Dual Similarity Regularization for Recommendation

Jing Zheng[1], Jian Liu[1], Chuan Shi[1(✉)], Fuzhen Zhuang[2], Jingzhi Li[3], and Bin Wu[1]

[1] Beijing Key Lab of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876, China
oliviacheng@126.com, fullback@yeah.net, {shichuan,wubin}@bupt.edu.cn
[2] Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
zhuangfz@ics.ict.ac.cn
[3] Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, China
lijz@sustc.edu.cn

Abstract. Recently, social recommendation becomes a hot research direction, which leverages social relations among users to alleviate data sparsity and cold-start problems in recommender systems. The social recommendation methods usually employ simple similarity information of users as social regularization on users. Unfortunately, the widely used social regularization may suffer from several aspects: (1) the similarity information of users only stems from users' social relations; (2) it only has constraint on users; (3) it may not work well for users with low similarity. In order to overcome the shortcomings of social regularization, we propose a new dual similarity regularization to impose the constraint on users and items with high and low similarities simultaneously. With the dual similarity regularization, we design an optimization function to integrate the similarity information of users and items, and a gradient descend solution is derived to optimize the objective function. Experiments on two real datasets validate the effectiveness of the proposed solution.

Keywords: Social recommendation · Regularization · Heterogeneous information network

1 Introduction

Recommender system, as an effective way to tackle information overload problems, has attracted much attention from multiple disciplines. Many techniques have been proposed to build recommender systems. As a popular technique, the low rank matrix factorization has shown its effectiveness and efficiency, which factorizes user-item rating matrix into two low rank user-specific and item-specific matrices, then utilizes the factorized matrices to make further predictions [10].

J. Bailey et al. (Eds.): PAKDD 2016, Part II, LNAI 9652, pp. 542–554, 2016.
DOI: 10.1007/978-3-319-31750-2_43

With the boom of social media, social recommendation has become a hot research topic, which utilizes the social relations among users for better recommendation. Some researchers utilized trust information among users [5,6], and some began to use friend relationship among users [7,12] or other types of information [1,2]. Most of these social recommendation methods employ social regularization to confine similar users under the low rank matrix factorization framework. Specifically, we can obtain the similarity of users from their social relations as a constraint term to confine the latent factors of similar users to be closer. It is reasonable, since similar users should have similar latent features.

However, the social regularization used in social recommendation has several shortcomings. (1) The similarity information of users is only generated from social relations of users. But we can obtain users' similarity from many ways, such as users' contents. (2) The social regularization only has constraint on users. In fact, we can obtain the similarity of items to impose constraint on the latent factors of items. (3) The social regularization may be less effective for dissimilar users, which may lead to dissimilar users having similar factors. The analysis and experiments in Sect. 2 validate this point.

In order to address the limitations of traditional social recommendation, we propose a Dual Similarity Regularization based recommendation method (called DSR) in this paper. Inspired by the success of Heterogeneous Information Network (HIN) in many applications, we organize objects and relations in a recommender system as a HIN, which can integrate all kinds of information, including interactions between users and items, social relations among users and attribute information of users and items. Based on the HIN, we can generate rich similarity information on users and items by setting proper meta paths. Furthermore, we propose a new similarity regularization which can impose the constraint on users and items with high and low similarity. With the similarity regularization, DSR adopts a new optimization objective to integrate those similarity information of users and items. Then we derive its solution to learn the weights of different similarities. The experiments on real datasets show that DSR always performs best compared to social recommendation and HIN-based recommendation methods. Moreover, DSR also achieves the best performance for cold-start users and items due to the dual similarity regularization on users and items.

The rest of the paper is organized as follows. We analyze the limitations of social recommendation in Sect. 2 and introduce the rich similarity information of users and items generated from HIN in Sect. 3. Then we propose the similarity regularization and the DSR model in Sect. 4. We do experiments in Sect. 5, describe related work in Sect. 6 and finally draw the conclusion in Sect. 7.

