Exploiting Local and Global Social Context for Recommendation

Jiliang Tang, Xia Hu, Huiji Gao, Huan Liu
Computer Science and Engineering, Arizona State University, Tempe, AZ 85281
{Jiliang.Tang,Xia.Hu,Huiji.Gao,Huan.Liu}@asu.edu

Abstract

With the fast development of social media, the information overload problem becomes increasingly severe and recommender systems play an important role in helping online users find relevant information by suggesting information of potential interests. Social activities for online users produce abundant social relations. Social relations provide an independent source for recommendation, presenting both opportunities and challenges for traditional recommender systems. Users are likely to seek suggestions from both their local friends and users with high global reputations, motivating us to exploit social relations from local and global perspectives for online recommender systems in this paper. We develop approaches to capture local and global social relations, and propose a novel framework LOCABAL taking advantage of both local and global social context for recommendation. Empirical results on real-world datasets demonstrate the effectiveness of our proposed framework and further experiments are conducted to understand how local and global social context work for the proposed framework.

1 Introduction

With the increasing amount of information available, it becomes difficult for online users to find relevant information. Recommender systems, attempting to tackle the information overload problem by suggesting to online users the information that is potentially of interests, attract more and more attention in recent years [Sarwar *et al.*, 2001; Golbeck, 2006; Massa and Avesani, 2007; Koren, 2009]. Such systems are widely implemented in various domains including product recommendation at Amazon[1] and movie recommendation at Netflix[2].

The pervasive usage of social media allows users to participate in online activities which produce a large amount of social relations such as trust relations in Epinions[3]. Social relations provide an independent source of information about users beyond rating information. Social correlation theories such as homophily [Miller McPherson and Cook, 2001] and social influence [Marsden and Friedkin, 1993] indicate that there are correlations between two socially connected users, which can potentially be used to exploit social relations for recommender systems. The availability of social relations presents opportunities for traditional recommender systems.

Social relations can be viewed from different perspectives. An important classification of views of social relations is from local and global perspectives [Massa, 2007]. The former reveals the correlations among the user and his/her neighborhood while the latter reveals the reputation of the user in the whole social network. Users in the physical world are likely to ask for suggestions from their local friends while they also tend to seek suggestions from users with high global reputations, indicating both local and global views of social relations can be potentially exploited to improve the performance of online recommender systems.

There is recent work exploiting the local perspective of social relations for online recommender systems [Ma *et al.*, 2011; Jamali and Ester, 2010; Yang *et al.*, 2011; Tang *et al.*, 2012a; 2012b]. Exploiting local social context can reduce the number of cold-start users [Tang *et al.*, 2012a; 2012b]; hence, it can improve recommendation performance [Golbeck, 2006; Ma *et al.*, 2008; Jamali and Ester, 2009; Yang *et al.*, 2011; Jiang *et al.*, 2012]. Although the global perspective of social relations is widely exploited by various online applications such as identifying influential bloggers [Agarwal *et al.*, 2008], ranking online content [Weng *et al.*, 2010] and predicting trustworthiness of users [Massa, 2007], there is little work exploiting global social context for recommendation.

In this paper, we investigate how to exploit social relations from local and global perspectives in recommender systems and propose a novel framework taking advantage of both local and global social context for recommendation. Our contributions are summarized next.

- Introducing an alternative approach to exploit local social context and investigating its connections with existing representative approaches;
- Providing a principled way to exploit global social context for recommendation;
- Proposing a novel recommender system LOCABAL

[1] http://www.amazon.com/
[2] http://www.netflix.com
[3] http://www.epinions.com

which exploits local and global social context simultaneously for recommendation;

- Evaluating LOCABAL extensively using real-world datasets to understand the working of LOCABAL.

The rest of this paper is organized as follows. We introduce the details of our proposed framework LOCABAL in Section 2. In Section 3, we investigate the connections between our approach and existing approaches in terms of modeling local social context. We show empirical evaluation with discussion in Section 4. In Section 5, we present the conclusion with future work.

