Embedding Based Personalized New Paper Recommendation

Yi Xie[1,2], Shaoqing Wang[2], Wei Pan[1], Huaibin Tang[3], and Yuqing Sun[1,4](✉)

[1] School of Software, Shandong University, Jinan, China
heilongjiangxieyi@163.com, 530834115@qq.com, sun_yuqing@sdu.edu.cn

[2] School of Computer Science and Technology, Shandong University, Jinan, China
492358046@qq.com

[3] School of Microelectronics, Shandong University, Jinan, China
tanghuaibin@sdu.edu.cn

[4] Engineering Research Center of Digital Media Technology, Shandong University, Jinan, China

Abstract. It is meaningful for researchers to find the interested and high quality new papers. We propose the Joint Text and Influence Embedding recommendation model (JTIE) to consider both the paper quality and the content correlation. We train a paper embedding based on its core elements: contents, authors and publication venues. The quality of a new paper is evaluated based on the author authority and the venue reputation. The citation relationships between papers are considered asymmetric such that they can reflect the user's consideration on the intrinsic influence of a paper. We learn user interests by one's historical references or a set of query keywords. Finally, papers are recommended according to the relatedness between user interests and paper embeddings. We perform experiments against three real-world datasets. The results show that our model outperforms baseline methods on both the personalized recommendation and the query keywords based retrieval.

Keywords: Academic paper · Recommendation · Embedding

1 Introduction

There is a large number of academic papers published every year. To continuously have creative ideas, researchers are interested in the state-of-the-art theory and technology. It is important to find the high quality new papers. Personalized paper recommendation has become a major technique for helping researchers handle huge amounts of papers. To improve user experience, it is essential that the recommendation model predicts users' preferences on papers and provides explainable results. To satisfy the above requirements, there are some challenges, such as the assessment of a new paper quality and the inherent correlation between a user interest and the paper content.

Y. Sun et al. (Eds.): ChineseCSCW 2020, CCIS 1330, pp. 558–570, 2021.
https://doi.org/10.1007/978-981-16-2540-4_40

The recent attention on explainability has lead to the development of a series of explainable recommendation models. A fundamental question explainable recommendation aims to answer is how we balance accuracy and explainability. The purpose of this paper is to illustrate how to effectively optimize accuracy and explainability in a joint and unified framework. The key idea is that fully exploiting the correlations between the recommendation task and the explanation task potentially enables both tasks to be better off than when they are considered separately.

To evaluate a paper quality, most of existing works consider paper citations, including the citation amount and the quality of them [17,24,29]. The more citation a paper receives, the better quality it has. However, such metric is not applicable for a new paper quality evaluation, because there is less citation. Many works based on their publication texts and calculate the content correlation between the profiles and papers [12,14,23]. Generally, representative publications and references well illustrate one's research interests. But these content-based methods neither model the user's preference on paper quality nor consider the inherent semantic connection between user interests and papers for recommendation.

To solve the problems, we propose a Joint Text and Influence Embedding (JTIE) based recommendation method for providing related new papers. Given an academic corpus, which includes paper contents and citation relationships, our model learns the latent influence of new papers and recommends the most related papers to researchers. Based on the contents and the citation relationships, we model the semantics and influence of the elements related to papers, such as the authors and the venues. We embedding paper contents, authors and venues as vectors in the same latent space. Then the representation of a paper is fused on the basis of the former elements. Paper citation reflects the preferences of authors's consideration on the intrinsic influence of a paper, which consists the author authority and the venue reputation. The idea behind is that a paper and its references always share some common features, the embedding of them should be close in the latent space. We model the influence as the asymmetric probabilistic propagation based on citation relationships.

When recommending new papers to a user, we first consider the user interests. Research interests could be learned from user published papers and references. We utilize these information to predict his interests and judge if he could be interested in a new paper. Since a user always follow the academic works or authors in a specific area, we profile a user interest from the papers he ever cited, and recommend the most similar new papers to him. We also consider a general query requirement for a researcher in the form of query words. We perform experiments on three real-world datasets and the results illustrate that our method performs better than the existing methods.

