CrossMark

REGULAR PAPER

Recommendation in heterogeneous information network via dual similarity regularization

Jing Zheng[1] · Jian Liu[1] · Chuan Shi[1] · Fuzhen Zhuang[2] · Jingzhi Li[3] · Bin Wu[1]

Abstract Recommender system has caught much attention from multiple disciplines, and many techniques are proposed to build it. Recently, social recommendation becomes a hot research direction. The social recommendation methods tend to leverage social relations among users obtained from social network to alleviate data sparsity and cold-start problems in recommender systems. It employs simple similarity information of users as social regularization on users. Unfortunately, the widely used social regularization suffers from several aspects: (1) The similarity information of users only stems from social relations of related users; (2) it only has constraint on users without considering the impact of items for recommendation; (3) it may not work well for dissimilar users. To overcome the shortcomings of social regularization, we design a novel dual similarity regularization to impose the constraint on users and items with high and low similarities simultaneously. With the dual similarity regularization, we further propose an optimization function to integrate the similarity information of users and items under different semantic meta-paths, and a gradient descend solution is derived to optimize the objective function. Experiments with differ- ent meta-paths validate the superiority of integrating much available information, and the experiments conducted on three real data sets validate the effectiveness of the proposed solution.

Keywords Social recommendation · Regularization · Heterogeneous information network

This paper is an extension version of the PAKDD2016 Long Presentation paper "Dual Similarity Regularization for Recommendation" [40].

✉ Chuan Shi
shichuan@bupt.edu.cn

1 Beijing Key Lab of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing, China

2 Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

3 Southern University of Science and Technology, Shenzhen, China

1 Introduction

Recommender system, which aims to recommend items to user personality, has attracted much attention from multiple disciplines. Many entertainment Web sites, such as Netflix and YouTube, have deployed their own recommender systems to increase users' stickiness. Also many e-commerce companies, such as Amazon and Alibaba, try to improve their own recommender systems to increase users' click conversion rate. Recommender system has become popular in recent years, and many techniques have been proposed to build recommender systems [1,22]. Among them, the low-rank matrix factorization, as a popular technique, which factorizes user–item rating matrix into two low-rank user-specific and item-specific matrices and then utilizes the factorized matrices to make further predictions, has shown its effectiveness and efficiency [29].

With the boom of social media, social recommendation has become a hot research topic. The basic idea of social recommendation is to model social network information as regularization terms to constrain the matrix factorization framework. Social recommendation generates similarity of users through leveraging rich social relations among users, such as friendships in Facebook, following relations in Twitter. Some researchers utilized trust information among users [18,19], and some began to use friend relationship

🖄 Springer

among users [20,34] or other types of information [2,5] to enhance recommendation quality. Most of these social recommendation methods employ the similarity of users as 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, we can find that the social regularization used in social recommendation has following limitations, which may lead to poor performance.

(1) The similarity information of users is only generated from social relations of users. On the one hand, in real world, much more useful information can be utilized to calculate the similarity of users. For example, we can obtain the similarity of two users through analyzing their attribute information or rating information. On the other hand, some users' social relations are weak, which is hard for similarity analysis. In addition, the characteristics of items connecting to two users who have close social relation may be very different, so that we cannot simply employ the similarity generated from their social relation to constrain the latent factor of these two users. Therefore, it is more reasonable to integrate more useful information of users for users' similarity analysis.

(2) The social regularization only has constraint on users. In fact, we can obtain the similarity information of items as well and then we can use it to impose constraint on the latent factors of items. When the information of users is sparse, it is difficult to obtain users' similarity information. In this condition, we cannot employ the social recommendation. On the other hand, available information of items can be collected to analyze the similarity of items, which can be used to confine the latent factors of items for better recommendation.

(3) The social regularization may be less effective for dissimilar users. When two users are similar, the social regularization takes effect and forces the latent factors of these two users to be close. However, when two users are not similar, the social regularization may fail and still forces the latent factors of these two ones to be close, which is inconsistent with their social relation in reality. The analysis and experiments in Sect. 3 validate this point.

In order to address the above limitations we discussed, 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 various information, including interactions between users and items, social relations among users and attribute information of users and items. Based on the HIN, we generate rich similarity information on users and items by setting proper meta-paths. Furthermore, we propose a new dual similarity regularization which can constrain the latent features of users and items with arbitrary similarity simultaneously. With the similarity regularization, DSR adopts a new optimization objective to integrate the similarity information of users and items. Then, we derive its solution to learn the weights of different similarities. The experiments on three real data sets (i.e., Douban-Movie, Douban-Book and Yelp) show that DSR always performs better compared to social recommendation and two HIN-based recommendation methods (i.e., HeteCF and HeteMF). Moreover, DSR also achieves the best performance for cold-start users and items due to the dual similarity regularization on users and items. The main contributions of the paper are summarized as follows:

(1) We design a novel dual similarity regularization, which imposes constraints on users and items with high and low similarities simultaneously. The dual similarity regularization would force the latent factors of two dissimilar objects to be far and force the latent factors of two similar ones to be close, which can avoid the weakness of the social regularization for two dissimilar users.

(2) We propose a unified HIN-based social recommendation method with dual similarity regularization, called DSR for short, which is based on the basic matrix factorization framework. DSR can integrate multiple types of information in HIN to enhance recommendation performance. In addition, DSR can obtain more accurate integrated similarity information of objects through flexible weights assigned to similarities of users or items under different semantic paths.

(3) Extensive experimentation conducted on three real data sets demonstrates the effectiveness of DSR. In particular, in terms of providing recommendation to cold-start objects (i.e., users, items or both of them), DSR can achieve comparable performance, which shows the superiority of our proposed method in alleviating cold-start problem.

The remainder of this paper is organized as follows. We review the related work in Sect. 2. Next, we analyze the limitations of social recommendation in Sect. 3 and introduce the rich similarity information of users and items generated from HIN in Sect. 4. Then, we propose the similarity regularization and the DSR model in Sect. 5. Experimental results are reported in Sect. 6, and the conclusion is drawn in Sect. 7.

