



Social recommendation algorithm based on stochastic gradient matrix decomposition in social network

Tian-wu Zhang¹ · Wei-ping Li^{2,4} · Lu Wang^{3,4} · Jie Yang²

Received: 8 July 2018 / Accepted: 10 December 2018 / Published online: 11 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

The revenue of an e-commerce system is affected directly by the prediction accuracy of recommendation system. Although recommendation systems have been comprehensively analyzed in the past decade, the study of social-based recommendation systems just started. In this paper, aiming at providing a general method for improving recommendation systems by incorporating social network information, we propose a social recommendation algorithm based on stochastic gradient matrix decomposition in social network so as to improve the prediction accuracy. This paper considered the social network as auxiliary information, and proposed a matrix factorization based on social recommendation algorithm, which systematically illustrate how to design a matrix factorization objective function with social regularization. It constructed a matrix with the social network and the user scoring matrix, and proposed a stochastic gradient descent algorithm for matrix factorization. The empirical analysis on two large datasets demonstrates our proposed algorithm has lower prediction error, and is obviously better than other state-of-the-art methods.

Keywords Matrix decomposition · Recommendation system · Social network · Stochastic gradient

1 Introduction

With the explosive growth of internet information, users are in the world with the information overload. Therefore, how to extract the online information that users care about from the massive information is an important and difficult task. As an indispensable type of information filtering technique, recommendation systems have attracted a lot of attention in the past decade. Related recommendation techniques have

been widely studied in research communities of information retrieval machine learning and data mining. Due to their great commercial value, recommendation systems have also been successfully deployed in industry, such as Amazon, iTunes, Netflix, etc.

As an effective tool to mine potential valuable information, the recommendation system has drawn great attention in academia and industry. Currently, many existing online recommendation system applications have gone up formally line, which recommend movie, book and music to users respectively. Social network systems, like last.fm, play a significant role in Web 2.0, containing large amounts of multimedia-enriched data that are enhanced both by explicit user-provided annotations and implicit aggregated feedback describing the personal preferences of each user. It is also a common tendency for these systems to encourage the creation of virtual networks among their users by allowing them to establish bonds of friendship and thus provide a novel and direct medium for the exchange of data. The traditional recommendation system mainly includes users and items, where the item can be movie, book, music and so on. Each user give a score to a subset of the item sets, and each item is scored by a subset of the user sets. If a user's score on an item-set is denoted as a row vector and a score on a user's set

✉ Jie Yang
jiejyang@whut.edu.cn
Wei-ping Li
liweiping@rpc.edu.cn
Lu Wang
wanglu@rpc.edu.cn

¹ School of Computing, Henan University of Engineering, Zhengzhou 451191, Henan, China

² School of Information Engineering, Wuhan University of Technology, Wuhan 430070, Hubei, China

³ Department of Electrical Automation, Shanghai Maritime University, Shanghai 201306, China

⁴ Department of Police Technology, Railway Police College, Zhengzhou 450003, Henan, China

as a column vector, thus the user's score-set for the item is a matrix. The task of the recommendation system is to predict the scores of unknown items based on known elements in the matrix and recommend the items with high predicted scores to the user.

In the past few years, the dramatic expanding of Web 2.0 Web sites and applications poses new challenges for traditional systems. Traditional recommendation systems always ignore social relationships among users. But in our real life, when we are asking our friends for recommendations of nice digital cameras or touching movies, we are actually requesting verbal social recommendations. Social recommendation is a daily occurrence, and we always turn to our friends for recommendations. Hence, in order to improve recommendation systems and to provide more personalized recommendation results, we need to incorporate social network information among users.

With the emergence of social networks, more of websites have introduced social network services to promote the competitive power. In generally, the recommendation systems with social network service will be called as social recommendation system. In the social system, in addition to the user-item scoring matrix, there is a user-user social network that not only can utilize social networks to improve the performance of the score prediction (for example, suggestions from friends in the network tend to be preferable to strangers' suggestions), but also can judge the user's preference on the basis of the score of the user to item, then recommend the friend to the user according to the user's preference.

Recommending suitable materials by using the learning styles of users and the difficulty of the materials is a complex combinatorial problem that needs to take into account many design parameters. Various approaches have been proposed to deal with such combinatorial problems, and in recent years artificial bee colony algorithms have attracted increasing interest. Karaboga suggested that the performance of an artificial bee colony algorithm is better than, or at least similar to, that of a particle swarm optimization algorithm (PSO), genetic algorithm (GA), and evolution strategy (ES). However, these methods have the incentives of the widespread adoption of social networks and of the lack of some previous study that directly addresses the problem of efficiently integrating the added value knowledge provided by those networks in the field of collaborative recommendation.