The rest of this paper is organized as follows. First, we introduce related works about embedding and paper recommendation systems in Sect. 2. Then we present definitions and propose our models in Sect. 3, detail our recommendation strategies in Sect. 4, and show our experiments and evaluations in Sect. 5. Finally, we draw our conclusions in Sect. 6.

2 Related Work

Content-Based Embedding. Content-based Embedding is high related with our work, such as word embedding models, which are very helpful means to embed items into lower-dimension latent vectors seeing that the relationships between neighbor nodes remain close in the latent space, such as ISOMAP [28] and Laplacian eigenmap [3]. Word2vec is a method of word embedding, of which the most frequently mentioned two models are CBOW and Skip-gram. Proposed by Mikolov et al. [18], the vectors learned by these models have been utilized to find semantically similar words and proved their great performance.

Most of the current methods focus on the relatedness between a user interest and paper contents. A widely adopted method with paper recommendation is Content Based Filtering (CBF). CBF calculates similarity among items and recommends similar items to target users. Chakraborty et al. [4] propose a diversified citation recommendation system for scientific queries, which considers the semantically correlated articles. Sugiyama et al. proposed to use TF-IDF values of keywords as elements in a paper's vector and represent an author with his publications and citations [24]. However, this method does not consider the semantic ambiguity on user chosen words. However, these paper recommendation methods based on explicit factors of similar items or users without learning various kinds of relationships in the Citation Network.

Structure-Based Embedding. Many network embedding methods were inspired by word embedding. LINE [26] and DeepWalk [20] are network embedding models which be can applied to large-scale networks. They are designed to preserve the proximity between vertices and keep the structure of the network. Supervised PTE were then proposed as an expanding method for the heterogeneous networks but it is not specific for a bibliometric task [25]. Embedding has also been used in research of author identification [5]. However, they only consider homogeneous networks where all the nodes are of the same type and cannot scale to a heterogeneous network containing various kinds of nodes and relations such as author-write-paper and author-published in-venue. To recommend related high-quality works to researchers, some works use network structure and relationships between items and users to promote the accuracy of recommendation. Network-based recommender systems [2,9,10] are also related closely with our work. However, their network refers to social network, which cannot be widely applied to networks that do not contain such explicit social relationships, such as "following" or "mutual following" relationships. Besides, they only leverage the network structure without considering the semantics of the textual items such as papers or books.

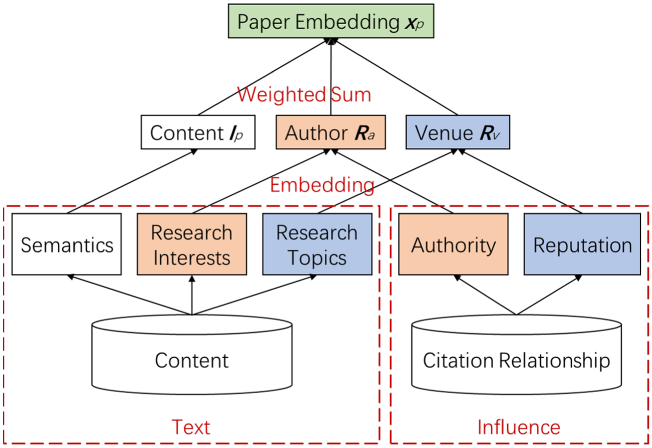

Fig. 1. The joint text and influence embedding model

Paper Recommendation. Another related personalised paper recommendation method is CF based systems. Sugiyama et al. [23] propose a comprehensive evaluation of scholarly paper recommendation by an adaptive neighbor selection method in a collaborative filtering framework. Matrix Factorization (MF) is a model-based Collaborative Filtering and was used in recommender system in previous works [7,14,16]. It considers the latent semantics of users and items. For example, Zhang et al. [32] use a low-dimensional linear model to describe the user rating matrix in a recommendation system. They present a hybrid approach, which learns an additive linear combination of canonical to represent each user's rating profile. However, the dataset in this problem contains millions of authors and papers. Even if we extract hundreds of authors amongst them, the matrix could be extremely sparse. Besides, MF is a CF method so it also could not solve new paper recommendation problem.