2 A Recommendation Framework with Local and Global Social Context

We first introduce notations used in this paper. Let $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ and $\mathcal{V} = \{v_1, v_2, \ldots, v_m\}$ be the sets of users and items respectively, where n is the number of users and m is the number of items. We assume that $\mathbf{R} \in \mathbb{R}^{n \times m}$ is the user-item rating matrix. If u_i gives a rating to v_j, $\mathbf{R}_{ij}$ is the rating score, otherwise we employ 0 to represent the unknown rating from u_i to v_j, i.e., $\mathbf{R}_{ij} = 0$. Let $\mathcal{O} = \{\langle u_i, v_j \rangle | \mathbf{R}_{ij} \neq 0\}$ be the set of known ratings and $\mathcal{T} = \{\langle u_i, v_j \rangle | \mathbf{R}_{ij} = 0\}$ be the set of unknown ratings. Users can establish social relations to each other. We use $\mathbf{T} \in \mathbb{R}^{n \times n}$ to denote user-user social relations where $\mathbf{T}_{ij} = 1$ if u_j has a relation to u_i and zero otherwise.

Before modeling social relations, we first introduce a state-of-the-art recommendation method based on matrix factorization as our basic model. Matrix factorization techniques have been widely employed for recommendation [Salakhutdinov and Mnih, 2008; Koren, 2008; Ma *et al.*, 2011] and they assume that a few latent patterns influence user rating behaviors and perform a low-rank matrix factorization on the user-item rating matrix. Let $\mathbf{u}_i \in \mathbb{R}^K$ and $\mathbf{v}_j \in \mathbb{R}^K$ be the user preference vector for u_i and item characteristic vector for v_j respectively, where K is the number of latent factors. Matrix factorization based recommender systems solve the following problem,

$$\min \sum_{\langle u_i, v_j \rangle \in \mathcal{O}} (\mathbf{R}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2 + \lambda(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2), \quad (1)$$

where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n] \in \mathbb{R}^{K \times n}$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m] \in \mathbb{R}^{K \times m}$. The terms $\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2$ are introduced to avoid over-fitting and the non-negative parameter λ is used to control the capability of $\mathbf{U}$ and $\mathbf{V}$.

There are several nice properties of matrix factorization methods: (1) simple optimization methods such as gradient based methods can be employed to find a well-worked optimal solution; (2) it has a nice probabilistic interpretation with Gaussian noise [Salakhutdinov and Mnih, 2008]; and (3) it is very flexible and enables us to integrate side information such as local and global social context introduced in the following subsections.

2.1 Exploiting Local Social Context

The local perspective of social relations reveals the correlations between users and their neighborhoods. Social correlation theories such as homophily [Miller McPherson and Cook, 2001] and social influence [Marsden and Friedkin, 1993] pave a way for us to exploit local social context for recommendation. Homophily indicates that users with similar tastes are more likely to be socially connected, and social influence suggests that users that are socially connected are more likely to share similar tastes. However, the low cost of social relation formation can lead to social relations with heterogeneous strengths (e.g., weak ties and strong ties mixed together) [Xiang *et al.*, 2010]. Since users with strong ties are more likely to share similar tastes than those with weak ties, treating all social relations equally is likely to lead to degradation in recommendation performance. Previous work demonstrated that exploiting social relations with equal weight cannot improve recommendation performance [Tang *et al.*, 2012a]. These observations suggest that we should consider heterogeneous strengths when exploiting local social context for recommendation.

In this paper, we simply use the rating cosine similarity to measure the social relation strength, although there are other more sophisticated measures in [Xiang *et al.*, 2010]. For the i-th user u_i, we use $\mathcal{N}_i = \{u_k | \mathbf{T}(i,k) = 1\}$ to denote his/her neighborhood. Let $\mathbf{S} \in \mathbb{R}^{n \times n}$ be the trust relation representation matrix where $\mathbf{S}_{ik}$ denotes the social relation strength between u_i and u_k.

For $u_k \in \mathcal{N}_i$, $\mathbf{S}_{ik}$ is calculated as the cosine similarity between the rating vectors of u_i and u_k while for $u_k \notin \mathcal{N}_i$, we set $\mathbf{S}_{ik}$ to 0. Therefore, $\mathbf{S}_{ik}$ can be formally defined as

$$\mathbf{S}_{ik} = \begin{cases} \frac{\sum_j \mathbf{R}_{ij} \cdot \mathbf{R}_{kj}}{\sqrt{\sum_j \mathbf{R}_{ij}^2} \sqrt{\sum_j \mathbf{R}_{kj}^2}} & \text{for } u_k \in \mathcal{N}_i, \\ 0 & \text{for } u_k \notin \mathcal{N}_i. \end{cases} \quad (2)$$

