Table 1, where  $k$  is the dimension of the feature space. The main observations of this experimental study are summarized in Table 1.

**Table 1.** Performance comparisons on DBLP dataset

$N$		10		20		50		80	
	Methods	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
$k=10$	IBCF	0.951	1.235	0.942	1.223	0.922	1.211	0.904	1.205
	UBCF	0.942	1.217	0.933	1.215	0.917	1.205	0.902	1.203
	TrustMF	0.832	1.072	0.821	1.079	<b>0.817</b>	1.062	0.806	1.059
	ADRCR	<b>0.815</b>	<b>1.051</b>	<b>0.811</b>	<b>1.047</b>	0.818	<b>1.034</b>	<b>0.798</b>	<b>1.023</b>

We can observe that in terms of MAE and RMSE ,the proposed ADRCR method always outperforms other state-of-the-art methods. In particular, compared with the best method TrustMF, When the recommended number  $N$  is 20, 50, 80, and  $k=10$ , on RMSE, the performance of the ADRCR method is improved by 2.1%, 2.8% and 3.6%, respectively. In addition, as the number of recommended papers increases, so does the recommendation efficiency. This shows that an accurate dual role and an approach that takes into account both users and items information modeling can improve recommendation performance.

**Impact of Data Sparsity (RQ2).** The problem of sparsity usually limits the expression of recommendation systems, since some papers are rarely cited by authors. Therefore, we investigate how our proposed ADRCR model can improve the recommendation performance of the paper with few citations. Specifically, we divide all authors into groups based on the number of citation records: [0-5, 6-10, 11-15, 16-20,>20]. In each group, the number of authors ranges from 100 to 200, which can eliminate the randomness of the experimental results. For every group, we compare the performance of our method with the benchmark methods. The results are shown in Fig.2.



**Fig. 2.** Performance of IBCF, UBCF, TrustMF and ADRCR on authors with different number of citation records

It can be seen from the results that when the author cites becomes sparse, the proposed ADRCR performance is better than other methods. Especially for RMSE, the performance of ADRCR is improved by 7.07% from the fifth group to the first group, while the performance of UBCF and TrustMF is improved by 4.70% and 5.06%, respectively. It is found that as the data becomes sparse, the performance gap between ADRCR and other methods becomes more apparent. Because the ADRCR model considers both authors and papers information, it can achieve good recommendation performance for authors with sparse citations.

## 6 Conclusions

The traditional method of recommending scientific papers does not think over the information of authors and papers simultaneously. To this end, we propose a recommendation method that considers the authority of authors and papers concurrently. Through the author's citations of scientific papers, find the scientific papers that the author is interested in, and recommend them to a large number of scientific papers. The experimental results show that compares with other traditional recommendation methods, the proposed method has achieved good recommendation results on both evaluation indicators. Especially in the topic of scientific papers, clustering is divided into different class by offline clustering. This not only enhances the recommendation speed, but also improves the recommendation efficiency.

**Acknowledgement.** This work is supported by the National Natural Science Foundation of China (61762078, 61363058, 61966004), Major project of young teachers' scientific research ability promotion plan (NWNLU-LKQN2019-2) and Research Fund of Guangxi Key Lab of Multi-source Information Mining and Security (MIMS18-08).

## References

1. Li, Z., Zhao, H., Liu, Q., Huang, Z., Mei, T., Chen, E.: Learning from History and Present: Next-item Recommendation via Discriminatively Exploiting User Behaviors. In: 24th International Conference on SIGKDD, pp. 1734-1743. ACM, London (2018).
2. Joseph, K., Jiang, H.: Content based News Recommendation via Shortest Entity Distance over Knowledge Graphs. In: the World Wide Web Conference, pp. 690-699. ACM, San Francisco (2019).
3. Huang, L., Ma, H., Li, N., Yu, L.: Collaborative filtering recommendation algorithm based on bipartite graph and joint clustering. *Computer Engineering and Science* 41(11), 2040–2047(2019).
4. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: *Recommender Systems: An Introduction*. Cambridge University Press(2010).
5. Li, H., Ma, X.P., Shi, J.: Incorporating Trust Relation with PMF to Enhance Social Network Recommendation Performance. *International Journal of Pattern Recognition and Artificial Intelligence* 30(06), 1659016(2016).
6. West, J.D., Wesley, S.I., Bergstrom, C.T.: A Recommendation System Based on Hierarchical Clustering of an Article-Level Citation Network. *IEEE Transactions on Big Data*, 2(2), 113-123(2017).
7. Dai, T., Li, Z., Cai, X., Pan, S., Yuan, S.: Explore semantic topics and author communities for citation recommendation in bipartite bibliographic network. *Journal of Ambient Intelligence & Humanized Computing* 9(9), 1-19(2017).
8. Wang, H., Shi, X., Yeung, D.Y.: Relational Stacked Denoising Autoencoder for Tag Recommendation. In: Twenty-ninth Aaai Conference on Artificial Intelligence, pp. 3052-3058. AAAI Press, Texas (2015).
9. Li, P., Tuzhilin, A.: DDTCDR: Deep Dual Transfer Cross Domain Recommendation. In: Proceedings of the 13th International Conference on Web Search and Data Mining, pp. 331-339. ACM, Houston (2020).
10. Nassar, N., Jafar, A., Rahhal, Y.: A novel deep multi-criteria collaborative filtering model for recommendation system. *Knowledge-Based Systems* 187, 104811(2020).
11. Tang, J., Hu, X., Gao, H., Liu, H.: Exploiting local and global social context for recommendation. In: Twenty-Third international joint conference on Artificial Intelligence, pp. 2712–2718. AAAI Press, Beijing(2013).
12. Yang, B., Lei, Y., Liu, J., Li, W.: Social collaborative filtering by trust. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(8), 1633–1647(2017).
13. Liu, H., Wang, Y., Peng, Q., Wu, F., Gan, L., Pan, L., Jiao, P.: Hybrid neural recommendation with joint deep representation learning of ratings and reviews. *Neurocomputing* 374, 77-85 (2020).
14. Pan, Y., He, F., Yu, H.: A correlative denoising autoencoder to model social influence for top-n recommender system. *Frontiers of Computer Science* 14(3), 143301 (2020).
15. Ma, F., Wu, Y.S.: Citation Identity—— A Noticeable Concept. *Library and information work*, 53(16), 27-115(2009).
16. Liu, Q., Xiang, B., Yuan, N.J., Chen, E., Xiong, H., Zheng, Y., Yang, Y.: An Influence Propagation View of PageRank. *ACM Transactions on Knowledge Discovery from Data* 11(3), 1-30(2017).
17. Shao, Z., Liu, S., Zhao, Y.: Correction to: Identifying influential nodes in complex networks based on Neighbours and edges. *Peer-to-Peer Networking and Applications* 12(6), 1538-1538(2019).

18. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, L.: Recommender systems with social regularization. In: the Forth International Conference on Web Search and Web Data Mining, pp. 287–296, ACM, Hong Kong(2011).