2 Limitations of Social Recommendation

Recently, with the increasing popularity of social media, there is a surge of social recommendations which leverage rich social relations among users to improve recommendation performance. Ma et al. [7] first proposed the social regularization to extend low-rank matrix factorization, and then it is widely used in a lot of work [4,13]. A basic social recommendation method is illustrated as follows:

$$\min_{U,V} \mathcal{J} = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i V_j^T)^2 + \frac{\alpha}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} S_U(i,j)\|U_i - U_j\|^2 + \frac{\lambda_1}{2}(\|U\|^2 + \|V\|^2), \quad (1)$$

where $m \times n$ rating matrix R depicts users' ratings on n items, R_{ij} is the score user i gives to item j. I_{ij} is an indicator function which equals to 1 if user i rated item j and equals to 0 otherwise. $U \in \mathbb{R}^{m\times d}$ and $V \in \mathbb{R}^{n\times d}$, where $d << min(m,n)$ is the dimension number of latent factor. U_i is the latent vector of user i derived from the ith row of matrix U while V_j is the latent vector of item j derived from the jth row of V. S_U is the similarity matrix of users and $S_U(i,j)$ denotes the similarity of user i and user j. $\|\cdot\|^2$ is the Frobenius norm. Particularly, the second term is the social regularization which is defined as follows:

$$SocReg = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} S_U(i,j)\|U_i - U_j\|^2. \quad (2)$$

As a constraint term in Eq. (1), *SocReg* forces the latent factors of two users to be close when they are very similar. However, it may have two drawbacks.

- The similarity information may be simple. In social recommendation, the similarity information of users is usually generated from rating information or social relations and only one type of similarity information is employed. However, we can obtain much rich similarity information of users and items from various ways, such as rich attribute information and interactions.
- The constraint term may not work well when two users are not very similar. The minimization of optimization objective should force the latent factors of two similar users to be close. But we note that when two users are not similar (i.e., $S_U(i,j)$ is small), *SocReg* may still force the latent factors of these two users to be close. In fact, these two users is dissimilar which means their latent factors should have a large distance.

In order to uncover the limitations of social regularization, we apply the model detailed in Eq. (1) to conduct four experiments each with different levels of similarity information ($None$, *Low*, $High$, *All*). $None$ denotes that we utilize no similarity information in the model (i.e., $alpha = 0$ in the model), *Low* denotes that we utilize bottom 20 % users' similarity information generated in the model, $High$ is that of top 20 %, *All* denotes we utilize all users' similarity information. The Douban dataset detailed in Table 1 is employed in the experiments and we report MAE and RMSE (defined in Sect. 5.2) in Fig. 1. The results of *Low*, $High$ and *All* are better than that of $None$, which implies social regularization really works in the model. However, in terms of performance improvement compared to $None$, *Low* does not improve as much as $High$ and *All* do. The above analysis reveals that the social regularization may not work well in recommender models when users are with low similarity.

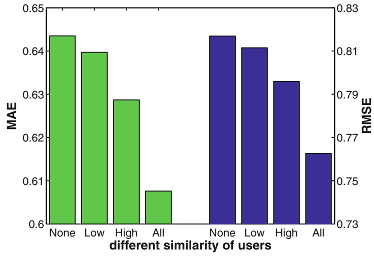
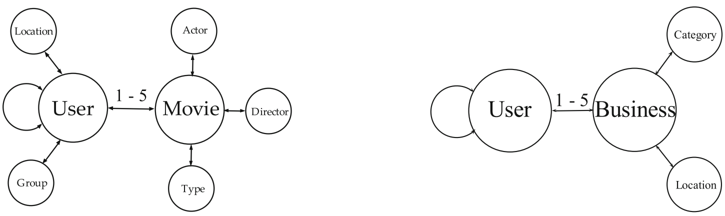


Fig. 1. Limitations of social regularization.

3 Rich Similarity Generated from HIN

Traditional social recommendations only consider the constraint of users with their social relations. However, rich similarity information on users and items can be generated in a heterogeneous information network. A **heterogeneous information network** [11] is a special information network with multiple types of entities and relations. Figure 2(a) shows a typical HIN extracted from a movie recommender system. The HIN contains multiple types of objects, e.g., users (U), movies (M), groups (G), and actors (A).