According to social correlation theories, the user preferences of two socially connected users are correlated thus we define the term to capture the local social context as

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} (\mathbf{S}_{ik} - \mathbf{u}_i^\top \mathbf{H} \mathbf{u}_k)^2, \quad (3)$$

In Eq. (3), for two socially connected users u_i and u_k, their preference vectors $\mathbf{u}_i$ and $\mathbf{u}_k$ are correlated through $\mathbf{H}$ controlled by their social strength $\mathbf{S}_{ik}$ where $\mathbf{H} \in \mathbb{R}^{K \times K}$ is the matrix to capture the user preference correlation. A large value of $\mathbf{S}_{ik}$, i.e., u_i and u_k with a strong connection, indicates that their preferences $\mathbf{u}_i$ and $\mathbf{u}_k$ should be tightly correlated via $\mathbf{H}$, while a small value of $\mathbf{S}_{ik}$ tells that $\mathbf{u}_i$ and $\mathbf{u}_k$ should be loosely correlated.

Eq.(3) can be rewritten to its matrix form as,

$$\min \|\mathbf{T} \odot (\mathbf{S} - \mathbf{U}^\top \mathbf{H} \mathbf{U})\|_F^2 \quad (4)$$

where $\odot$ is the Hadamard product where $(\mathbf{A} \odot \mathbf{B})_{ij} = \mathbf{A}_{ij} \times \mathbf{B}_{ij}$ for any two matrices $\mathbf{A}$ and $\mathbf{B}$ with the same size.

There is recent work also exploiting local social context under the matrix factorization framework for recommendation [Ma *et al.*, 2008; Jamali and Ester, 2010; Ma *et al.*, 2011]. We will study the connections between the proposed approach and representative existing methods to capture local social context in Section 3.

2.2 Exploiting Global Social Context

The global perspective of social relations reveals the reputation of a user in the whole social network [Massa, 2007]. User reputation is a sort of status that gives additional powers and capabilities in the online system [Massa, 2007], arousing various online applications such as predicting trustworthiness of users [Massa, 2007], identifying influential bloggers [Agarwal *et al.*, 2008], ranking online content [Weng *et al.*, 2010] and finding high-quality reviews [Lu *et al.*, 2010]. There are many algorithms to calculate the reputations of nodes in social networks according to their connections [Page *et al.*, 1999; Kleinberg, 1999] and we adopt one of the most popular algorithms PageRank [Page *et al.*, 1999] to compute the user reputation scores in this work. In detail, we first perform PageRank to rank users by exploiting the global perspective of social relations. We assume that $r_i \in [1, n]$ is the reputation ranking of u_i where $r_i = 1$ denotes that u_i has the highest reputation in the whole social network. Then we define user reputation score w_i as a function f of user reputation ranking r_i. The following definition of f works well in this work:

$$w_i = f(r_i) = \frac{1}{1 + log(r_i)}, \tag{5}$$

where the function f limits the value of the reputation score w_i within $[0, 1]$ and is a decreasing function of r_i, i.e., top-ranked users have high reputation scores.

In the physical world, user reputation plays an important role in recommendation and many companies employ people with high reputations to enhance consumers' awareness and understanding of their products. Seno and Lukas found that suggestions from people with high reputations positively affect a consumer's adoption of a brand [Seno and Lukas, 2007]. While in the online world, Massa found that ratings from users with high reputations are more likely to be trustworthy [Massa, 2007]. Therefore our solution to capture global social context is to weight the importance of user ratings according to their reputation scores as

$$\min \sum_{\langle u_i, v_j \rangle \in \mathcal{O}} w_i (\mathbf{R}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2, \tag{6}$$

where the importance of ratings from u_i is controlled by her/his reputation score w_i during the matrix factorization process. A large value of w_i, indicating the high reputation of u_i, will force $\mathbf{u}_i^\top \mathbf{v}_j$ to tightly fit the rating $\mathbf{R}_{ij}$ while $\mathbf{u}_i^\top \mathbf{v}_j$ will loosely approximate $\mathbf{R}_{ij}$ when w_i is small.