To solve the above problem, some works explore the use of deep neural networks to learn the interaction function from data. However, there is little related work on employing deep neural networks for recommendation in contrast to the vast amount of models on MF methods. He et al. [11] present a Neural network-based Collaborative Filtering architecture to model latent features of users and items in recommendation task. Yates et al. [30] adopt the PageRank method to quantify the authority of an author by the citation network. Xie et al. [29] consider the new paper recommendation problem by learning the inherent relativity of word usages from an academic dataset, and computing the potential influence of a new paper by contents, publication venue and author reputation. There are some other academic recommendation, which also seem related with our work, such as experts ranking [1,19,21], partner recommendation [6,31], hot research topic prediction [22] and etc.

3 Joint Text and Influence Embedding

We propose the JTIE method to generate the embedding of papers, authors and venues. Given a corpus, we learn both semantics and influence from the contents and the citation relationship. Then the deep paper semantics and the latent features of the elements related to the paper quality, such as the authors and the venue, are further learned. We model the representation vector of a paper on the basis of its content, authors and venue The diagram is illustrated in Fig. 1 in a down-top view.

3.1 Item Embedding

In this section, we present the notions that would be used in the following discussion, and learn the embedding of papers, authors and venues in a latent space.

For a given academic corpus, we construct a citation network $G = (P, E)$, where $=P$ denotes the paper set, $E = \{(p, r)|p, r \in P,\ p\ \textit{cites}\ r\}$ denotes the citation relationship between papers. We definite the keyword set, the author set and the venue set of the academic corpus as K, A and V, respectively. Each paper p is accompanied with an attribute set (K_p, a_p, v_p), in which $K_p \subseteq K$ denotes the set of keywords mentioned in p, a_p means the author of p and v_p means the publication venue of p.

At first, we model the representation of a paper based on the learned embedding of its elements. The contents and the quality are two main concerns when users refer papers, thus these are key points when we model the representation vector for a paper. With respect to the paper contents, keywords are direct features in depicting paper contents, which are characteristic words extracted from the paper. Let $\mathbf{I}_p$ represents content vector of p, which is the mean value of all keyword vectors extracted from this paper. $\mathbf{I}_p$ is calculated as $\mathbf{I}_p = \frac{\sum_{k_i \in K_p} \mathbf{I}_{k_i}}{|K_p|}$, where $\mathbf{I}_{k_i}$ denote the vector representation of keyword k_i.

With respect to the paper quality, we consider both the authority of the author and the reputation of the venue. The author of a paper reflects its authority and its research directions. A paper is more likely to be a good work if its author has published many high citation papers. And an author usually focus on a few research directions, which will reflect on his papers. Also, the venue of a paper can reflect both its reputation and research topics. If a paper is published in top conference, it tends to be more influential in a certain research field. Besides, each conference concentrates on the topics within a specific direction, so papers published in similar venues share similar research contents. Let $\mathbf{x}_p$ denotes the representation vector of p, which is calculated by a weighted sum of its content vector, author embedding and venue embedding: $\mathbf{x}_p = \alpha\mathbf{I}_p + \beta\mathbf{R}_{a_p} + \gamma\mathbf{R}_{v_p}$. $\mathbf{R}_{a_p}$ and $\mathbf{R}_{v_p}$ denote the author and the venue embedding vector for paper p, respectively. $\alpha, \beta, \gamma \in [0, 1]$ are parameter for balancing the weight of keywords, author and venue, $\alpha + \beta + \gamma = 1$.