2 Related work

Recommender system aims to tackle the information overload problem, and many recommender methods have been presented to improve recommender performance. Traditional recommender methods usually utilize the user–item rating information for recommendation [7,21]. To further enhance the recommendation performance, external information such as the attributes of users and items is considered in many methods [8,11]. Collaborative filtering [10] as one of the most successful approaches, which utilizes information of users preferences to predict items a target user might like, can be grouped into two types: memory-based methods and model-based methods. However, most of them cannot deal with the cold-start problem which is a situation that recommender system fails to draw recommendation for new items, users or both.

With the prevalence of social media, social recommendation has attracted many researchers. Ma et al. [19] fused user–item matrix with users' social trust networks. Bernardi et al. [3] proposed a recommendation algorithm to solve the cold-start problem with social tags. In [20], the social regularization ensures the latent feature vectors of two friends with similar tastes to be closer. Yang et al. [34] inferred category-specific social trust circles from available rating data combined with friend relations. Moreover, Zhang et al. [38] proposed a novel BiFu scheme for the cold-start problem based on the bi-clustering and fusion techniques in large-scale social recommender systems. Considering that users may be affiliated with multiple groups in online social networks, Yang et al. [33] utilized social groups with richer information to improve the recommendation performance.

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 [25]. Zhang et al. [39] 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 [13] proposed HETEROMF to integrate a general latent factor and context-dependent latent factors. Yu et al. [36,37] proposed an implicit feedback recommendation model with extracted latent features from HIN. Moreover, they [35] proposed HeteMF through combining rating information and items' similarities derived from meta-paths in HIN. Luo et al. [17] proposed a collaborative filtering-based social recommendation method, called HeteCF, using heterogeneous relations. Recently, Shi et al. [27] proposed the SemRec on HIN to improve recommendation performance through flexibly integrating information with the help of weighted meta-paths. Furthermore, the SimMF proposed by Shi et al. [26] integrates heterogeneous information via flexible regularization of users and items for better recommendation.

With the development of recommender systems, much more additional information and learning tools have been exploited to improve the recommendation quality. Guo et al. [9] proposed an effective algorithm called LBMF to explore review texts and rating scores simultaneously, which directly correlate user and item latent dimensions with each word in review texts and ratings. Wu et al. [32] presented a novel method for top-N recommendation, called collaborative denoising auto-encoder (CDAE), which utilized the idea of denoising auto-encoders. Cao et al. [31] noticed not only the heterogeneity but also the couplings between objects. They proposed the non-IID recommendation theories and systems [6]. Considering the non-IID context, Li et al. [16] presented a coupled matrix factorization framework, which integrated user couplings and item couplings into the basic matrix factorization model.

Contemporary recommender methods tend to consider single type or two types of information, even for heterogeneous methods. For example, HeteMF proposed by Yu et al. only takes similarity of items into account to improve recommendation performance. Following the existing HIN-based methods, we proposed a novel recommendation method called DSR in HIN, where we can obtain rich useful information of users and items. In order to make full use of the available information, DSR flexible integrates the information of users and items under different semantic paths to improve the recommendation performance. Ma et al. proposed SoMF for social recommendation and designed the social regularization to ensure the latent feature vectors of two friends with similar tastes to be closer; however, it may not work well for dissimilar users. Compared with the social regularization, the dual similarity regularization we proposed imposes the constraint not only on similar users but also on dissimilar users.

3 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. [20] first proposed the social regularization to extend low-rank matrix factorization, and then, it is widely used in a lot of work [17,35]. A basic social recommendation method is illustrated as follows:

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

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:

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

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 rela-tions and only one type of similarity information is employed. However, we can obtain much rich similar-ity 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 are 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 exper-iments each with different levels of similarity information (*None*, *Low*, *High*, *All*). *None* denotes that we do not uti-lize any similarity information in the model (i.e., $\alpha=0$ in the model), *Low* denotes that we utilize the bottom 20% users' similarity information generated in the model, *High* is that of top 20%, and *All* denotes we utilize all users' similarity information. The Douban-Movie data set detailed in Table 1 is employed in the experiments, and we report MAE and RMSE (defined in Sect. 6.2) 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 improve-ment compared to *None*, *Low* does not improve as much as *High* and *All* do. The above analysis reveals that the social

Table 1 Statistics of Douban and Yelp data set

| Data sets	Relations (A–B)	Number of A/B/A–B	Ave. degrees of A/B
Douban-Movie	User–Movie	3022/6971/195,493	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	5438/3004/15,585	2.87/5.19
	Movie–Type	6787/36/15,598	2.29/433.28
Douban-Book	User–Book	13,024/22,347/792,062	60.82/35.44
	User–User	12,748/12,748/169,150	13.27/13.27
	User–Group	13,024/2936/1,189,271	91.31/405.07
	User–Location	10592/453/10592	1.00/23.38
	Book–Publisher	21,773/1815/21,773	1.00/12.00
	Book–Author	21,907/10,805/21,907	1.00/2.03
	Book–Year	21,192/64/21,192	1.00/331.13
Yelp	User–Business	14,085/14,037/194,255	4.6/20.7
	User–User	9581/9581/150,532	10.0/10.0
	Business–Category	14,037/575/39,406	2.8/73.9
	Business–Location	14,037/62/14,037	1.0/236.1
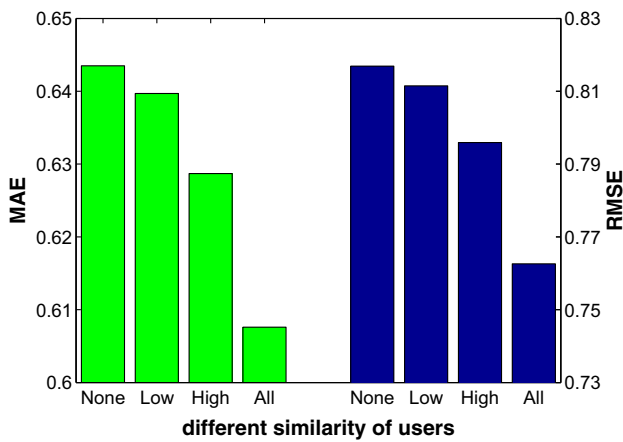
Fig. 1 Limitations of social regularization

regularization may not work well in recommender models when users are with low similarity.