Eq.(6) can be also rewritten to its matrix form as

$$\min \|\mathbf{W} \odot (\mathbf{R} - \mathbf{U}^\top \mathbf{V})\|_F^2, \tag{7}$$

where $\mathbf{W} \in \mathbb{R}^{n \times m}$ is constructed as,

$$\mathbf{W}_{ij} = \begin{cases} \sqrt{w_i} & \text{if } \mathbf{R}_{ij} \neq 0 \\ 0 & \text{if } \mathbf{R}_{ij} = 0 \end{cases}. \tag{8}$$

2.3 Our Framework

In the above, we introduce our solutions to capture local and global social context mathematically. With these solutions, we propose a recommendation framework LOCABAL exploiting local and global social context simultaneously and the proposed framework is to solve the follow optimization problem

$$\begin{aligned} \min_{\mathbf{U},\mathbf{V},\mathbf{H}} \quad & \sum_{\langle u_i, v_j \rangle \in \mathcal{O}} w_i (\mathbf{R}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2 \\ & + \alpha \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} (\mathbf{S}_{ik} - \mathbf{u}_i^\top \mathbf{H} \mathbf{u}_k)^2 \\ & + \lambda (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{H}\|_F^2), \end{aligned} \tag{9}$$

where the reputation score of u_i w_i in the first term is used to exploit global social context and local social context is captured by the second term. The parameter α is introduced to control the contribution from local social context.

We use $\mathcal{J}$ to denote the objective function in Eq. (9) and it can be rewritten to its matrix form as

$$\begin{aligned} \mathcal{J} = & \|\mathbf{W} \odot (\mathbf{R} - \mathbf{U}^\top \mathbf{V})\|_F^2 + \alpha \|\mathbf{T} \odot (\mathbf{S} - \mathbf{U}^\top \mathbf{H} \mathbf{U})\|_F^2 \\ & + \lambda (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{H}\|_F^2) \end{aligned} \tag{10}$$

The derivations of $\mathcal{J}$ with respect to $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{H}$ are

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial \mathbf{U}} = & 2\big(- \mathbf{V}(\mathbf{W} \odot \mathbf{W} \odot \mathbf{R})^\top + \mathbf{V}(\mathbf{W} \odot \mathbf{W} \odot (\mathbf{U}^\top \mathbf{V}))^\top \\ & - \alpha \mathbf{H}^\top \mathbf{U}(\mathbf{T} \odot \mathbf{S}) - \alpha \mathbf{H} \mathbf{U} (\mathbf{T} \odot \mathbf{S})^\top + \lambda \mathbf{U} \\ & + \alpha \mathbf{H}^\top \mathbf{U} (\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H} \mathbf{U})) + \alpha \mathbf{H} \mathbf{U} (\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H} \mathbf{U}))^\top \big), \\ \frac{\partial \mathcal{J}}{\partial \mathbf{V}} = & 2\big(- \mathbf{U}(\mathbf{W} \odot \mathbf{W} \odot \mathbf{R}) \\ & + \mathbf{U}(\mathbf{W} \odot \mathbf{W} \odot (\mathbf{U}^\top \mathbf{V})) + \lambda \mathbf{V} \big), \\ \frac{\partial \mathcal{J}}{\partial \mathbf{H}} = & 2\big(- \lambda \mathbf{U}(\mathbf{T} \odot \mathbf{S}) \mathbf{U}^\top \\ & + \alpha \mathbf{U} (\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H} \mathbf{U})) \mathbf{U}^\top + \lambda \mathbf{H} \big) \end{aligned} \tag{11}$$

A local minimum of the objective function $\mathcal{J}$ in Eq. (9) can be obtained through a gradient decent optimization method, which usually works well for recommender systems [Koren, 2008; 2009]. The detailed algorithm is shown in Algorithm 1.

Algorithm 1 The Proposed Social Recommendation Framework LOCABAL

Input: The rating information $\mathbf{R}$, the social information $\mathbf{T}$, the number of latent factors k and α

Output: The user preference matrix $\mathbf{U}$ and the item characteristic matrix $\mathbf{V}$