Two types of objects in a HIN can be connected via various **meta path** [11], which is a composite relation connecting these two types of objects. A meta path $\mathcal{P}$ is a path defined on a schema $\mathcal{S} = (\mathcal{A}, \mathcal{R})$, and is denoted in the form of $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \cdots \xrightarrow{R_l} A_{l+1}$ (abbreviated as $A_1 A_2 \cdots A_{l+1}$), which defines a composite relation $R = R_1 \circ R_2 \circ \cdots \circ R_l$ between type A_1 and A_{l+1}, where $\circ$ denotes the composition operator on relations. As an example in Fig. 2(a), users can be connected via "User-User" (UU), "User-Movie-User" (UMU) and so on. Different meta paths denote different semantic relations, e.g., the UU path means that users have social relations while the UMU path means that users have watched the same movies. Therefore we can evaluate the similarity of users (or movies) based on different meta paths. For example, for users, we can consider UU, UGU, UMU, etc. Similarly, meaningful meta paths connecting movies include MAM, MDM, etc.



(a) Douban (b) Yelp

Fig. 2. Network schema of HIN examples.

Several path-based similarity measures have been proposed to evaluate the similarity of objects under given meta path in HIN [9,11]. We assume that $S_U^{(p)}$ denotes similarity matrix of users under meta path $\mathcal{P}_U^{(p)}$ connecting users, and $S_U^{(p)}(i,j)$ denotes the similarity of users i and j under the path $\mathcal{P}_U^{(p)}$. Similarly, $S_I^{(q)}$ denotes similarity matrix of items under the path $\mathcal{P}_I^{(q)}$ connecting items, and $S_I^{(q)}(i,j)$ denotes the similarity of items i and j.

Since users (or items) have different similarities under various meta paths, we combine their similarities on all paths through assigning weights on these paths. For users and items, we define S_U and S_I to represent the similarity matrix of users and items on all meta paths, respectively.

$$S_U = \sum_{p=1}^{|\mathcal{P}_U|} \boldsymbol{w}_U^{(p)} S_U^{(p)}, \tag{3}$$

$$S_I = \sum_{q=1}^{|\mathcal{P}_I|} \boldsymbol{w}_I^{(q)} S_I^{(q)}, \tag{4}$$

where $\boldsymbol{w}_U^{(p)}$ denotes the weight of meta path $\mathcal{P}_U^{(p)}$ among all meta paths $\mathcal{P}_U$ connecting users, and $\boldsymbol{w}_I^{(q)}$ denotes the weight of meta path $\mathcal{P}_I^{(q)}$ among all meta paths $\mathcal{P}_I$ connecting items.

4 Matrix Factorization with Similarity Regularization

In this section, we propose our dual similarity regularization based matrix factorization method **DSR** and infer its learning algorithm.

4.1 Similarity Regularization

Due to the limitations of social regularization, we design a new similarity regularization to constrain users and items simultaneously with much similarity information of users and items. The basic idea of similarity regularization is that the distance of latent factors of two users (or items) should be negatively correlated to their similarity, which means two similar users (or items) should have a short distance while two dissimilar ones should have a long distance with their latent factors. We note that the Gaussian function meet above requirement and the range of it is [0,1], which is the same with the range of similarity function. Following the idea, we design a similarity regularization on users as follows:

$$SimReg^{\mathcal{U}} = \frac{1}{8}\sum_{i=1}^{m}\sum_{j=1}^{m}(S_U(i,j) - e^{-\gamma\|U_i - U_j\|^2})^2, \tag{5}$$

where γ controls the radial intensity of Gaussian function and the coefficient $\frac{1}{8}$ is convenient for deriving the learning algorithm. This similarity regularization can enforce constraint on both similar and dissimilar users. In addition, the similarity matrix S_U can be generated from social relations or the above HIN. Similarly, we can also design the similarity regularization on items as follows:

$$SimReg^{\mathcal{I}} = \frac{1}{8}\sum_{i=1}^{n}\sum_{j=1}^{n}(S_I(i,j) - e^{-\gamma\|V_i - V_j\|^2})^2. \tag{6}$$