4 Rich similarity generated from HIN

An information network represents an abstraction of the real world, focusing on the objects and the interactions among them. The information network analysis has become a hot spot in data mining and information retrieval fields in the past decades, whose base idea is to mine link relations from massive data in network to find the hidden patterns of different objects.

A number of contemporary recommendation methods model their problems in homogeneous information networks, which ignore the heterogeneity of objects and merely contain one type of relations among objects with the same type, such as the friendship network [15]. Traditional social recommendations only consider the constraint of users with their social relations in an information network. However, most of the recommender systems contain multiple types of objects and relation links. It is hard to model them as homogeneous networks while rich similarity information on users and items

can be generated in a heterogeneous information network. A **heterogeneous information network** [12] is a special information network with multiple types of entities and relations. Compared to traditional homogeneous information networks, the HIN can integrate much more available information effectively and contain richer semantics in objects and links. Moreover, the homogeneous information network is usually constructed from single data domain, while the HIN can fuse information across domains. For example, users always tend to participate in different networks, such as YouTube, Facebook and Twitter.

Recently, the HIN has aroused general interests and many recommendation tasks have been exploited in such networks for better recommendation. Figure 2a 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). Also, it illustrates different kinds of relations among objects, e.g., rating information, attribute information and social relation.

Two types of objects in a HIN can be connected via various **meta-path** [30], 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 $A_1 A_2 \ldots A_{l+1}$), 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. 2a, 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 and MDM.

Several path-based similarity measures have been proposed to evaluate the similarity of objects under given meta-path in HIN [14,24,30]. Sun et al. [30] implement
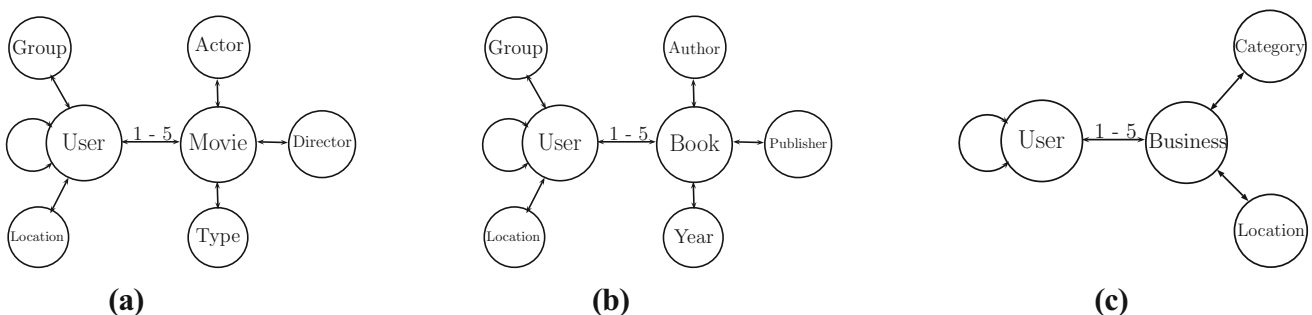

Fig. 2 Network schema of HIN examples. **a** Douban-Movie, **b** Douban-Book, **c** Yelp

the PathSim to measure the similarity of objects with same type based on symmetric paths. Lao et al. [14] propose a path-constrained random walk (PCRW) model to measure the similarity of two objects in a labeled directed graph. Shi et al. [28] presented a semantic-based recommendation system HeteRecom, which utilizes the semantics information of meta-path to evaluate the similarities between movies. The HeteSim [24] is proposed to measure the relatedness of heterogeneous objects based on an arbitrary meta-path. Burke et al. [4] proposed an approach for recommendation through incorporating multiple relations in a weighted hybrid system.

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.

$$
\begin{aligned}
S_U &= \sum_{p=1}^{|\mathcal{P}_U|} \boldsymbol{w}_U^{(p)} S_U^{(p)}, \\
S_I &= \sum_{q=1}^{|\mathcal{P}_I|} \boldsymbol{w}_I^{(q)} S_I^{(q)},
\end{aligned}
\tag{3}
$$

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.

5 Matrix factorization with similarity regularization

In this section, we propose our dual similarity regularization-based matrix factorization method **DSR** and infer its learning algorithm. We first review the basic low-rank matrix factorization method. Then, we introduce dual similarity regularization which imposes the constraint on users and items with high and low similarity. Based on the basic matrix factorization framework and the regularization we designed, we propose the dual similarity regularization for recommendation (called DSR). Finally, we infer the learning algorithm of DSR and show the unified optimization algorithm framework.

5.1 Low-rank matrix factorization

The low-rank matrix factorization has been widely studied in recommendation system. Its basic idea is to factorize the user–item rating matrix R into two matrices (U and V) representing users' and items' distributions on latent semantic, respectively. Then, the rating prediction can be made through these two specific matrices. Assuming an $m \times n$ rating matrix R to be m users' ratings on n items, this approach mainly minimizes the objective function $\mathcal{J}$ as follows.

$$
\begin{aligned}
\min_{U,V} \mathcal{J} =& \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij} \left(R_{ij} - U_i V_j^T \right)^2 \\
& + \frac{\lambda_1}{2} (\|U\|^2 + \|V\|^2),
\end{aligned}
\tag{4}
$$

where I_{ij} is the indicator function that is equal to 1 if user i rated item j and equal to 0 otherwise. $U \in \mathbb{R}^{m \times d}$ and $V \in \mathbb{R}^{n \times d}$, where d is the dimension number of latent factors and $d << \min(m, n)$. U_i is a row vector derived from ith row of matrix U, and V_j is a row vector derived from jth row of matrix V. λ_1 represents the regularization parameter.

In summary, the optimization problem minimizes the sum-of-squared-errors objective function with quadratic regularization terms which aim to avoid over-fitting problem. This problem can be effectively solved by a simple stochastic gradient descent technique.

5.2 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 with 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 meets 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:

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

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:

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

5.2.1 The proposed DSR model

We propose the dual similarity regularization for recommendation (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:

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

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

5.3 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 $\boldsymbol{w}_U = \left[\boldsymbol{w}_U^{(1)}, \boldsymbol{w}_U^{(2)}, \ldots, \boldsymbol{w}_U^{(|\mathcal{P}_U|)} \right]^T$ and $\boldsymbol{w}_I = \left[\boldsymbol{w}_I^{(1)}, \boldsymbol{w}_I^{(2)}, \ldots, \boldsymbol{w}_I^{(|\mathcal{P}_I|)} \right]^T$.