1: Initialize $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{H}$ randomly
2: Construct $\mathbf{S}$ from $\mathbf{R}$ and $\mathbf{T}$
3: **while** Not convergent **do**
4:     Calculate $\frac{\partial \mathcal{J}}{\partial \mathbf{U}}$, $\frac{\partial \mathcal{J}}{\partial \mathbf{V}}$ and $\frac{\partial \mathcal{J}}{\partial \mathbf{H}}$
5:     Update $\mathbf{U} \leftarrow \mathbf{U} - \gamma_u \frac{\partial \mathcal{J}}{\partial \mathbf{U}}$
6:     Update $\mathbf{V} \leftarrow \mathbf{V} - \gamma_v \frac{\partial \mathcal{J}}{\partial \mathbf{V}}$
7:     Update $\mathbf{H} \leftarrow \mathbf{H} - \gamma_h \frac{\partial \mathcal{J}}{\partial \mathbf{H}}$
8: **end while**

In Algorithm 1, γ_u, γ_v and γ_h are learning steps, which are chosen to satisfy Goldstein Conditions [Nocedal and Wright, 1999]. After learning the user preference matrix $\mathbf{U}$ and the item characteristic matrix $\mathbf{V}$ via Algorithm 1, an unknown rating $\hat{\mathbf{R}}_{i'j'}$ from the user u_i' to the item v_j' will be predicted as $\hat{\mathbf{R}}_{i'j'} = \mathbf{u}_{i'}^\top \mathbf{v}_{j'}$.

3 Connections to Existing Approaches

Exploiting local social context under the matrix factorization framework for recommendation has attracted increasingly attention recently [Ma *et al.*, 2008; Jamali and Ester, 2010; Ma *et al.*, 2011]. Next we study the connections between our approach and two representative methods [Ma *et al.*, 2008; 2011] w.r.t. capturing local social context. The purpose of this study is two-fold - to investigate their main differences, and to seek the unique properties of our proposed approach.

In [Ma *et al.*, 2008], the local social context is captured as

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} (\mathbf{S}_{ik} - \mathbf{u}_i^\top \mathbf{z}_k)^2, \tag{12}$$

where $\mathbf{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n\} \in \mathbb{R}^{K \times n}$ is the factor-specific latent feature matrix [Ma *et al.*, 2008]. The underlying assumption of this approach is that the i-th user u_i should share the same user preference vector $\mathbf{u}_i$ in the rating space and the social space, which is different from that of our method based on social correlation theories. We develop the following theorem considering their connections.

Theorem 3.1 *Eq. (3) and Eq. (12) have the same form by factorizing* $\mathbf{z}_k$ *into* $\mathbf{H}\mathbf{u}_k$

Proof By factorizing $\mathbf{z}_k$ into $\mathbf{H}\mathbf{u}_k$ and plugging into Eq. (12), Eq. (3) and Eq. (12) are equivalent. □.

In [Ma *et al.*, 2011], social regularization is defined to capture local social context as

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} \mathbf{S}_{ik} \|\mathbf{u}_i - \mathbf{u}_k\|_2^2, \tag{13}$$

the intuition of social regularization is that two users with social relations should share the same or similar preferences. Therefore social regularization forces the preference vectors of two connected users to be close to each other, which is controlled by their social strength and we also develop the following theorem to investigate their connections.

Theorem 3.2 *Eq. (3) and Eq. (13) can be rewritten into the following form as*

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} -2\mathbf{S}_{ik}\mathbf{u}_i^\top \mathbf{A} \mathbf{u}_k + \mathbf{u}_i^\top \mathbf{B}_k \mathbf{u}_i + \mathbf{u}_k^\top \mathbf{C}_i \mathbf{u}_k \tag{14}$$

where for Eq. (3), $\mathbf{A}$, $\mathbf{B}_k$, *and* $\mathbf{C}_i$ *are defined as,*

$$\mathbf{A} = \mathbf{H},\ \mathbf{B}_k = \frac{1}{2}\mathbf{H}\mathbf{u}_k\mathbf{u}_k^\top \mathbf{H}^\top,\ \mathbf{C}_i = \frac{1}{2}\mathbf{H}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{H}, \tag{15}$$

while for Eq. (13), $\mathbf{A}$, $\mathbf{B}_k$, *and* $\mathbf{C}_i$ *are defined as,*

$$\mathbf{A} = \mathbf{I},\ \mathbf{B}_k = \mathbf{C}_i = \frac{1}{\mathbf{S}_{ik}} diag(\mathbf{S}_{ik}). \tag{16}$$

Proof By removing constants, Eq. (3) can be rewritten as

$$\min \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} -2\mathbf{S}_{ik}\mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k + \mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k \mathbf{u}_k^\top \mathbf{H}^\top \mathbf{u}_i. \tag{17}$$