In this paper, aiming at solving the problems mentioned above, we propose two social recommendation methods that utilize social information to improve the prediction accuracy of traditional recommendation systems. More specifically, the social network information is employed in designing two social regularization terms to constrain the matrix factorization objective function. In order to realize the aim, the social network between users and the scoring matrix of users are merged into a large matrix (user + item). The

matrix decomposition technique is used to decompose the merged matrix into two low-rank matrices, and then the low rank matrix predicts the unknown score in the original matrix. Our proposed approaches are quite general, and they can also be applied to trust-aware recommendation systems. The experimental analysis on two large data-sets shows that our methods outperform other state-of-the-art algorithms.

The rest of our paper is organized as follows: next section presents a brief description for related works. Section 3 shows the general outline of the proposed method, while Sect. 4 describes the experiments performed to evaluate the accuracy of the proposal and the quality of the obtained result. Finally, Sect. 5 discusses the results and proposes some conclusions of the whole dissertation.

2 Related works

In this section, we review several major approaches for recommendation systems, especially for collaborative filtering. Two types of collaborative filtering approaches are widely studied: memory-based and model-based.

The memory-based approaches are the most popular prediction methods and are widely adopted in commercial collaborative filtering systems. The most analyzed examples of memory-based collaborative filtering include user-based approaches and item-based approaches. User-based approaches predict the ratings of active users based on the ratings of similar users found, and item-based approaches predict the ratings of active users based on the computed information of items similar to those chosen by the active user. User-based and item-based approaches often use the Pearson correlation coefficient (PCC) algorithm and the VSS algorithm as the similarity computation methods. PCC-based collaborative filtering generally can achieve higher performance than the other popular algorithm, since it considers the differences of user rating style.

The traditional recommendation system only contains user's score matrix for items. (Wang et al. 2012). Since the collaborative filtering based recommendation algorithm can well apply the preference of other users to the items, so its accuracy is much higher and it is also a commonly used recommendation method (Hailing et al. 2009). The recommendation algorithm based on collaborative filtering can be divided into heuristic collaborative filtering algorithm and model-based collaborative filtering algorithm (Xia et al. 2017).

Heuristic collaborative filtering algorithm is also called as memory-based collaborative filtering algorithm, which searches for similar nearest neighbors in the matrix and scores the evaluated items on the basis of the nearest neighbor's score (Bellogín et al. 2013). In order to search the nearest neighbor, some similarity measure methods, fox

example Pearson correlation coefficient, cosine similarity index, can be used to find the nearest neighbor of the item to be estimated (Huang Chuanguang et al. 2010). Nearest neighbor selection process can be either by the line (looking for similar users), or by column (looking for similar items). The model-based collaborative filtering algorithm builds a model on the basis of the scoring matrix of a known user and applies the model to score the evaluated items. The most common method for model-based collaborative filtering are matrix decomposition, classification algorithm (Park et al. 2012), clustering algorithm (Tsai and Hung 2012), Bayesian network (Hsu et al. 2012) and so on.

We first demonstrate our social recommendation framework using a simple but illustrative toy example. Then we introduce the factor analysis method for social recommendation using probabilistic matrix factorization. In recent years, matrix decomposition techniques such as singular value decomposition and non-negative matrix factorization have been gradually applied to the recommendation system (Zhang et al. 2013; Tu et al. 2013; Zhou et al. 2012). Given a user scoring matrix $R_{n \times m}$, the singular value decomposition is to decompose the matrix $R_{n \times m}$ into $R_{n \times m} = A_{n \times n} \lambda_{n \times m} B_{m \times m}$, where $\lambda_{n \times m}$ is the diagonal matrix of singular values contained by $R_{n \times m}$, and $\lambda_{i,i} \geq \lambda_{j,j} \geq 0, 0 \leq i \leq j \leq \min(n, m)$. At the end of the decomposition process, the matrix $\lambda_{n \times m}$ is truncated into $\lambda_{k \times k}$, which includes only the top k row and the top k column; In addition, $A_{n \times n}$ and $B_{m \times m}$ will also truncated, where the top k column of $A_{n \times n}$ is reserved to get $A_{n \times k}$ and the top k column of $B_{m \times m}$ is reserved to get $B_{k \times m}$. Thus, we can get $R_{n \times m} \doteq A_{n \times k} \lambda_{k \times k} B_{k \times m}$, so the unknown item in $R_{n \times m}$ can be predicted by applying this formula. Non-negative matrix factorization decomposes the matrix $R_{n \times m}$ into $R_{n \times m} = U_{n \times k} V_{k \times m}$, where $k \ll n, m$ represents the number of potential factors for user scoring. Again, the unknown item in $R_{n \times m}$ can be also predicted applying this formula.