Based on the paper embedding, we consider the asymmetric relationships between papers. We define a relevance score between a paper and its reference, which is expected to be high if there is citation relationship between them. Let pair (p, r) means two papers $p, r \in P$. The relevance score $g(p, r)$ denotes the extent of p citing r $g(p, r)$ should reflect three features: the semantic relatedness between two papers, the research interest of a_p and v_p, the influence of a_r and v_r. We adapt the dot product of $\mathbf{x}_p$ and $\mathbf{x}_r$ to model their semantic relevance score. The reference relationship is unidirectional and asymmetric, so we add local bias vectors $\mathbf{M_1}$ and $\mathbf{M_2}$, which shift $\mathbf{x}_p$ and $\mathbf{x}_r$ due to their different semantics in the path. Then the relevance score between the two papers is defined as

$$g(p, r) = \mathbf{I}_p^T \mathbf{I}_r + \mathbf{M_1}^T \mathbf{x}_p + \mathbf{M_2}^T \mathbf{x}_r \tag{1}$$

We adopt soft-max function to model the probability that there exists citation relationship between pair (p, r), which is denoted as $Pr(p|r)$.

$$Pr(p|r) = \frac{\exp(g(p, r))}{\sum_{p' \in P} \exp(g(p', r))} \tag{2}$$

3.2 Objective Function

After embedding the items into a latent space, we formulate the loss function as

$$\mathbb{L} = -\sum_{(p,r) \in G} \log Pr(p|r) \tag{3}$$

Then we have the objective function as follow:

$$\arg\min_{\theta} \mathbb{L} = \arg\min_{\theta} -\sum_{(p,r) \in G} \log \frac{\exp(g(p, r))}{\sum_{p' \in P} \exp(g(p', r))} \tag{4}$$

where θ is the parameter to be learned. Since the time cost to re-calculate the normalization part in Eq. 2 is intolerable, we adopt negative sampling as in NCE [8] which solves the problem of estimation of un-normalized data. Let $Pr^+(p|r) = Pr(p|r)$ denotes the Probability Distribution Function(PDF) of every positive sample $(p, r) \in E$, $Pr^-(p|r)$ denotes the artificial noise distribution of each negative sample $(p, r) \notin E$, and G' denote the set of all negative samples. The loss function can be re-written as follows

$$\mathbb{L} = -\sum_{(p,r) \in G} \log \sigma(\log Pr^+(p|r)) - \sum_{(p',r') \in G'} \log(1 - \sigma(\log Pr^-(p'|r'))) \tag{5}$$

Stochastic gradient descent method is adopted to update the parameters $\mathbf{R}_a$, $\mathbf{R}_v$, $\mathbf{M_1}$ and $\mathbf{M_2}$. There might be millions of pairs in the form of (p, r), so storage for all these pairs is intolerable. Sampling must be applied to solve this problem. A straightforward and efficient way of sampling is to sample nodes following the sequence of the tuple, i.e., to r, p, and a one by one. Papers with more citations are more likely to be sampled as the first node of the sampled tuple, i.e., r in our instance. So we pre-compute the citations of all papers and then sample r according to distributions of numbers of their citations. Once r has been sampled, positive sampling and negative sampling are performed next.

4 Personalized New Paper Recommendation

4.1 Personalized Recommendation Based on User Interests

We recommend related new papers to users based on their research interests. Deep research interests could be learned from user publications and references. We utilize these information to predict their interests and judge if they could be interested in the new paper. Let U denote the set of users who have publications. Each user in $a \in U$ has a reference set RE_a, i.e., all papers he referenced before the publish time of the new paper. The representation of user interest is calculated by the mean of reference vectors as $\mu_a = \frac{\sum_{r \in RE_a} \mathbf{x_r}}{|RE_a|}$, where $\mathbf{x_r}$ is representation vector of a paper $r \in RE_a$. With the above metric, we calculate the similarity between new papers and researcher's interests. Given a researcher, we recommend the most similar new papers to him.