(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 $\boldsymbol{w}_I$.

5.3.1 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.

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

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

5.3.2 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 nonnegative 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:

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

Let $S_{UU}^{(ij)} = S_U^{(i)} \odot S_U^{(j)} (1 \leq i, j \leq |\mathcal{P}_U|)$. Here H_U is a $|\mathcal{P}_U| \times |\mathcal{P}_U|$ symmetric matrix as follows:

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

$\odot$ 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 $\boldsymbol{w}_I$. Algorithm 1 shows the optimization algorithm framework.

6 Experiments

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

6.1 Data preparation

Although many public data sets are available, some of them may not contain enough objects attributes information or social relation information. In order to get more available information of objects and ensure the generality of our method, we use three real data sets in our experiments. Two

Algorithm 1 Alogrithm of DSR

Require:
 G: heterogeneous information network
 $\mathcal{P}_U, \mathcal{P}_I$: meta-path sets of users and items
 η: learning rate for gradient descent
 $\alpha, \beta, \lambda_1, \lambda_2$: see details in above section
 ϵ: convergence tolerance
Ensure:
 U, V: the latent factor of users and items
 $\boldsymbol{w}_U, \boldsymbol{w}_I$: the weights of meta-paths on users and items
 1: **for** $\mathcal{P}_U^{(p)} \in \mathcal{P}_U$ and $\mathcal{P}_I^{(q)} \in \mathcal{P}_I$ **do**
 2:    Calculate the similarity of users and items ($S_U^{(p)}$ and $S_I^{(q)}$)
 3: **end for**
 4: Initialize $U, V, \boldsymbol{w}_U, \boldsymbol{w}_I$
 5: **repeat**
 6:    $U_{old} := U, V_{old} := V$
 7:    Calculate $\frac{\partial \mathcal{J}}{\partial U}, \frac{\partial \mathcal{J}}{\partial V}$
 8:    Update $U := U - \eta * \frac{\partial \mathcal{J}}{\partial U}, V := V - \eta * \frac{\partial \mathcal{J}}{\partial V}$
 9:    Update $\boldsymbol{w}_U, \boldsymbol{w}_I$ with quadratic optimization
 10: **until** $\|U - U_{old}\|^2 + \|V - V_{old}\|^2 < \epsilon$

data sets are from Douban,[1] a well-known social media Web site in China. In this Web site, data in two domains including movie and book are collected and extracted for our experiments. The data set in movie domain includes 3022 users and 6971 movies with 195,493 ratings ranging from 1 to 5, while the data set in book domain comprises 792,062 ratings (scales 1–5) by 13,024 users on 22,347 books. And another real data set is employed from Yelp,[2] a famous user review Web site in America, which includes 14,085 users and 14,037 movies with 194,255 ratings ranging from 1 to 5.

The description of three data sets is shown in Table 1, and their network schemas are shown in Fig. 2. The Douban-Movie data set has sparse social relationship with dense rating information, while the Douban-Book data set and the Yelp data set have dense social relationships with sparse rating information, which leads to different experimental results.

6.2 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 three recent HIN-based methods HeteCF, HeteMF and CMF. 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.
- **ItemCF** [22]. Sarwar et al. proposed the item-based collaborative filtering recommendation algorithm.
- **PMF** [21]. Salakhutdinov and Minh proposed the basic low-rank matrix factorization method for recommendation.
- **SoMF** [20]. Ma et al. proposed the social recommendation method with social regularization on users.
- **HeteCF** [17]. Luo et al. proposed the social collaborative filtering algorithm using heterogeneous relations.
- **HeteMF** [35]. Yu et al. proposed the HIN-based recommendation method through combining user ratings and items' similarity matrices.
- **CMF** [16]. Li et al. proposed the coupled matrix factorization recommendation method integrating user couplings and item couplings into the basic MF model.
- **SoMF$_{SR}$**. It adapts SoMF through only replacing the social regularization with the similarity regularization SimReg$^{\mathcal{U}f}$.

For Douban-Movie data set, we utilize 7 meta-paths for user (i.e., UU, UGU, ULU, UMU, UMDMU, UMTMU and UMAMU) and 5 meta-paths for item (i.e., MTM, MDM, MAM, MUM and MUUM). For Douban-Book data set, we utilize 7 meta-paths for user (i.e., UU, UGU, ULU, UBU, UBPBU, UBABU and UBYBU) and 7 meta-paths for item (i.e., BUB, BPB, BAB, BYB, BUUB, BUGUB and BULUB). For Yelp data set, we utilize 4 meta-paths for user (i.e., UU, UBU, UBCBU and UBLBU) and 4 meta-paths for item (i.e., BUB, BCB, BLB and BUUB). HeteSim [24] 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-Movie and Douban-Book data sets. In the experiments on Yelp data set, we set the parameters $\gamma = 1$, $\alpha = 10$ and $\beta = 10$. 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:

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

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

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. $\hat{R}_{u,i}$ can be calculated by $U_i V_j^T$. A smaller MAE or RMSE means a better performance.

6.3 Effectiveness experiments

For Douban data sets of movie and book, we use different ratios (80, 60, 40%) of data as training sets and the rest of the data set for testing. For example, the training ratio 80% means that we randomly select 80% of rating information as training data while the other 20% of rating information as testing data in our experiments. Considering the sparse density of Yelp data set, we use 90, 80, 70% of data as training sets and the rest as test sets for Yelp data set. The random selection is carried out 10 times independently. We report the average results on Douban and Yelp data sets in Table 2. From the results, we have following observations.