As the user relationship accumulates, the social network between users generated in the recommendation system plays a very important role in scoring the unknown items in the scoring matrix. In social networks, the link relation between users represents a trust relationship, and this trust relationship can expand along the network. Although some literature (Ma et al. 2008, 2009) applies the matrix factorization technique combined with the social network to recommend the item, it does not consider the expansion of the trust relationship in the network. TidalTrust (Golbeck et al. 2005) introduced the direct neighbor node of u in the network when calculating the trust value between nodes u and v , and weights the trust value from nodes u to neighbor nodes. The idea of MoleTrust (Massa et al. 2007) is similar to that of TidalTrust, which introduced the neighbor node of v and weights the trust value from u to the neighbor node, where the weight value is the trust value from the trust value from neighbor nodes to v . TrustWalker (Jamali and Ester 2009) divides a user's score of an item into two parts,

where one is a friend's score of similar items and another is a stranger's score for the same item. The weight of the friend's score is based on the similarity of random walks.

3 Our proposed social recommendation algorithm

The problem we study in this paper is different from traditional recommendation systems since the latter normally only considers the user-item rating matrix. In this paper, we will also incorporate users' social network information to improve recommendation systems. Given a user set $U = \{u_1, \dots, u_n\}$, and a item set $I = \{i_1, \dots, i_m\}$, thus a recommendation system contains social networks $G = (U, E)$ and score matrix $R = [r_{u,i}]_{n \times m}$, where G can be denoted as a matrix M_G with the size of $n \times n$:

$$m_{i,j} = \begin{cases} \frac{1}{d_{out}(i)}, (i, j) \in E \\ 0, (i, j) \notin E \end{cases}, \tag{1}$$

where $d_{out}(i)$ is the out-degree of node i in graph G . The matrix R is a sparse matrix, where $r_{u,i} = \text{null}$ denote the item i has not already been scored for the user u .

The matrix M_G and R will be merged into a bigger matrix $M = [M_G \ R]_{n \times (n+m)}$ in this paper, where the top n columns are the social relations matrix between users, the bottom m columns are user's scoring matrix. The purpose of this paper is to decompose the matrix M into two low-rank matrix $A_{n \times k}$ and $B_{k \times (n+m)}$, namely:

$$M = A_{n \times k} \cdot B_{k \times (n+m)}, \tag{2}$$

where $k \ll n, m$ is a predefined constant.

An efficient and effective approach to recommendation systems is to factorize the user-item rating matrix, and utilize the factorized user-specific and item-specific matrices to make further missing data prediction. The premise behind a low-dimensional factor model is that there is only a small number of factors influencing the preferences, and that a user's preference vector is determined by how each factor applies to that user.

After the low-rank matrices $A_{n \times k}$ and $B_{k \times (n+m)}$ are obtained, the following equation can be adopted to predict the unknown score of item i for user u :

$$\hat{m}_{u,i} = \sum_{r=1}^k A_{u \times r} \cdot B_{r \times i}. \tag{3}$$

where $n \leq i \leq n + m$.

Similarly, if the matrix M_G of the social network can be written as:

$$m_{i,j} = \begin{cases} \frac{1}{d_{out}(i)} & (i,j) \in E, \\ \text{null} & (i,j) \notin E. \end{cases} \quad (4)$$

where the low-rank matrices $A_{n \times k}$ and $B_{k \times (n+m)}$ can also predict the link relation between users by using Eq. (3), so as to recommend potential users to the user as friends.