4.2 Personalized Recommendation Based on Keyword Query

We consider a general query requirement for recommendation. Users could provide a set of query keywords, denoted by $Q \subset K$. We solve the problem in two aspects: contents relevance and venue reputation. We model the user query as the mean of keyword representation vector, which is calculated as $\mu_q = \frac{\sum_{k \in Q} \mathbf{I_k}}{|Q|}$, where $|Q|$ denotes the number of keywords in Q.

Then we calculate the similarity between μ_q and the content vectors of new papers, choose the most related ones from the user as candidates. Besides user interests, it is better to recommend papers with high reputation and trusts, so we choose top k papers published in top conferences from the candidate set as our results. This recommending method could also be applied to new researchers who have not published any paper yet.

5 Experiments and Analysis

5.1 Datasets

We adopt three datasets to verify our model. The first dataset is the citation network of DBLP downloaded from AMiner [27]. It contains information about title, abstract, venue, author, year of publication, reference information about a paper. The second dataset is crawled from *Scopus* website dataset[1]. Each paper contains a title, authors, an abstract, keywords, citation information and discipline labels.

Another dataset is the internal patent database released by United States Patent and Trademark Office (PT for short)[2], which includes the information on all published patents at the PT. Each patent contains the ownership (referred to as authors in this paper), mark characteristics, classification, prosecution events,

[1] https://www.scopus.com/.
[2] https://bulkdata.USPTO.gov.

Table 1. Dataset statistics

Dataset	#papers/patents	#authors	#keywords	#venues
AMiner	3056388	1752401	354693	11397
Scopus	1304907	482602	127630	7653
PT	182260	73974	-	-

references, renewal and maintenance history. The patent dataset does not contain the venues and keywords information, so we consider the impact of the authors.

We adopt the Seg-phrase model [15] to extract the keywords, which iterates the process of phrase segmentation and keyword extraction to obtain the most representative phrases. We combine the phrases (two or more words) extracted by Seg-phrase and unigrams with relatively higher tf-idf values, and take this as the keyword set. Stopwords are removed from the keyword set.

After cleaning, the statistics of these datasets are shown in Table 1. As a matter of convenience, papers and patents are named collectively as papers.

5.2 Baseline Methods and Metrics

We compare our model with the following methods.

- *WNMF-5* [32]: This is the weighted nonnegative matrix decomposition method, where each entry in Matrix is set 1 if the researcher has cited the paper, and 0 otherwise. The feature number is set to 5.
- *WNMF-10*: The same method with the above with the feature number being 10.
- *NBCF* [23]: A neighborhood-based Collaborative Filtering algorithm, which is a comprehensive evaluation of scholarly paper recommendation by an adaptive neighbor selection method in a collaborative filtering framework.
- *MLP* [11]: It uses Multi-Layer Perceptron to learn the non-linear interaction function of embeddings from data.

All the parameters of baseline methods are empirically set to the optimal values.

For each user, we prepare k candidate papers/patents. Each candidate set contains 1 really cited paper/patent at least. The candidate papers are ranked according to the relatedness between user interests and paper vectors. We dopt $nDCG@k$ as a measurement [13]. $nDCG@k$ is often used to measure effectiveness of web search engine algorithms. $nDCG@k$ is calculated as $nDCG@k = \frac{DCG@k}{IDCG}$, $DCG@k = \sum_{i=1}^{k} \frac{rel_i}{log_2(i+1)}$, where $rel_i = 5$ if the i-th paper is really cited by the researcher, otherwise $rel_i = 0$. $IDCG = \sum_{i=1}^{|Cite|} \frac{5}{log_2(i+1)}$ is the ideal discounted cumulative gain, in which $|Cite|$ means the number of papers that are really cited by the researcher among the candidate papers.