The Proposed DSR Model. We propose the **D**ual **S**imilarity regularization for **R**ecommendation (called DSR) through adding the similarity regularization on users and items into low-rank matrix factorization framework. Specifically, the optimization model is proposed as follows:

$$\begin{aligned}
\min_{U,V,\boldsymbol{w}_U,\boldsymbol{w}_I} \mathcal{J} = & \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i V_j^T)^2 \\
& + \frac{\lambda_1}{2}(\|U\|^2 + \|V\|^2) + \frac{\lambda_2}{2}(\|\boldsymbol{w}_U\|^2 + \|\boldsymbol{w}_I\|^2) \\
& + \alpha SimReg^{\mathcal{U}} + \beta SimReg^{\mathcal{I}} \\
s.t. \quad & \sum_{p=1}^{|\mathcal{P}_U|} \boldsymbol{w}_U^{(p)} = 1, \boldsymbol{w}_U^{(p)} \geq 0 \\
& \sum_{q=1}^{|\mathcal{P}_I|} \boldsymbol{w}_I^{(q)} = 1, \boldsymbol{w}_I^{(q)} \geq 0,
\end{aligned} \tag{7}$$

where α and β control the ratio of similarity regularization term on users and items, respectively.

4.2 The Learning Algorithm

The learning algorithm of DSR can be divided into two steps. (1) Optimize the latent factor matrices of users and items (i.e., U, V) with the fixed weight vectors $\boldsymbol{w}_U = [\boldsymbol{w}_U^{(1)}, \boldsymbol{w}_U^{(2)}, \ldots, \boldsymbol{w}_U^{(|\mathcal{P}_U|)}]^T$ and $\boldsymbol{w}_I = [\boldsymbol{w}_I^{(1)}, \boldsymbol{w}_I^{(2)}, \ldots, \boldsymbol{w}_I^{(|\mathcal{P}_I|)}]^T$. (2) Optimize the weight vectors $\boldsymbol{w}_U$ and $\boldsymbol{w}_I$ with the fixed latent factor matrices U and V. Through iteratively optimizing these two steps, we can obtain the optimal U, V, $\boldsymbol{w}_U$, and $\boldsymbol{w}_I$.

Optimize *U* and *V*. With the fixed $\boldsymbol{w}_U$ and $\boldsymbol{w}_I$, we can optimize U and V by performing stochastic gradient descent.

$$\begin{aligned}
\frac{\partial \mathcal{J}}{\partial U_i} = & \sum_{j=1}^{n} I_{ij}(U_i V_j^T - R_{ij})V_j \\
& + \alpha\sum_{j=1}^{m}\gamma[(S_U(i,j) - e^{-\gamma\|U_i - U_j\|^2})e^{-\gamma\|U_i - U_j\|^2}(U_i - U_j)] + \lambda_1 U_i,
\end{aligned} \tag{8}$$

$$\frac{\partial \mathcal{J}}{\partial V_j} = \sum_{i=1}^{m} I_{ij}(U_i V_j^T - R_{ij})U_i + \beta \sum_{i=1}^{n} \gamma[(S_I(i,j) - e^{-\gamma\|V_i - V_j\|^2})e^{-\gamma\|V_i - V_j\|^2}(V_i - V_j)] + \lambda_1 V_j. \tag{9}$$

Optimize $\boldsymbol{w}_U$ and $\boldsymbol{w}_I$. With the fixed U and V, the minimization of $\mathcal{J}$ with respect to $\boldsymbol{w}_U$ and $\boldsymbol{w}_I$ is a well-studied quadratic optimization problem with non-negative bound. We can use the standard trust region reflective algorithm to update $\boldsymbol{w}_U$ and $\boldsymbol{w}_I$ at each iteration. We can simplify the optimization function of $\boldsymbol{w}_U$ as the following standard quadratic formula:

$$\min_{\boldsymbol{w}_U} \ \frac{1}{2}\boldsymbol{w}_U^T H_U \boldsymbol{w}_U + f_U^T \boldsymbol{w}_U \qquad s.t. \ \sum_{p=1}^{|\mathcal{P}_U|} \boldsymbol{w}_U^{(p)} = 1, \boldsymbol{w}_U^{(p)} \geq 0. \tag{10}$$