- It is clear that four HIN-based methods (DSR, HeteCF, HeteMF and CMF) all achieve significant performance improvements compared to PMF, UserMean, ItemMean, ItemCF and SoMF. With more additional information to confine the latent factors of users or items, the performance can be improved to some extent. It implies that integrating heterogeneous information is a promising way to improve recommendation performance. Furthermore, we can see that DSR and CMF always perform better than HeteMF and SoMF, which indicates that fusing richer information can improve the recommendation quality better. We can also find that CMF and DSR always achieve the best performances. It also shows the benefits of exploring the feature correlation in CMF.
- DSR outperforms most of the baselines, including three HIN-based methods (DSR, HeteCF and HeteMF), which 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 SoMF$_{SR}$ over SoMF.
- On three real-world data sets, SoMF$_{SR}$ always performs better than SoMF. Although the superiority of SoMF$_{SR}$ over SoMF is not significant, the improvement is achieved on the very weak social relations in Douban-Movie data set. Note that, when training data becomes sparser (for example, when the training ratio is 70% in Yelp data set), SoMF$_{SR}$ performs better or not bad than PMF, while SoMF performs worse than PMF. These observations further confirmed the superiority of the dual similarity regularization.
- Observing the results on three data sets, we can also find that DSR has better performance improvement for less training data. On Douban-Movie data set, four HIN-based methods (DSR, HeteCF, HeteMF and CMF) and the SoMF all have better results than PMF. However, on Yelp data set, HeteCF and SoMF improve performances very limitedly compared with PMF. What is

worse, the results of HeteMF are even worse than PMF. But we can see that DSR always achieves significant performance improvement compared with PMF. It reveals that DSR has the potential to alleviate the cold-start problem.

- CMF takes constraints on users and items through utilizing similarity of users and items. When there are abundant training data, CMF can achieve relatively better performance than DSR and other HIN-based methods, since it compromises non-IID context into basic matrix factorization framework which is different from other HIN-based methods which do not consider the coupling relationships between two objects. However, when training data are sparse (e.g., 70% Yelp training data), the experimental improvement in CMF is less than DSR, because DSR can integrate not only the attribute information of objects but also the social information and rating information of them.

6.4 Study on cold-start problem

Cold start is an important issue in recommender system. It is giving recommendations to novel users who have no or little preference on any items, or recommending items that little or no user of the community has seen yet [23]. To validate the superiority of DSR on cold-start problem, we run PMF, SoMF, HeteCF, HeteMF, DSR on Douban-Movie data set with 40% training ratio. We set four levels of users, including: 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 three HIN-based methods all are effective for cold-start users and items. Moreover, DSR always has the highest MAE and RMSE improvement on almost all conditions, due to dual similarity regularization on users and items. It is 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.

6.5 Impact of different meta-paths

In this section, we study the impact of different meta-paths on the Douban-Movie data set. In each experiment, we run our DSR with the similarity information from only one meta-path. UU* and MM* represent the integrated similarity matrix of users and items on all meta-paths, respectively. Optimal parameters are set for all experiments. The experiment results are shown in Fig. 4.

Table 2 Effectiveness experiments on Douban and Yelp (the improvement is based on PMF)

Data set	Training	Metrics		PMF	UserMean	ItemMean	ItemCF	SoMF	HeteCF	HeteMF	CMF	SoMF$_{SR}$	DSR
Douban-Movie	80%	MAE		0.6444	0.6954	0.6284	0.7007	0.6396	0.6101	0.5941	**0.5684**	0.6336	0.5856
			Improve		−7.92%	2.47%	−8.73%	0.73%	5.32%	7.79%	11.80%	1.68%	9.12%
		RMSE		0.8151	0.8658	0.7928	0.8781	0.8098	0.7657	0.7520	**0.7194**	0.8000	0.7379
			Improve		−6.23%	2.73%	−7.73%	0.64%	6.05%	7.73%	11.74%	1.85%	9.46%
	60%	MAE		0.6780	0.6967	0.6370	0.7022	0.6696	0.6317	0.6056	**0.5732**	0.6648	0.5946
			Improve		−2.76%	6.05%	−3.57%	1.25%	6.84%	10.68%	15.46%	1.96%	12.31%
		RMSE		0.8569	0.8687	0.8135	0.8809	0.8445	0.7901	0.7665	**0.7254**	0.8358	0.7483
			Improve		−1.37%	5.07%	2.80%	1.45%	7.80%	10.56%	15.35%	2.46%	12.68%
	40%	MAE		0.7364	0.7009	0.6629	0.7064	0.7245	0.6762	0.6255	**0.5833**	0.7141	0.6092
			Improve		4.83%	9.99%	4.08%	1.63%	8.18%	15.07%	20.79%	3.03%	17.28%
		RMSE		0.9221	0.8747	0.8747	0.8863	0.9058	0.8404	0.7891	**0.7373**	0.8950	0.7629
			Improve		5.14%	5.13%	3.88%	1.76%	8.86%	14.42%	20.04%	2.94%	17.27%
Douban-Book	80%	MAE		0.5729	0.6207	0.6055	0.6244	0.5713	0.5667	0.5781	0.5729	0.5709	**0.5578**
			Improve		−8.34%	−5.70%	−8.99%	0.28%	1.07%	−0.90%	0.01%	0.35%	2.64%
		RMSE		0.7322	0.7899	0.7517	0.7912	0.7259	0.7174	0.7352	0.7235	0.7266	**0.7001**
			Improve		−7.89%	−2.67%	−8.05%	0.86%	2.02%	−0.42%	1.19%	0.77%	4.38%
	60%	MAE		0.5882	0.6244	0.6092	0.6273	0.5886	0.5775	0.5993	0.5878	0.5872	**0.5661**
			Improve		−6.16%	−3.58%	−6.65%	−0.08%	1.81%	−1.88%	0.07%	0.17%	3.76%
		RMSE		0.7510	0.7991	0.7588	0.7967	0.7477	0.7323	0.7608	0.7431	0.7468	**0.7108**
			Improve		−6.40%	0.44%	−6.09%	−0.94%	2.50%	−1.30%	1.05%	0.57%	5.35%
	40%	MAE		0.6149	0.6330	0.6232	0.6344	0.6171	0.5927	0.6425	0.6584	0.6156	**0.5788**
			Improve		−2.95%	−1.35%	−3.16%	−0.37%	3.60%	−4.49%	−7.07%	−0.13%	5.87%
		RMSE		0.7844	0.8209	0.7944	0.8111	0.7858	0.7546	0.8127	0.8292	0.7836	**0.7293**
			Improve		−4.64%	−1.27%	−3.44%	−0.17%	3.81%	−3.60%	−5.71%	0.11%	7.02%
Yelp	90%	MAE		0.8475	0.9543	0.8822	0.9443	0.8460	0.8461	0.8960	**0.8099**	0.8459	0.8158
			Improve		−12.60%	−4.09%	−11.42%	0.18%	0.17%	−5.72%	4.44%	0.18%	3.74%
		RMSE		1.0796	1.3138	1.2106	1.2932	1.0772	1.0773	1.1272	**1.0343**	1.0772	1.0369
			Improve		−21.69%	−12.13%	−19.78%	0.22%	0.21%	−4.41%	4.20%	0.22%	3.95%
	80%	MAE		0.8528	0.9621	0.8931	0.9658	0.8527	0.8528	0.8907	**0.8197**	0.8526	0.8206
			Improve		−12.82%	−4.72%	−13.25%	0.01%	0.00%	−4.44%	3.88%	0.01%	3.78%
		RMSE		1.0850	1.3255	1.2304	1.3372	1.0849	1.0850	1.1195	1.0467	1.0848	**1.0413**
			Improve		−22.17%	−13.40%	−23.24%	0.01%	0.00%	−3.18%	3.53%	0.02%	4.03%
	70%	MAE		0.8576	0.9706	0.9062	0.9860	0.8575	0.8576	0.8976	0.8298	0.8575	**0.8250**
			Improve		−13.17%	−5.67%	−26.65%	0.01%	0.00%	−4.66%	3.24%	0.01%	3.80%
		RMSE		1.0894	1.3395	1.2547	1.3798	1.0936	1.0894	1.1313	1.0592	1.0894	**1.0461**
			Improve		−22.96%	−15.17%	−26.65%	−0.39%	0.00%	−3.85%	2.77%	0.00%	3.97%