It is easy to verify that $\mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k = \mathbf{u}_k^\top \mathbf{H}^\top \mathbf{u}_i$, and then we have

$$\mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k\mathbf{u}_k^\top \mathbf{H}^\top \mathbf{u}_i = \frac{1}{2}\mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k\mathbf{u}_k^\top \mathbf{H}^\top \mathbf{u}_i + \frac{1}{2}\mathbf{u}_k^\top \mathbf{H}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{H}\mathbf{u}_k. \tag{18}$$

By constructing $\mathbf{A}$, $\mathbf{B}_k$, and $\mathbf{C}_i$ according to Eq. (15), we can demonstrate that Eq. (3) can be rewritten into the form of Eq. (14). Similarly we can verify that Eq. (13) has the form of Eq. (14) by constructing $\mathbf{A}$, $\mathbf{B}_k$, and $\mathbf{C}_i$ based on Eq. (16), which completes the proof. □

The above theorems not only suggest that our proposed approach is different from existing ones but also indicate that it can be considered as a bridge to connect existing approaches.

4 Experiments

In this section, we conduct experiments to answer the following questions: (1) how does the proposed framework LOCABAL perform compared to the state-of-the-art recommender systems? and (2) how do local and global social context contribute to the proposed framework. After introducing experimental settings, we compare the performance of different recommender systems and then investigate the effects of local and global social context on the proposed framework, followed by parameter selection for LOCABAL.

4.1 Experimental Settings

Two datasets are chosen to evaluate our proposed framework, i.e., Epinions and Ciao[4], which are publicly available from the first author's homepage[5]. Users in Epinions and Ciao can rate products with scores from 1 to 5 and they can also establish social relations (i.e., trust relations) with others. Some statistics of these two datasets are presented in Table 1. The rating data is very sparse for both Epinions and Ciao. We compute the number of ratings from each user and the distribution suggests a power-law-like distribution - a few users contribute a large number of ratings while most users contribute few ratings.

Table 1: Statistics of the Datasets

	Epinions	Ciao
# of Users	21,882	7,252
# of Items	59,104	21,880
# of Ratings	632,663	183,749
Rating Density	0.0005	0.0012
# of Social Relations	348,197	110,536
Social Relation Density	0.0007	0.0021

For each dataset, we choose $x\%$ as the training set and the remaining $1 - x\%$ as the testing set where x is varied as $\{50, 70, 90\}$ in this paper. Two popular metrics, the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE), are chosen to evaluate the prediction performance.

[4]http://www.ciao.co.uk/

[5]http://www.public.asu.edu/∼jtang20/datasetcode/truststudy.htm

The metric RMSE is defined as

$$RMSE(\mathcal{T}) = \sqrt{\frac{\sum_{(u_i,v_j)\in\mathcal{T}} (\mathbf{R}_{ij} - \hat{\mathbf{R}}_{ij})^2}{|\mathcal{T}|}}, \tag{19}$$

where $\mathcal{T}$ is the set of ratings in the testing set, $|\mathcal{T}|$ is the size of $\mathcal{T}$ and $\hat{\mathbf{R}}_{ij}$ is the predicted rating from u_i to v_j. The metric MAE is defined as

$$MAE(\mathcal{T}) = \frac{1}{|\mathcal{T}|} \sum_{(u_i,v_j)\in\mathcal{T}} |\mathbf{R}_{ij} - \hat{\mathbf{R}}_{ij}|, \tag{20}$$

A smaller RMSE or MAE value means better performance. Note that previous work demonstrated that *small improvement in RMSE or MAE terms can have a significant impact on the quality of the top-few recommendation* [Koren, 2008].

4.2 Comparisons of Different Recommender Systems

We compare the proposed framework LOCABAL with the following representative recommender systems:

MF: This method performs matrix factorization on the user-item rating matrix as shown in Eq. (1) [Salakhutdinov and Mnih, 2008]. It only utilizes rating information.