Here H_U is a $|\mathcal{P}_U| \times |\mathcal{P}_U|$ symmetric matrix as follows:

$$H_U(i,j) = \begin{cases} \frac{\alpha}{4}(\sum\sum S_U^{(i)} \odot S_U^{(j)}) & i \neq j, 1 \leq i,j \leq |\mathcal{P}_U| \\ \frac{\alpha}{4}(\sum\sum S_U^{(i)} \odot S_U^{(j)}) + \lambda_2 & i = j, 1 \leq i,j \leq |\mathcal{P}_U|, \end{cases}$$

$\odot$ denotes the dot product. f_U is a column vector with length $|\mathcal{P}_U|$, which is calculated as follows:

$$f_U(p) = -\frac{\alpha}{4}\sum_{i=1}^{m}\sum_{j=1}^{m} S_U^{(p)}(i,j)e^{-\gamma\|U_i - U_j\|^2}.$$

Similarly, we can also infer the optimization function of $\boldsymbol{w}_I$.

5 Experiments

In this section, we conduct experiments to validate the effectiveness of DSR and further explore the cold-start problem.

5.1 Dataset

We use a real dataset from Douban[1], a well known social media network in China, which includes 3,022 users and 6,971 movies with 195,493 ratings ranging from 1 to 5. And another real dataset is employed from Yelp[2], a famous user review website in America, which includes 14,085 users and 14,037 movies with 194,255 ratings ranging from 1 to 5. The description of two datasets can be seen in Table 1 and their network schemas are shown in Fig. 2. The Douban dataset has sparse social relationship with dense rating information while the Yelp dataset has dense social relationships with sparse rating information.

[1] http://movie.douban.com/.
[2] http://www.yelp.com/.

Table 1. Statistics of Douban and Yelp dataset

Datasets	Relations of (A − B)	Number of A/B/A − B	Ave. degrees of A/B
Douban	User-Movie	3022/6971/195493	64.69/28.04
	User-User	779/779/1366	1.75/1.75
	User-Group	2212/2269/7054	3.11/3.11
	User-Location	2491/244/2491	1.00/10.21
	Movie-Director	3014/789/3314	1.09/4.20
	Movie-Actor	5438/3004/15585	2.87/5.19
	Movie-Type	6787/36/15598	2.29/433.28
Yelp	User-Business	14085/14037/194255	4.6/20.7
	User-User	9581/9581/150532	10.0/10.0
	Business-Category	14037/575/39406	2.8/73.9
	Business-Location	14037/62/14037	1.0/236.1

5.2 Comparison Methods and Metrics

In order to validate the effectiveness of DSR, we compare it with following representative methods. Besides the classical social recommendation method SoMF, the experiments also include two recent HIN based methods HeteCF and HeteMF. In addition, we include the revised version of SoMF with similarity regularization (i.e., SoMF_{SR}) to validate the effectiveness of similarity regularization.

- **UserMean**. It employs a user's mean rating to predict the missing ratings directly.
- **ItemMean**. It employs an item's mean rating to predict the missing ratings directly.
- **PMF** [8]. Salakhutdinov and Minh proposed the basic low-rank matrix factorization method for recommendation.
- **SoMF** [7]. Ma et al. proposed the social recommendation method with social regularization on users.
- **HeteCF** [4]. Luo et al. proposed the social collaborative filtering algorithm using heterogeneous relations.
- **HeteMF** [13]. Yu et al. proposed the HIN based recommendation method through combining user ratings and items' similarity matrices.
- **SoMF_{SR}**. It adapts SoMF through only replacing the social regularization with the similarity regularization $SimReg^{\mathcal{U}}$.

For Douban dataset, we utilize 7 meta paths for user (i.e., UU, UGU, ULU, UMU, UMDMU, UMTMU, UMAMU) and 5 meta paths for item (i.e., MTM, MDM, MAM, MUM, MUUM). For Yelp dataset, we utilize 2 meta paths for user (i.e., UB, UU) and 2 meta paths for item (i.e., BC, BL). HeteSim [9] is employed to evaluate the object similarity based on above meta paths. These similarity matrices are fairly utilized for HeteCF, HeteMF, and DSR. We set $\gamma = 1$, $\alpha = 10$, and $\beta = 10$ through parameter experiments on Douban dataset.