Bold values represent the best results in each group of experiments
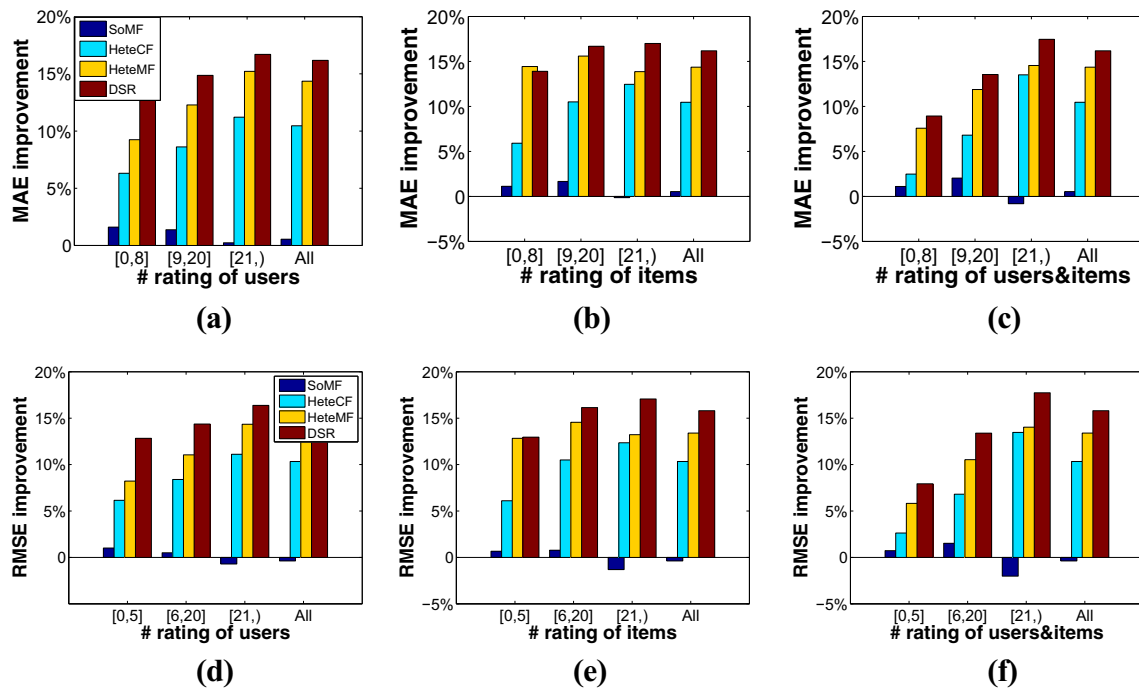
Fig. 3 MAE & RMSE improvement against PMF on various cold-start levels. **a** Users_MAE, **b** Items_MAE, **c** Users&Items_MAE, **d** Users_RMSE, **e** Items_RMSE, **f** Users&Items_RMSE

It is obvious that DSR with similarity constraint always achieves better performances than the basic matrix factorization (i.e., PMF). When integrating more meta-paths, DSR will has better performances (see UU* and MM*). Observing the experimental results of DSR with different meta-paths, we can find that the meaningful and dense similarity generated by meta-path will make DSR have better performances. For example, the UMTMU gets a slightly better result, while UGU gets the worst results. This is consistent with the real world, since users liking the same type of movies (i.e., UMTMU) always have similar preferences for movies. However, users in the same group (i.e., UGU) together with various reasons may have huge differences in movie preference. Another example is that the MTM path has better performance, while the MUUM path has the worst. It suggests that the similarity of movies with the same type (i.e., MTM) plays a leading role in movie recommendation, while the similarity of movies generated by the same movies watched by two friends is not such meaningful and useful.

6.6 Parameter study on α, β and γ

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

regularization on Douban data set. Considering the impact of γ in dual similarity regularization, we also take parameter study experiments on it.

Figure 5 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 optimal. When α and β are set the proper values (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 with the optimal results, 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.