Table 2. The experiments on the personalised new paper recommendation

nDCG@k	Aminer			Scopus			PT		
	k = 20	k = 30	k = 50	k = 20	k = 30	k = 50	k = 20	k = 30	k = 50
WNMF-5	0.7914	0.7852	0.7256	0.7797	0.7686	0.6819	0.6797	0.6586	0.5419
WNMF-10	0.8265	0.7892	0.7316	0.7889	0.7725	0.7052	0.6789	0.6625	0.5652
NBCF	0.8331	0.7994	0.7322	0.7932	0.7856	0.7272	0.6932	0.6756	0.6272
MLP	0.8391	0.8011	0.7649	0.8263	0.8201	0.7305	0.7063	0.6801	0.6504
Our Method	**0.8693**	**0.8512**	**0.8053**	**0.8309**	**0.8257**	**0.7399**	**0.7309**	**0.7057**	**0.6899**

Table 3. Query based recommendation compared with MLP

User Query	Our method	MLP
J. Han Information Network Knowledge Discoverry Data Mining	**1.** DeepTutor: An Effective, Online Intelligent Tutoring System That Promotes Deep Learning **2.** Quantifying Robustness of Trust Systems against Collusive Unfair Rating Attacks Using Information Theory **3.** Deep Learning Architecture with Dynamically Programmed Layers for Brain Connectome Prediction **4.** Evaluating the statistical significance of biclusters **5.** A Gaussian Process Latent Variable Model for BRDF Inference	**1.** Convergence properties of general network selection games **2.** Quantifying the Targeting Performance Benefit of Electrostatic Haptic Feedback on Touchscreens **3.** The Role of Environmental Predicta-bilityand Costs in Relying on Automation **4.** Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection **5.** When hybrid cloud meets flash crowd: Towards cost-effective service provisioning
D. Patterson High Performance Computing Computer Architecture Parallel	**1.**Markov Mixed Membership Models **2.**The effect of head mounted display weight and locomotion method on the perceived naturalness of virtual walking speeds **3.**Influence at Scale: Distributed Computation of Complex Contagion in Networks	**1.**Convergence properties of general network selection games **2.**Quantifying the Targeting Performance Benefit of Electrostatic Haptic Feedback on Touchscreens **3.**The Role of Environmental Predicta-bilityand Costs in Relying on Automation

5.3 Evaluation on Personalized New Paper Recommendation

Recommendation Based on User Publications and References. To evaluate the performance of our recommendation, we compare our model with other methods on new research papers in this subsection. We randomly select 300, 100 and 50 researchers in Aminer, Scopus and PT datasets, respectively, to verify the performance of our model. The authors must satisfy the following conditions:

– Have published at least 5 papers and have cited at least 5 papers before year Y;
– Have cited at least 1 paper after year Y;
– The above academic papers should contain titles, authors, venues and abstracts, while the patents should contain titles, abstracts and authors.

The dataset is separated into two parts according to their published years. Papers published before year Y are used for training and after year Y are used

Table 4. The comparison between methods with different number of representative papers (#rp)

Average nDCG	Aminer			Scopus		
	#rp = 1	#rp = 3	#rp = 5	#rp = 1	#rp = 3	#rp = 5
WNMF-5	0.664	0.754	0.770	0.633	0.706	0.765
WNMF-10	0.680	0.760	0.790	0.659	0.715	0.761
NBCF	0.690	0.769	0.821	0.676	0.721	0.782
MLP	0.759	0.853	0.871	0.681	0.747	0.805
Our Method	**0.771**	**0.861**	**0.874**	**0.706**	**0.752**	**0.828**

Table 5. The comparison between methods with different ratios between samples (positive samples: negative samples)