In the experiments on Yelp dataset, we set the parameters $\gamma = 1$, $\alpha = 10$, $\beta = 10$. Meanwhile, optimal parameters are set for other models in the experiments.

We use Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to evaluate the performance of rating prediction:

$$MAE = \frac{\sum_{(u,i)\in R} |R_{u,i} - \hat{R}_{u,i}|}{|R|}, \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{(u,i)\in R} (R_{u,i} - \hat{R}_{u,i})^2}{|R|}}, \tag{12}$$

where R denotes the whole rating set, $R_{u,i}$ denotes the rating user u gave to item i, and $\hat{R}_{u,i}$ denotes the rating user u gave to item i as predicted by a certain method. A smaller MAE or RMSE means a better performance.

5.3 Effectiveness Experiments

For Douban dataset, we use different ratios (80 %, 60 %, 40 %) of data as training sets and the rest of the dataset for testing. Considering the sparse density of Yelp dataset, we use 90 %, 80 %, 70 % of data as training sets and the rest of the dataset for testing for Yelp dataset. The random selection is carried out 10 times independently and we report the average results in Table 2.

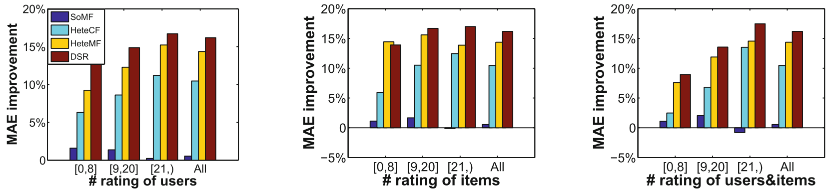
It is clear that three HIN based methods (DSR, HeteCF, and HeteMF) all achieve significant performance improvements compared to PMF, UserMean, ItemMean and SoMF. It implies that integrating heterogeneous information is a promising way to improve recommendation performance. Particularly, DSR always has the best performance on all conditions compared to other methods. It indicates that the dual similarity regularization on users and items may be more effective than traditional social regularization. It can be further confirmed by the better performance of $\mathrm{SoMF_{SR}}$ over SoMF. Although the superiority of $\mathrm{SoMF_{SR}}$ over SoMF is not significant, the improvement is achieved on the very weak social relations in Douban dataset. In addition, we can also find that DSR has better performance improvement for less training data. It reveals that DSR has the potential to alleviate the cold-start problem.

5.4 Study on Cold-Start Problem

To validate the superiority of DSR on cold-start problem, we run PMF, SoMF, HeteCF, HeteMF, DSR on Douban dataset with 40 % training ratio. Four levels of users are seted: three types of cold-start users with various numbers of rated movies (e.g., [0,8] denotes users rated no more than 8 movies and "All" means all users in Fig. 3). We conduct similar experiments on cold-start items and users & items (users and items are both cold-start). The experiments are shown in Fig. 3. Once again, we find that 3 HIN-based methods all are effective for cold-start

Table 2. Effectiveness experimental results on Douban and Yelp (The improvement is based on PMF)