In the matrix decomposition process, the error between the estimated value and the true value of each item in the matrix is denoted as follows:

$$e_{u,i} = m_{u,i} - \hat{m}_{u,i}. \quad (5)$$

In order to get two optimal matrix in the matrix decomposition, it is hoped that the total error of M is the smallest, namely, $\min \sum_{u=1}^n \sum_{i=1}^{n+m} I_{u,i} (e_{u,i})^2$, thus it can be rewritten as follows:

$$\min \sum_{u=1}^n \sum_{i=1}^{n+m} I_{u,i} \left(\sum_{r=1}^k A_{u \times r} \cdot B_{r \times i} - m_{u,i} \right)^2, \quad (6)$$

where I is selection function, and satisfy the following restricted condition:

$$I_{u,i} = \begin{cases} 1 & m_{u,i} \neq \text{null}, \\ 0 & m_{u,i} = \text{null}. \end{cases} \quad (7)$$

In order to solve the overfitting problem for training data, the formula (6) can be modified, which we can get:

$$e_{u,i} = \min \sum_{u=1}^n \sum_{i=1}^{n+m} I_{u,i} \left(\sum_{r=1}^k A_{u \times r} B_{r \times i} - m_{u,i} \right)^2 + \lambda \sum_{r=1}^k ((A_{u \times r})^2 + (B_{r \times i})^2), \quad (8)$$

where λ is the predefined normalized parameter. The optimization problem in Eq. (8) minimizes the sum-of-squared-errors objective function with quadratic regularization terms. Gradient based approaches can be applied to find a local minimum. It also contains a nice probabilistic interpretation with Gaussian observation noise, which is detailed in (Sarwar et al. 2001). The above algorithm is perhaps one of the most popular methods in collaborative filtering.

In order to get the optimal matrices $A_{n \times k}$ and $B_{k \times (n+m)}$ satisfied the Eq. (8), the common methods are the gradient descent method and the least square method. In this paper, the stochastic gradient descent method is used to calculate the decomposed matrices $A_{n \times k}$ and $B_{k \times (n+m)}$ for M .

Given two matrices $A_{n \times k}^{(0)}$ and $B_{k \times (n+m)}^{(0)}$ randomly during initialization, the training samples $m_{u,i}$ can be inputted one by one and $A_{n \times k}$ and $B_{k \times (n+m)}$ can be adjusted according to the error of training sample (Eq. 5). Let us suppose $e_{u,i}$ is the error of sample. $m_{u,i}$, so the u th row A_u in matrix $A_{n \times k}$ and

the i column B_i in matrix $B_{k \times (n+m)}$ can be rewritten and denoted as follows:

$$\begin{aligned} A_u &\leftarrow A_u + \gamma \cdot (e_{u,i} \cdot B_i^T - \lambda \cdot A_u) \\ B_i &\leftarrow B_i + \gamma \cdot (e_{u,i} \cdot A_u^T - \lambda \cdot B_i), \end{aligned} \quad (9)$$

where λ is the predefined normalized parameter from Eq. (8), γ is a predefined step length. After all the training sample data have been trained, $A_{n \times k}^{(1)}$ and $B_{k \times (n+m)}^{(1)}$ can be obtained.

Repeat the training process for k times, we can get $A_{n \times k}^{(k)}$ and $B_{k \times (n+m)}^{(k)}$. If $A_{n \times k}^{(k)}$ and $B_{k \times (n+m)}^{(k)}$ in Eq. (8) are less than a predefined small constant, thus the training process is over, and we get the $A_{n \times k}^{(k)}$ and $B_{k \times (n+m)}^{(k)}$. In other words, the obtained matrix is the matrix model we need for matrix decomposition.

After the $A_{n \times k}^{(k)}$ and $B_{k \times (n+m)}^{(k)}$ are obtained by decomposing the matrix M , we can apply the low-rank matrices $A_{n \times k}^{(k)}$ and $B_{k \times (n+m)}^{(k)}$ obtained from the decomposition to predict the unknown item in the scoring matrix R . First, the matrix $B_{k \times (n+m)}$ is divided into two matrices $C_{k \times n}$ and $D_{k \times m}$, where $C_{k \times n}$ is the top n column of $B_{k \times (n+m)}$, and $D_{k \times m}$ is the bottom m row of $B_{k \times (n+m)}$. The matrices $A_{n \times k}$ and $D_{k \times m}$ can be adopted to predict the unknown item $r_{u,i}$ in the scoring matrix R :

$$\hat{r}_{u,i} = \sum_{r=1}^k A_{u \times r} \cdot D_{r \times i} \dots n \dots \quad (10)$$

When users score to an item, it can be affected by several potential factors. In the proposed matrix decomposition model in this paper, the decomposed matrix parameter k is the number of potential features that affect user scoring. The value k that satisfies Eq. (8) is the k most important features in these factors.

The proposed social recommendation algorithm based on matrix decomposition considers social network as a part of matrix decomposition and can well solve the problem of user information transfer in network. In this paper, we propose that the potential factor that determines the user's score is determined by the eigenvectors of their direct neighbors. The eigenvectors of their neighbors are recursively determined by the neighbors' direct neighbors. In the scoring system of real society, users tend to only evaluate some of the items. If user scoring values are calculated only by the user's eigenvectors of direct neighbor, the amount of available information is very small and does not solve the data sparse problem.