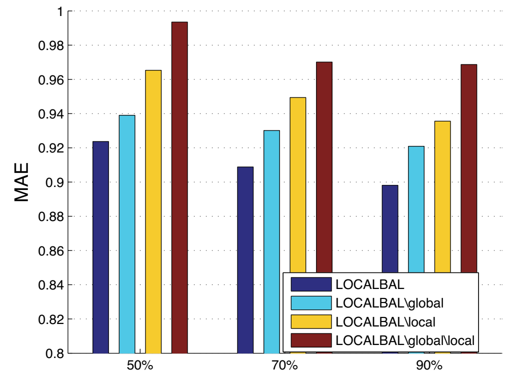
SoRec: This method is based on matrix factorization and exploits local social context by performing a factorization on the social matrix as shown in Eq. (12) [Ma *et al.*, 2008].

SoReg: this method is also based on matrix factorization and defines social regularization to capture local social context as shown in Eq. (13) [Ma *et al.*, 2011].

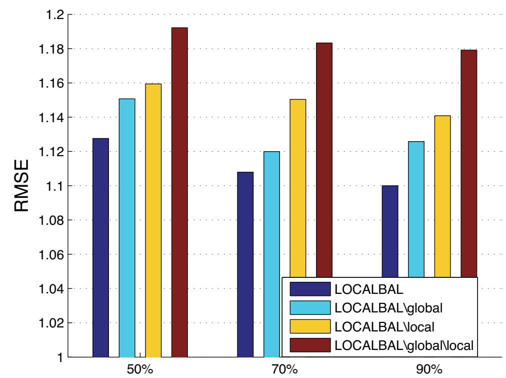
Note that we use cross-validation to determine parameters for all baseline methods. For LOCABAL, α is set to 0.7 and 0.5 for Ciao and Epinions respectively and the number of latent factors K is set to 10 for both datasets. More details about parameter selection for LOCABAL will be discussed later. According to empirical experience, λ is set to 1. The comparison results are demonstrated in Table 2 and we have the following observations,

- Exploiting social context can significantly improve recommendation performance in terms of both RMSE and MAE. For example, **SoRec**, **SoReg** and LOCABAL obtain 3.36%, 5.56%, 7.03% relative improvement in terms of MAE in Epinions with 50% as the training set respectively, compared to **MF**.
- Our proposed framework LOCABAL always obtains the best performance. Compared to the second best performance of baseline methods, on average, LOCABAL gains 0.023 and 0.0186 absolute RMSE improvement in Ciao and Epinions, respectively. The major reason is that LOCABAL exploits local and global social context simultaneously. More details about the effects of these two types of social context on the performance of LOCABAL will be discussed in the following subsection.

With these observations, we can draw an answer to the first question - our proposed framework outperforms the state-of-the-art recommender systems by exploiting local and global social context simultaneously.



(a) *MAE*



(b) *RMSE*

Figure 1: The Performance of Variants of Our Framework in Epinions.

4.3 The Impact of Global and Local Social Context on LOCABAL

The experimental results in the above subsection show that our proposed framework LOCABAL outperforms the state-of-the-art recommender systems. In this subsection, we investigate the effects of local and global social context on the proposed framework LOCABAL to answer the second question. In detail, we eliminate the effect of global and local social context systematically from LOCABAL by defining the following variants of LOCABAL,

- *LOCABAL$\backslash$global* - Eliminating the effect of global social context by setting $w_i = 1$ when $\mathbf{R}_{ij} \neq 0$ in Eq. (9).
- *LOCABAL$\backslash$local* - Eliminating the effect of local social context by setting $\alpha = 0$ in Eq. (9).
- *LOCABAL$\backslash$global$\backslash$local* - Eliminating the effects of both local and global social context by setting $\alpha = 0$ and setting $w_i = 1$ when $\mathbf{R}_{ij} \neq 0$ in Eq. (9). Actually this variant is the **MF** recommendation framework as defined in Eq. (1).

The comparison results in Epinions are shown in Figures 1(a) and 1(b) for MAE and RMSE, respectively. Note

Table 2: Comparison of Different Recommender Systems

Datasets	Training Set	Metrics	Algorithms			
			MF	**SoRec**	**SoReg**	LOCABAL
Ciao	50%	MAE	0.9927	0.9619	0.9552	0.9356
		RMSE	1.1742	1.1375	1.1291	1.1088
	70%	MAE	0.9715	0.9446	0.9328	0.9234
		RMSE	1.1478	1.1140	1.1097	1.0861
	90%	MAE	0.9614	0.9433	0.9232	0.9076
		RMSE	1.1384	1.1028	1.0999	1.0758
Epinions	50%	MAE	0.9935	0.9574	0.9383	0.9237
		RMSE	1.1922	1.1581	1.1479	1.1276
	70%	MAE	0.9701	0.9480	0.9296	0.9088
		RMSE	1.1833	1.1482	1.1277	1.1079
	90%	MAE	0.9687	0.9397	0.9188	0.8981
		RMSE	1.1791	1.1387	1.1170	1.1000