Dataset	Training	Metrics	PMF	UserMean	ItemMean	SoMF	HeteCF	HeteMF	$SoMF_{SR}$	DSR
Douban	80 %	MAE	0.6444	0.6954	0.6284	0.6396	0.6101	0.5941	0.6336	**0.5856**
		Improve		−7.92 %	2.47 %	0.73 %	5.32 %	7.79 %	1.68 %	9.12 %
		RMSE	0.8151	0.8658	0.7928	0.8098	0.7657	0.7520	0.8000	**0.7379**
		Improve		−6.23 %	2.73 %	0.64 %	6.05 %	7.73 %	1.85 %	9.46 %
	60 %	MAE	0.6780	0.6967	0.6370	0.6696	0.6317	0.6056	0.6648	**0.5946**
		Improve		−2.76 %	6.05 %	1.25 %	6.84 %	10.68 %	1.96 %	12.31 %
		RMSE	0.8569	0.8687	0.8135	0.8445	0.7901	0.7665	0.8358	**0.7483**
		Improve		−1.37 %	5.07 %	1.45 %	7.80 %	10.56 %	2.46 %	12.68 %
	40 %	MAE	0.7364	0.7009	0.6629	0.7245	0.6762	0.6255	0.7141	**0.6092**
		Improve		4.83 %	9.99 %	1.63 %	8.18 %	15.07 %	3.03 %	17.28 %
		RMSE	0.9221	0.8747	0.8747	0.9058	0.8404	0.7891	0.8950	**0.7629**
		Improve		5.14 %	5.13 %	1.76 %	8.86 %	14.42 %	2.94 %	17.27 %
Yelp	90 %	MAE	0.8475	0.9543	0.8822	0.8460	0.8461	0.8960	0.8459	**0.8158**
		Improve		−12.60 %	−4.09 %	0.18 %	0.17 %	−5.72 %	0.18 %	3.74 %
		RMSE	1.0796	1.3138	1.2106	1.0772	1.0773	1.1272	1.0772	**1.0369**
		Improve		−21.69 %	−12.13 %	0.22 %	0.21 %	−4.41 %	0.22 %	3.95 %
	80 %	MAE	0.8528	0.9621	0.8931	0.8527	0.8528	0.8907	0.8526	**0.8206**
		Improve		−12.82 %	−4.72 %	0.01 %	0.00 %	−4.44 %	0.01 %	3.78 %
		RMSE	1.0850	1.3255	1.2304	1.0849	1.0850	1.1195	1.0848	**1.0413**
		Improve		−22.17 %	−13.40 %	0.01 %	0.00 %	−3.18 %	0.02 %	4.03 %
	70 %	MAE	0.8576	0.9706	0.9062	0.8575	0.8576	0.8976	0.8575	**0.8250**
		Improve		−13.17 %	−5.67 %	0.01 %	0.00 %	−4.66 %	0.01 %	3.80 %
		RMSE	1.0894	1.3395	1.2547	1.0936	1.0894	1.1313	1.0894	**1.0461**
		Improve		−22.96 %	−15.17 %	−0.39 %	0.00 %	−3.85 %	0.00 %	3.97 %



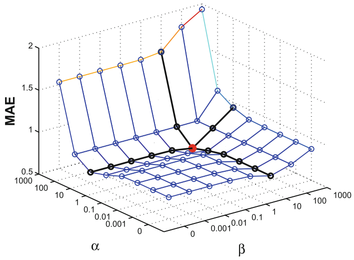
(a) Users_MAE (b) Items_MAE (c) Users&Items_MAE

Fig. 3. MAE improvement against PMF on various cold-start levels.

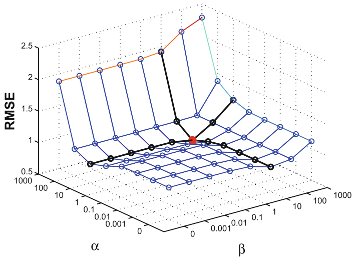
users and items. Moreover, DSR always has the highest MAE improvement on almost all conditions, due to dual similarity regularization on users and items. It's reasonable since the DSR method takes much constraint information of users and items into account which would play a crucial role when there's little available information of users or items.

5.5 Parameter Study on α and β

The DSR model is based on the low-rank matrix factorization framework and the similar regularization on users and items is applied to constrain the model



(a) MAE



(b) RMSE

Fig. 4. Parameter study on MAE and RMSE

learning process. The relevant parameters of the basic matrix factorization have been studied in other matrix factorization methods. In this section we only study α and β which are the parameters of dual similarity regularization on Douban dataset.

Figure 4 shows that the impacts of α and β on MAE and RMSE are quite similar. When the values of α and β are both around 10, the experiment has the best performance. When the values of α and β are quite large or small, the results are not ideal. When α and β set the proper value (in our experiments they are both 10), regularization and rating information take effect on the learning process simultaneously so that the experiments could get better performance. It indicates that integrating the similarity information of users and items in a HIN has a significant impact on recommender systems.

Compared to the optimal result, the experimental results decline sharply when the values of α and β are increased from 10. On the other hand, when α and β are quite small, DSR performs like basic matrix factorization method but the experimental results are not too bad.