Figure 6 shows the value of γ and also has a significant influence on experimental result. γ controls the radial intensity of Gaussian function on dual similarity regularization, which further controls the similarity of latent factors of two users (or items). Nineteen different values of γ are set (from 0.1 to 10) in our experiments, and optimal values are set for other parameters. We can see that when the value of γ is in [0.5, 4], it has good performance. When the value of γ tends to be quite large or small, the constraints on users or items
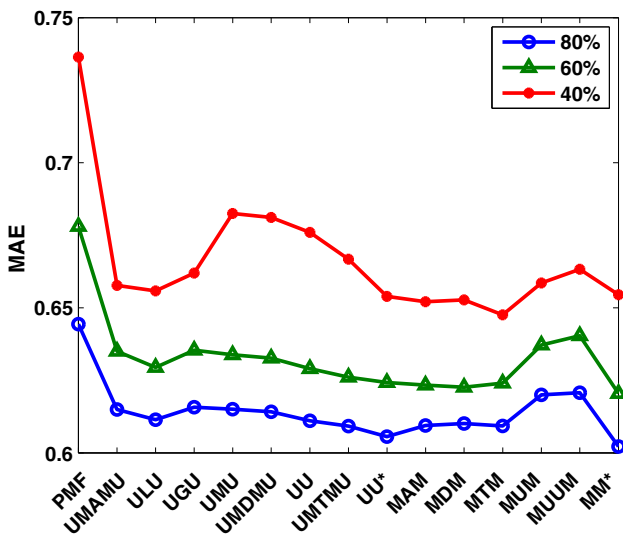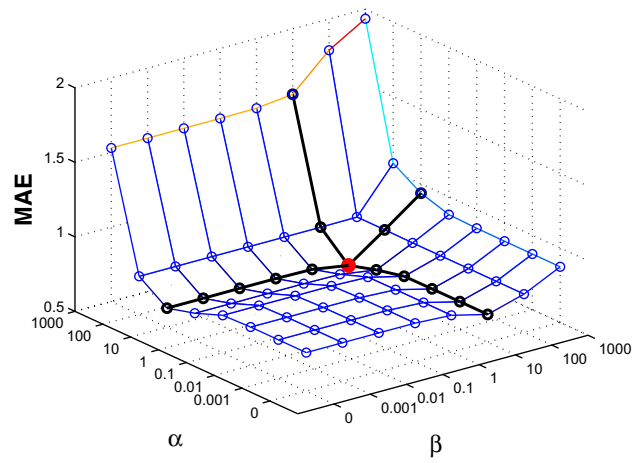
(a)
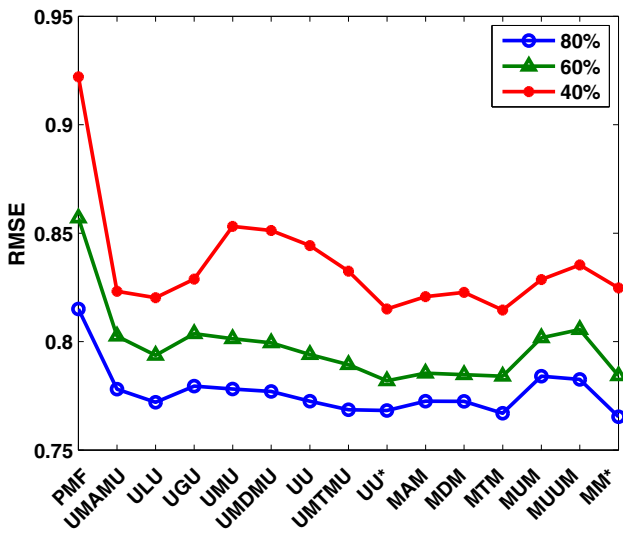
(b)

Fig. 4 Impact of different meta-paths. **a** MAE, **b** RMSE

(a)

(b)

Fig. 5 Parameter study on α and β. **a** MAE, **b** RMSE
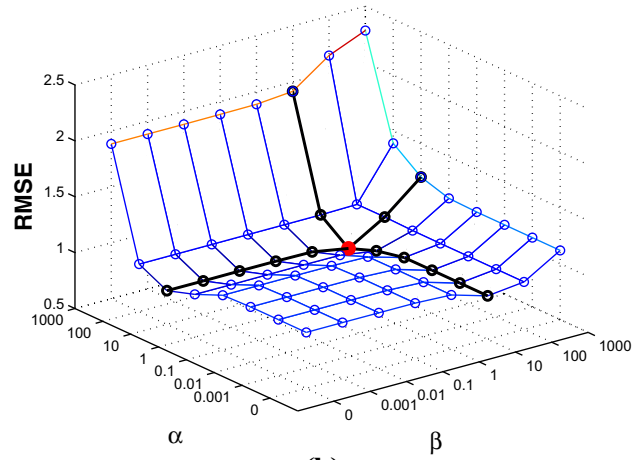

Fig. 6 Parameter study on γ

cannot work well, which leads to poor performance. Considering the experimental results on MAE and RMSE, we set $\gamma = 1$ in the former experiments.

7 Conclusions

In this paper, we analyzed the limitations of social regularization, including three aspects: First, the social regularization only takes constraint on users with their social relation; second, much available information of users and items can be used in HIN, while the social regularization d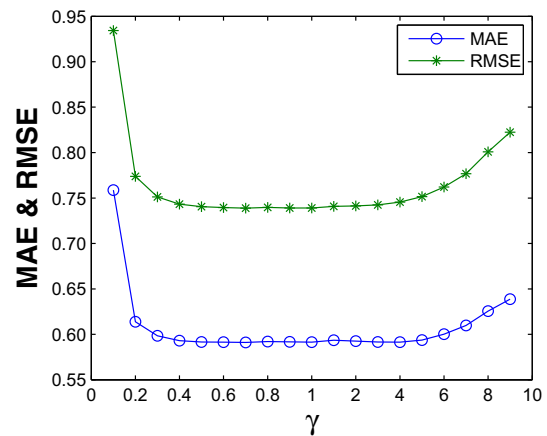id not make full use of it; third, the social regularization cannot work well for users with low similarity. Therefore, we design a dual similarity regularization whose basic idea is to enforce the constraints on both similar and dissimilar objects simul-

taneously. Then, encouraged by the successful of low-rank matrix factorization method, we employ the similarity regularization on the low-rank matrix factorization framework and proposed the DSR method, which integrates much useful information and takes constraint on users more effectively. Experiments on three real-world data sets validate the effectiveness of DSR, especially on alleviating the cold-start problem.