Our proposed model in this paper can recursively apply the eigenvectors of neighbors, so that the available information is very large and can well solve the data sparse problem, which is quite general, and it can be easily extended to incorporate other contextual information, including social tags issued by users, movie genres, user demographic

information, etc. By taking advantages of these information, we can use Cosine similarity of other similarity calculation methods to compute the affinities between users or items. Then we can plug in those similar users or items into our social regularization framework to further performance of recommendation systems.

4 Experiments results analysis

In this section, the performance of our proposed recommendation scheme is tested and analyzed. The main innovation in the framework is the combination of adaptive social network and low-rank constraint. We conduct several experiments to compare the recommendation qualities of our approaches with other state-of-the-art recommendation methods.

4.1 Evaluation criteria

All of the experiments are run under MATLAB v7.8 (R2009a) on PCs with an Intel quad-core i5 CPU at 3.3 GHz and 4 GB memory. In order to verify the accuracy of the proposed algorithm, the adopted evaluation criteria in our paper is root mean squared error (RMSE), which definition is shown as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{m_{u,i} \in T_{\text{test}}} (m_{u,i} - \hat{m}_{u,i})^2}{|T_{\text{test}}|}}, \quad (11)$$

where T_{test} is the set of training samples, and each of these elements meets $m_{u,i} \neq \text{null}$. From the definitions we can see that a smaller RMSE value means a better performance.

In the experiment, the social recommendation algorithm based on matrix factorization proposed in this paper is called as SocialMF. The 3 evaluated recommenders are PearsonCF, BasicMF and TrustWalker algorithms. It is worth noticing that the most challenging data-sets from the existing works are used for evaluation. All parameters in comparative algorithms are fixed for all the experiments to demonstrate the robustness and stability of our proposed recommendation method.

4.2 Data sets

Our proposed models are quite general, and can be utilized to both social recommendation systems and trust-aware recommendation systems.

The data sets Epinions and Flixster (Xia and Wang 2017) are adopted in our experimental, both of which contain a social network and user scoring matrix between users; the Epinions data set is a directed social network, whose score is an integer value between 1 and 5. The Flixster dataset is the user's score

to a movie whose social network is undirected and has a score with an integer value between 1 and 10 divided by 2. At Epinions, visitors can read reviews about a variety of items to help them decide on a purchase or they can join for free and begin writing reviews that may earn them reward and recognition (Agarwal and Chen 2010). To post a review, members need to first rate the product or service on a rating scale from 1 to 5 stars. The dataset we collected from Epinions consists of 51,670 users who have rated a total of 83,509 different items. The total number of ratings is 631,064. As to the user social trust network, the total number of issued trust statements is 511,799. Other statistics of this dataset are respectively shown in Table 1.

4.3 Experimental results

In this section we present a detailed experimental evaluation of the different algorithmic choices for the steps of the stochastic gradient matrix decomposition based recommendation systems and compare its performance to that achieved by traditional association-rule based approaches. Our main goal is to explore the possibilities of combining different sub-tasks to formulate an efficient recommendation algorithm. As the combination of different parameters and tasks is enormous, we experimentally evaluate each parameter by making reasonable choices for the rest.

As discussed in Bedi et al. (2007) and Xia et al. (2018) the number of dimensions is critical for the effectiveness of the low dimensional representation. We are interested in determining the number of dimensions that is large enough to capture all the latent relationships in the matrix yet small enough to avoid over-fitting errors. Unfortunately there is no direct analytical method to determine the value of the optimal number of dimensions so the optimal value has to be experimentally evaluated. Furthermore, the optimal value of the lower dimensional space (i.e., the optimal rank of the customer-product matrix) is different for different data sets. Determination of the optimal value of dimension can be done by using similar technique used to determine the optimal value of nearest neighbors. We performed an experiment where we divided the data set into a train and test portion first and further divide the train data set into a train and validation portion. We repeated the experiment for different number of dimensions and noted the impact on the RMSE metric and from the plot we determined optimum number of dimensions. To show that this approach leads to the correct estimation of the optimal value

Table 1 Statistical information for data-set

Dataset	#User	#Edge	#Item	#Scoring
Epinions	49,290	487,183	139,738	664,824
Flixster	786,936	7,058,819	48,794	8,196,077