Average nDCG	Aminer			Scopus		
	10:1	1:1	1:10	10:1	1:1	1:10
WNMF-5	0.679	0.754	0.730	0.633	0.726	0.693
WNMF-10	0.702	0.761	0.753	0.659	0.732	0.706
NBCF	0.720	0.775	0.753	0.676	0.742	0.731
MLP	0.743	0.821	0.770	0.680	0.775	0.734
Our Method	**0.791**	**0.869**	**0.780**	**0.708**	**0.801**	**0.754**

for testing. Then for each user, we prepare k candidate papers, each candidate paper set must satisfy the following conditions:

- Each candidate set contains 1 really cited paper at least;
- All candidates should be published after year Y, the papers contain titles, authors, venues and abstracts, while the patents should contain titles, abstracts and authors;
- The authors of these candidates should have been embedded during the training process;
- The candidates should be textually similar with the papers that cited by the user before year Y. The cosine similarity is adopted to calculate the semantic similarity between candidates and really cited papers.

Year Y is set 2014, then we rank the candidate papers for new paper recommendation.

The evaluation results on the three datasets are shown in Table 2. We can see that our method is very helpful in improving the performance of new paper recommendation.

Recommendation Based on Keyword Query. Our method could recommend for users who provide their queries. However, we do not have comparing metrics in aspect of contents relevance and we do not have their publication information so we cannot examine through their reference. Therefore, we test if our search results of some queries could retrieve content-relevant papers with higher reputation. We rank all the papers in a given area and see if papers retrieved by our method are more relevant to the query and querist. The results in Table 3 shows that our recommendations are more likely to the querists.

Evaluation on Parameters Settings. To quantify the influence by different parameter settings, we perform the above experiments with different settings on Aminer and Scopus. As the results shown in Table 4, the performance of our proposed personalized cross domain paper recommendation increases with a increasing number of papers. It is easy to understand that our recommendation method better grasp an author's requirement when she has more publications. Table 5 shows that our model performs best when the ratio between positive and negative samples is 1 : 1.

6 Conclusion

In this paper, we propose the Joint Text and Influence Embedding method for personalised recommendation on new papers. A paper is represented by the vectors of content, authors and published venue in the same latent space and the objective function is designed for the consistency of content semantics and paper influences in the citation network. Then we adopt the stochastic gradient descent method for optimization. A new paper is evaluated by its contents, authors' authority and the reputation of the publication venue. A user interest is learned either by one's historical references or a set of query keywords. Then we recommend the top-k related new papers. The results of the experiments show our method outperforms other methods.

Acknowledgement. This work was supported by the National Key R&D Program of China (2018YFC0831401), the Key R&D Program of Shandong Province (2019JZZY010107), the National Natural Science Foundation of China (91646119), the Major Project of NSF Shandong Province (ZR2018ZB0420), and the Key R&D Program of Shandong province (2017GGX10114). The scientific calculations in this paper have been done on the HPC Cloud Platform of Shandong University.

References

1. Asiwal, K., Suresh, B.K., Reddy, G.R.M.: Analysis of academic research networks to find collaboration partners. In: Stephanidis, C. (ed.) HCI 2016. CCIS, vol. 618, pp. 8–14. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40542-1_2
2. Bedi, P., Kaur, H., Marwaha, S.: Trust based recommender system for semantic web. IJCAI **7**, 2677–2682 (2007)
3. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. NIPS **14**, 585–591 (2001)