References

1. Balabanovic, M.: Content-based, collaborative recommendation. Commun. Acm **40**(3), 66–72 (1997)
2. BellogíN, R., Cantador, I., Castells, P.: A comparative study of heterogeneous item recommendations in social systems. Inf. Sci. **221**, 142–169 (2013)
3. Bernardi, L., Kamps, J., Kiseleva, J., Mller, M.J.: Solving the cold-start problem in recommender systems with social tags. Epl **92**(2), 28,002–28,007(6) (2010)
4. Burke, R., Vahedian, F., Mobasher, B.: Hybrid recommendation in heterogeneous networks. In: International Conference on User Modeling, Adaptation, and Personalization, pp. 49–60. Springer, New York (2014)
5. Cantador, I., Bellogin, A., Vallet, D.: Content-based recommendation in social tagging systems. In: RecSys, pp. 237–240 (2010)
6. Cao, L., Yu, P.S.: Non-iid recommendation theories and systems. Intell. Syst. IEEE **31**(2), 81–84 (2016)
7. Cplex, B.: Algorithms for non-negative matrix factorization. In: NIPS, pp. 556–562 (2010)
8. Diao, Q., Qiu, M., Wu, C.Y., Smola, A.J., Jiang, J., Wang, C.: Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars). In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 193–202 (2014)
9. Guo, Y., Wang, X., Xu, C.: Lbmf: Log-bilinear matrix factorization for recommender systems. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 502–513 (2016)
10. Herlocker, J.L., Konstan, J.A., Borchers, A., Riedl, J.: An algorithmic framework for performing collaborative filtering. In: SIGIR '99: Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, August 15–19, 1999, Berkeley, Ca, Usa, pp. 230–237 (1999)
11. Hong, L., Doumith, A.S., Davison, B.D.: Co-factorization machines: modeling user interests and predicting individual decisions in twitter. In: ACM International Conference on Web Search and Data Mining, pp. 557–566 (2013)
12. Jaiwei, H.: Mining heterogeneous information networks: the next frontier. In: SIGKDD, pp. 2–3 (2012)
13. Jamali, M., Lakshmanan, L.V.: Heteromf: recommendation in heterogeneous information networks using context dependent factor models. In: WWW, pp. 643–653 (2013)
14. Lao, N., Cohen, W.W.: Fast query execution for retrieval models based on path-constrained random walks. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 881–888 (2010)
15. Leroy, V., Cambazoglu, B.B., Bonchi, F.: Cold start link prediction. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 393–402. ACM (2010)
16. Li, F., Xu, G., Cao, L.: Coupled matrix factorization within non-iid context. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 707–719. Springer, New York (2015)
17. Luo, C., Pang, W., Wang, Z.: Hete-cf: social-based collaborative filtering recommendation using heterogeneous relations. In: ICDM, pp. 917–922 (2014)
18. Ma, H., King, I., Lyu, M.R.: Learning to recommend with social trust ensemble. In: SIGIR, pp. 203–210 (2011)
19. Ma, H., Yang, H., Lyu, M.R., King, I.: Sorec: Social recommendation using probabilistic matrix factorization. In: CIKM, pp. 931–940 (2008)
20. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I.: Recommender systems with social regularization. In: WSDM, pp. 287–296 (2011)
21. Salakhutdinov, R., Mnih, A.: Probabilistic matrix factorization. In: NIPS, pp. 1257–1264 (2012)
22. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: International Conference on World Wide Web, pp. 285–295 (2001)
23. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: SIGIR 2002: Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, August 11–15, 2002, Tampere, Finland, pp. 253–260 (2002)
24. Shi, C., Kong, X., Huang, Y., Yu, P.S., Wu, B.: Hetesim: A general framework for relevance measure in heterogeneous networks. In: IEEE Transactions on Knowledge and Data Engineering, pp. 2479–2492 (2014)
25. Shi, C., Li, Y., Zhang, J., Sun, Y., Yu, P.S.: A survey of heterogeneous information network analysis. Comput. Sci. **134**(12), 87–99 (2015)
26. Shi, C., Liu, J., Zhuang, F., Philip, S.Y., Wu, B.: Integrating heterogeneous information via flexible regularization framework for recommendation. In: Knowledge and Information Systems, pp. 1–25 (2016)
27. Shi, C., Zhang, Z., Luo, P., Yu, P.S., Yue, Y., Wu, B.: Semantic path based personalized recommendation on weighted heterogeneous information networks. In: The ACM International, pp. 453–462 (2015)
28. Shi, C., Zhou, C., Kong, X., Yu, P.S., Liu, G., Wang, B.: Heterecom: a semantic-based recommendation system in heterogeneous networks. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1552–1555. ACM, New York (2012)
29. Srebro, N., Jaakkola, T.: Weighted low-rank approximations. In: ICML, pp. 720–727 (2003)
30. 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)
31. Wang, C., Cao, L., Wang, M., Li, J., Wei, W., Ou, Y.: Coupled nominal similarity in unsupervised learning. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, pp. 973–978. ACM, New York (2011)
32. Wu, Y., Dubois, C., Zheng, A.X., Ester, M.: Collaborative denoising auto-encoders for top-n recommender systems. In: ACM International Conference on Web Search and Data Mining (2016)
33. Yang, C., Zhou, Y., Chen, L., Zhang, X., Chiu, D.M.: Social group based video recommendation addressing the cold-start problem. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 515–527 (2016)

34. Yang, X., Steck, H., Liu, Y.: Circle-based recommendation in online social networks. In: KDD, pp. 1267–1275 (2012)
35. 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)
36. Yu, X., Ren, X., Sun, Y., Gu, Q., Sturt, B., Khandelwal, U., Norick, B., Han, J.: Personalized entity recommendation: a heterogeneous information network approach. In: ACM International Conference on Web Search and Data Mining, pp. 283–292 (2014)
37. Yu, X., Ren, X., Sun, Y., Sturt, B., Khandelwal, U., Gu, Q., Norick, B., Han, J.: Recommendation in heterogeneous information networks with implicit user feedback. In: ACM Conference on Recommender Systems, pp. 347–350 (2013)
38. Zhang, D., Hsu, C.H., Chen, M., Chen, Q., Xiong, N., Lloret, J.: Cold-start recommendation using bi-clustering and fusion for large-scale social recommender systems. IEEE Trans. Emerg. Top. Comput. **2**(2), 239–250 (2014)
39. 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, 2008. WAIM'08, pp. 309–316. IEEE (2008)
40. Zheng, J., Liu, J., Shi, C., Zhuang, F., Li, J., Wu, B.: Dual similarity regularization for recommendation. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 542–554. Springer, New York (2016)