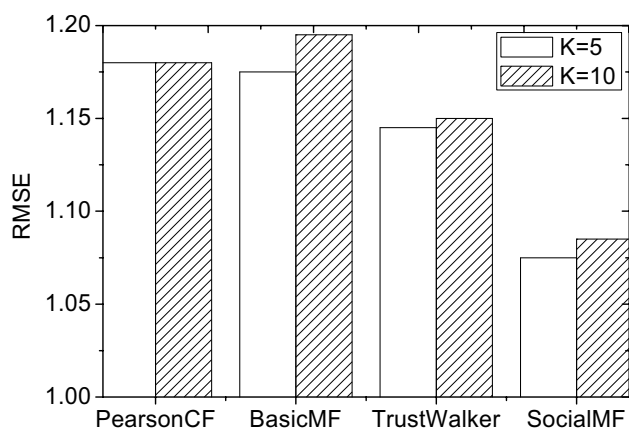


Fig. 1 Comparison accuracy for Epinions

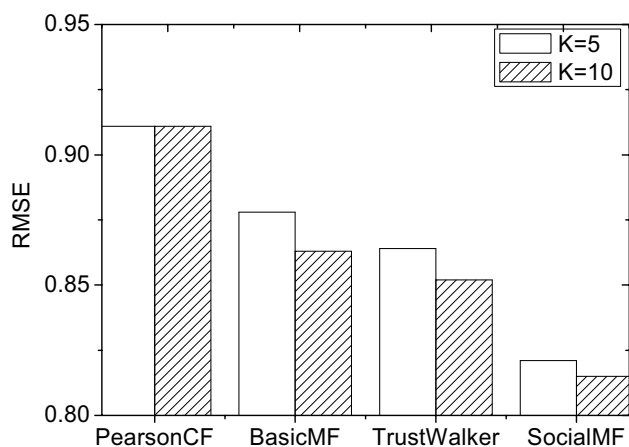


Fig. 2 Comparison accuracy for Flixster

of dimension, we conducted another experiment where we separate the entire data set into train and test data only and determine the sensitivity of dimensions. In the experiment, we use 100% of the social network and 80% of the user scoring matrix as the training data set, and the remaining 20% of the score matrix is adopted as the test set (Sattarpour et al. 2018).

First of all, the dimensions of the decomposed matrix are given as $K = 5$ and $K = 10$ in the matrix decomposition, respectively, and the prediction accuracy of the algorithm in these two cases is compared. Figures 1 and 2 are the comparison results for all of algorithms under the Epinions and Flixster datasets respectively it is shown from these two figures that our proposed matrix decomposition algorithm has the smallest prediction error.

Secondly, the normalized parameter λ of the proposed algorithm is evaluated. According to training sample set T_{test} , we adjust the parameters λ_T and observe the prediction error under the two data sets, whose results are shown in Figs. 3 and 4. It can be seen from these two figures that RMSE

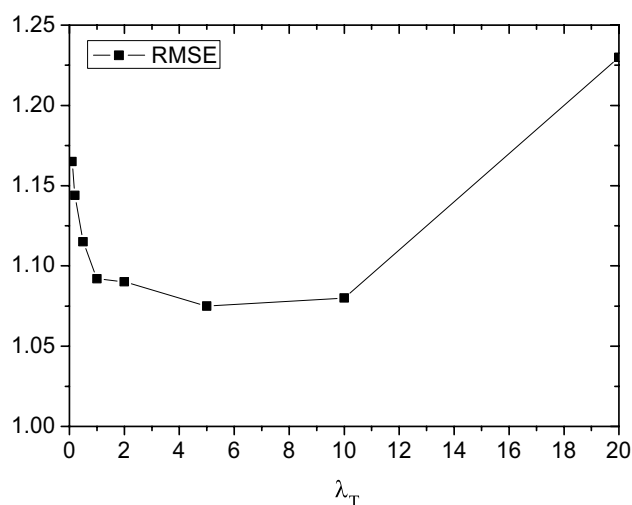


Fig. 3 RMSE for normalized parameter λ_T (Epinions)