4. Chakraborty, T., Modani, N., Narayanam, R., Nagar, S.: Discern: a diversified citation recommendation system for scientific queries. In: 2015 IEEE 31st International Conference on Data Engineering (ICDE), pp. 555–566. IEEE (2015)
5. Chen, T., Sun, Y.: Task-guided and path-augmented heterogeneous network embedding for author identification (2016)
6. Chikhaoui, B., Chiazzaro, M., Wang, S.: A new granger causal model for influence evolution in dynamic social networks: the case of DBLP. In: AAAI, pp. 51–57 (2015)
7. Gu, Q., Zhou, J., Ding, C.: Collaborative filtering: weighted nonnegative matrix factorization incorporating user and item graphs. In: Proceedings of the 2010 SIAM International Conference on Data Mining, pp. 199–210. SIAM (2010)
8. Gutmann, M.U., Hyvärinen, A.: Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics. J. Mach. Learn. Res. **13**(Feb), 307–361 (2012)
9. Guy, I., et al.: Personalized recommendation of social software items based on social relations. In: Proceedings of the Third ACM Conference on Recommender Systems, pp. 53–60. ACM (2009)
10. He, J., Chu, W.W.: A social network-based recommender system (SNRS). In: Memon, N., Xu, J., Hicks, D., Chen, H. (eds.) Data Mining for Social Network Data. Annals of Information Systems, vol. 12. Springer, Boston (2010). https://doi.org/10.1007/978-1-4419-6287-4_4
11. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S.: Neural collaborative filtering. In: Proceedings of the 26th International Conference on World Wide Web, pp. 173–182. International World Wide Web Conferences Steering Committee (2017)
12. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: Eighth IEEE International Conference on Data Mining. ICDM 2008, pp. 263–272. IEEE (2008)
13. Järvelin, K., Kekäläinen, J.: IR evaluation methods for retrieving highly relevant documents. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 41–48. ACM (2000)
14. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8) (2009)
15. Liu, J., Shang, J., Wang, C., Ren, X., Han, J.: Mining quality phrases from massive text corpora. In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, pp. 1729–1744. ACM (2015)
16. Ma, H., Yang, H., Lyu, M.R., King, I.: Sorec: social recommendation using probabilistic matrix factorization. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, pp. 931–940. ACM (2008)
17. McNee, S.M., et al.: On the recommending of citations for research papers. In: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, pp. 116–125. ACM (2002)
18. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems, pp. 3111–3119 (2013)
19. Moreira, C., Calado, P., Martins, B.: Learning to rank academic experts in the DBLP dataset. Expert. Syst. **32**(4), 477–493 (2015)
20. Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: online learning of social representations. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 701–710. ACM (2014)

21. Priem, J., Piwowar, H., Hemminger, B.: Altmetrics in the wild: Using social media to explore scholarly impact (2012)
22. Song, M., Heo, G., Kim, S.: Analyzing topic evolution in bioinformatics: investigation of dynamics of the field with conference data in DBLP. Scientometrics **101**(1), 397–428 (2014)
23. Sugiyama, K., Kan, M.: A comprehensive evaluation of scholarly paper recommendation using potential citation papers. Int. J. Digit. Libr. **16**(2), 91–109 (2015)
24. Sugiyama, K., Kan, M.Y.: Exploiting potential citation papers in scholarly paper recommendation. In: Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 153–162. ACM (2013)
25. Tang, J., Qu, M., Mei, Q.: PTE: predictive text embedding through large-scale heterogeneous text networks. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1165–1174. ACM (2015)
26. Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q.: Line: large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web, pp. 1067–1077. ACM (2015)
27. Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., Su, Z.: Arnetminer: extraction and mining of academic social networks. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 990–998. ACM (2008)
28. Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
29. Xie, Y., Sun, Y., Shen, L.: Predicating paper influence in academic network. In: 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pp. 539–544. IEEE (2016)
30. Yates, E., Dixon, L.: Pagerank as a method to rank biomedical literature by importance. Source Code Biol. Med. **10**(1), 16 (2015)
31. Zhang, A., Bhardwaj, A., Karger, D.: Confer: a conference recommendation and meetup tool. In: Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion, pp. 118–121. ACM (2016)
32. Zhang, S., Wang, W., Ford, J., Makedon, F.: Learning from incomplete ratings using non-negative matrix factorization. In: Proceedings of the 2006 SIAM International Conference on Data Mining, pp. 549–553. SIAM (2006)