6 Related Work

With the prevalence of social media, social recommendation has attracted many researchers. Ma et al. [6] fused user-item matrix with users' social trust networks. In [7], the social regularization ensures that the latent feature vectors of two friends with similar tastes to be closer. Yang et al. [12] inferred category-specific social trust circles from available rating data combined with friend relations.

To further improve recommendation performance, more and more researchers have been aware of the importance of heterogeneous information network (HIN), in which objects are of different types and links among objects represent different relations. Zhang et al. [14] investigated the problem of recommendation over heterogeneous network and proposed a random walk model to estimate the importance of each object in the heterogeneous network. Considering heterogeneous network constructed by different interactions of users, Jamali and Lakshmanan [3] proposed HETEROMF to integrate a general latent factor and context-dependent latent factors. Yu et al. [13] proposed Hete-MF through combining rating information and items' similarities derived from meta paths in HIN. More recently, Luo et al. [4] proposed a collaborative filtering-based social recommendation method, called Hete-CF, using heterogeneous relations.

7 Conclusions

In the paper, we analyzed the limitations of social regularization and designed a similarity regularization whose basic idea is to enforce the constraint on both similar and dissimilar objects. Then, we employ the similarity regularization on low-rank matrix factorization framework and proposed the DSR method. Experiments validate the effectiveness of DSR, especially on alleviating the cold-start problem.

Acknowledgments. This work is supported in part by National Key Basic Research and Department (973) Program of China (No. 2013CB329606), and the National Natural Science Foundation of China (No. 71231002, 61375058, 11571161), and the CCF-Tencent Open Fund, the Co-construction Project of Beijing Municipal Commission of Education, and Shenzhen Sci.-Tech Fund No. JCYJ20140509143748226.

References

1. BellogíN, R., Cantador, I., Castells, P.: A comparative study of heterogeneous item recommendations in social systems. Inf. Sci. **221**, 142–169 (2013)
2. Cantador, I., Bellogin, A., Vallet, D.: Content-based recommendation in social tagging systems. In: RecSys, pp. 237–240 (2010)
3. Jamali, M., Lakshmanan, L.V.: Heteromf: recommendation in heterogeneous information networks using context dependent factor models. In: WWW, pp. 643–653 (2013)
4. Luo, C., Pang, W., Wang, Z.: Hete-cf: Social-based collaborative filtering recommendation using heterogeneous relations. In: ICDM, pp. 917–922 (2014)
5. Ma, H., King, I., Lyu, M.R.: Learning to recommend with social trust ensemble. In: SIGIR, pp. 203–210 (2011)
6. Ma, H., Yang, H., Lyu, M.R., King, I.: Sorec: Social recommendation using probabilistic matrix factorization. In: CIKM, pp. 931–940 (2008)
7. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I.: Recommender systems with social regularization. In: WSDM, pp. 287–296 (2011)
8. Salakhutdinov, R., Mnih, A.: Probabilistic matrix factorization. In: NIPS, pp. 1257–1264 (2012)
9. Shi, C., Kong, X., Huang, Y., Yu, P.S., Wu, B.: Hetesim: a general framework for relevance measure in heterogeneous networks. IEEE Trans. Knowl. Data Eng. **26**(10), 2479–2492 (2014)
10. Srebro, N., Jaakkola, T.: Weighted low-rank approximations. In: ICML, pp. 720–727 (2003)
11. Sun, Y., Han, J., Yan, X., Yu, P., Wu, T.: Pathsim: meta path-based top-k similarity search in heterogeneous information networks. In: VLDB, pp. 992–1003 (2011)

12. Yang, X., Steck, H., Liu, Y.: Circle-based recommendation in online social networks. In: KDD, pp. 1267–1275 (2012)
13. Yu, X., Ren, X., Gu, Q., Sun, Y., Han, J.: Collaborative filtering with entity similarity regularization in heterogeneous information networks. In: IJCAI-HINA Workshop (2013)
14. Zhang, J., Tang, J., Liang, B., Yang, Z., Wang, S., Zuo, J., Li, J.: Recommendation over a heterogeneous social network. In: The Ninth International Conference on Web-Age Information Management, WAIM 2008, pp. 309–316. IEEE (2008)