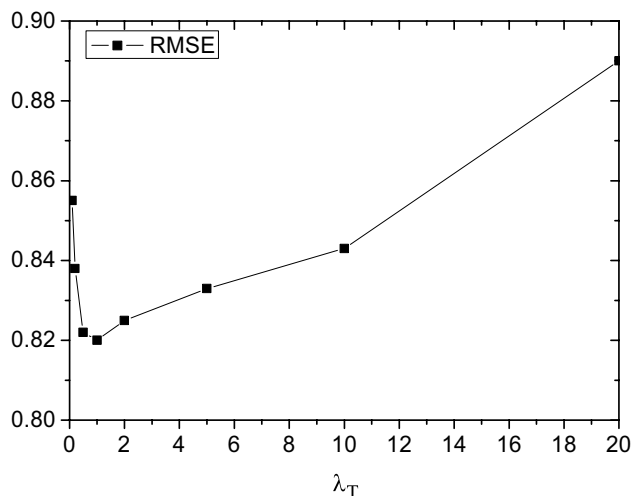


Fig. 4 RMSE for normalized parameter λ_T (Flixster)

initially decreases up to a certain point, and then increases because the parameter λ_T changes. In addition, the curve of RMSE can obtain the optimal value in a very small case.

In order to evaluate the number of features of the matrix decomposition result, the assessment diagram is done in Fig. 5. Since k represents the top largest feature that affects the user's score, the size of k affects the accuracy of the user's score. $k = 4, 8, 16$ and 20 are set to test the RMSE of the two data sets in the algorithm, and the result is shown in Fig. 5. It can be seen from the figure, with the increase of k , the RMSE of our algorithm is gradually decreased. RMSE changes fast when k is very small; the rate of change of RMSE also decreases when k is increasing gradually.

Finally, the accuracy of the algorithm in predicting Cold Start Nodes is also analyzed in experiment. In order

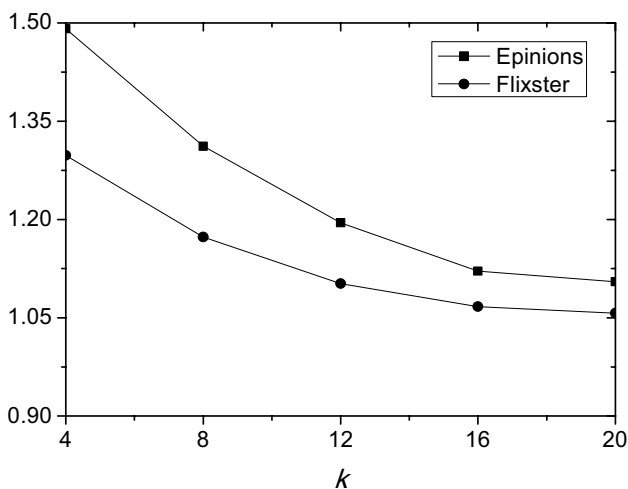


Fig. 5 RMSE for the number of feature *k*

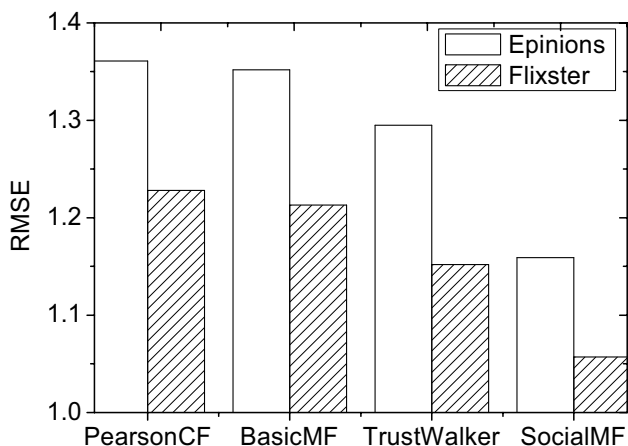


Fig. 6 Cold start nodes

to evaluate the performance of our algorithm in cold-start settings, we split the users into two disjoint subsets, the training set and the test set for each data set, containing 75% and 25% users, respectively. The users in the training set are assumed to be warm-start users whose ratings are known by the system. We learn the models and construct the recommendation process based on these training users. In contrast, the users in the test set are assumed to be cold-start users. When Pearson collaborative filtering algorithm is applied, the users with fewer than 5 neighbors is called as cold start node and the accuracy of the algorithm scoring is compared when dealing with these cold start nodes is given, where $\lambda_T = 2$ for the Epinions dataset and $\lambda_T = 1$ for the Flixster dataset. It can be seen from Fig. 6 that

the proposed matrix factorization algorithm works well to score the cold start node under both dataset.

5 Conclusion

With the explosion of internet data, users would be overwhelmed with the amount and type of data. As an effective tool to extract user-related information, the recommendation system is used by more and more websites. Traditional recommendation systems use the user scoring matrix to predict unknown user-item scores, however social networks can help improve the accuracy of the recommendation system as the social network information on the site continues to accumulate. In this paper, the social network between users is used as ancillary information combined with the user scoring matrix into a matrix, and we design a social algorithm based on the matrix factorization algorithm. Experiments show that the proposed algorithm has a good prediction effect, its performance is significantly better than the existing related research.

As the rapid growth of online social network sites continues, we believe the social-based research will become more and more popular and important. We also plan to develop similar techniques to allow one user’s friends to influence this user’s search results, query suggestions. This would be an interesting social search problem to explore in the future.

Acknowledgements The Scientific and Technological Research Program of Henan Province, China under Grant no. 172102210111. The Ministry of Public Security Technical Research Plan, under Grant no. 2016JSYJB38. The Scientific and Technological Research Program of Henan Province, China under Grant no. 172102210441.

References

Agarwal D, Chen B-C (2010) fLDA: matrix factorization through latent dirichlet allocation. In: Proc. of WSDM2010, New York, New York, USA, pp 91–100

Bedi P, Kaur H, Marwaha S (2007) Trust based recommender system for semantic web. In Proc. of IJCAI, pp 2677–2682, 2007

Bellogín A, Castells P, Cantador I (2013) Improving memory-based collaborative filtering by neighbour selection based on user preference overlap. In: Proceedings of the 10th conference on open research areas in information retrieval, pp 145–148

Golbeck J (2005) Computing and applying trust in web-based social networks. PhD thesis, University of Maryland College Park

Hailing X, Xiao W, Xiaodong L et al (2009) Research on internet recommendation system. J Softw 20(2):350–362

Hsu FM, Lin YT, Ho TK (2012) Design and implementation of an intelligent recommendation system for tourist attractions: the integration of EBM model, Bayesian network and Google Maps. Expert Syst Appl 39(3):3257–3264

Huang Chuanguang Y, Jian W, Jing et al (2010) Research on collaborative filtering recommendation algorithm for indefinite neighbors. Comput Sci 33(8):1369–1377

- Jamali M, Ester M (2009) TrustWalker: a random walk model for combining trust-based and item-based recommendation. In: Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 397–406
- Ma H, Yang H, Lyu MR et al (2008) Sorec: social recommendation using probabilistic matrix factorization. In: Proceedings of the 17th ACM conference on information and knowledge management, ACM, pp 931–940
- Ma H, King I, Lyu MR (2009) Learning to recommend with social trust ensemble. In: Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval. ACM, pp 203–210
- Massa P, Avesani P (2007) Trust-aware recommender systems. In: Proceedings of the 2007 ACM conference on recommender systems, ACM, pp 17–24
- Park DH, Kim HK, Choi IY et al (2012) A literature review and classification of recommender systems research. *Expert Syst Appl* 39(11):10059–10072
- Sarwar B, Karypis G, Konstan J et al (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international conference on world wide web, ACM, pp 285–295
- Sattarpour T, Nazarpour D, Golshannavaz S et al (2018) A multi-objective hybrid GA and TOPSIS approach for sizing and siting of DG and RTU in smart distribution grids. *J Ambient Intell Human Comput* 9:105. <https://doi.org/10.1007/s12652-016-0418-8>
- Tsai CF, Hung C (2012) Cluster ensembles in collaborative filtering recommendation. *Appl Soft Comput* 12(4):1417–1425
- Tu D, Shu C, Yu H (2013) Context advertisement algorithm based on joint probability matrix decomposition. *J Softw* 24(3):454–464
- Wang L, Meng X, Zhang Y (2012) Context perception recommendation system. *J Softw* 23(1):103–115
- Xia KJ, Wang JQ (2017) A novel medical image enhancement algorithm based on improvement correction strategy in wavelet transform domain. *Clust Comput*. <https://doi.org/10.1007/s10586-017-1264-y>
- Xia KJ, Wang JQ, Wu Y (2017) Robust Alzheimer disease classification based on feature integration fusion model for magnetic. *J Med Imaging Health Inf* 7:1–6
- Xia KJ, Yin HS, Wang JQ (2018) A novel improved deep convolutional neural network model for medical image fusion. *Clust Comput* 3:1–13
- Zhang Y, Zhang M, Liu Y et al (2013) Localized matrix factorization for recommendation based on matrix block diagonal forms. In: Proceedings of the 22nd international conference on world wide web, pp 1511–1520
- Zhou T, Shan H, Banerjee A et al (2012) Kernelized probabilistic matrix factorization: exploiting graphs and side information. In: Proceedings of the 2012 SIAM international conference on data mining (SDM), vol 36. pp 403–414

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.