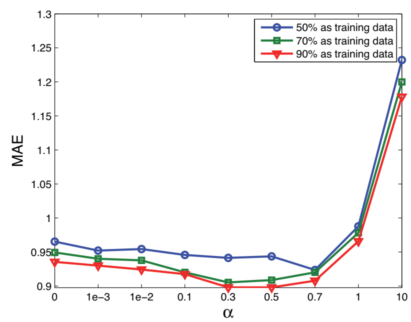
that we only show the results in Epinions since we has similar observations in Ciao. When eliminating the effect of local social context from the proposed framework, the performance of *LOCABAL\global* degrades. We have the similar observation for *LOCABAL\local* when eliminating the effect of global social context. For example, compared to LOCABAL, on average, *LOCABAL\global* and *LOCABAL\local* have 3.46% and 1.83% relative performance reduction in terms of RMSE, respectively. When eliminating the effects of both global and local social context, *LOCABAL\global\local* obtains worse performance than both *LOCABAL\global* and *LOCABAL\local*, suggesting that local and global social context contain complementary information to each other for recommendation.

With the evidence from Figure 1, we can answer the second question - both local and global social context can help improve the recommendation performance.

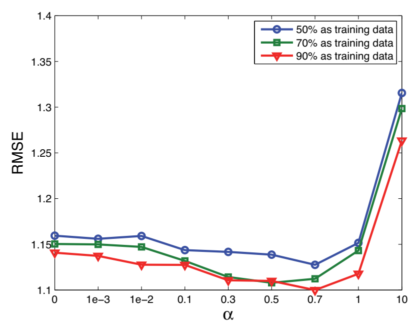
4.4 Parameter Selection

There is an important parameter for the proposed framework LOCABAL α, controlling the contribution from local social context. In this subsection, we investigate how changes of α affect the performance of LOCABAL. We vary the value of α as {0,1e-3,1e-2,0.1,0.3,0.5,0.7,1,10}. The results in Epinions w.r.t. RMSE and MAE are demonstrated in Figures 2(a) and 2(b), respectively. Since we have similar observations in Ciao, we only show the results in Epinions to save space.

With increasing values of α, we put more and more weight on local social context and we have the following observations: (1) the performance first increases, suggesting that integrating local social context can improve the recommendation performance; (2) when α is from 0.3 to 0.7, the performance is not sensitive to α; and (3) when α is from 1 to 10, the performance reduces dramatically. A large value of α will lead to social relations dominating the user preference learning process in Eq. (9), resulting in inaccurate estimates of the user preference matrix $\mathbf{U}$ and the item characteristic matrix $\mathbf{V}$. For example when $\alpha \to \infty$, the user preference matrix $\mathbf{U}$ is only learnt from social relations and the item characteristic matrix $\mathbf{V} = 0$



(a) *MAE*



(b) *RMSE*

Figure 2: The Performance of the Proposed Framework vs α in Epinions.

5 Conclusion and Future Work

The availability of social relations presents both challenges and opportunities for traditional recommender systems. In this paper, we investigate how to exploit local and global social context for recommendation. To capture local social context, we force that the user preferences of two socially connected users are correlated as suggested by social correlation theories and we also study the connections between our proposed approach and existing approaches. Ratings from users with high reputations are more likely to be trustworthy; therefore, to capture global social context, we use user reputation scores to weight the importance of their ratings. With these solutions, we propose a framework LOCABAL to integrate local and global social context for recommendation. Experimental results on real-world data sets show that the proposed framework LOCABAL outperforms representative social recommender systems. Further experiments are conducted to understand the working of LOCABAL.

There are several directions to investigate in the future. First, as user preferences and social relations might evolve, incorporating temporal information into social recommender systems will be an interesting direction [Tang *et al.*, 2012b]. Second, most existing social recommender systems exploit only positive social relations between users such as friendships and trust relations. However users can also have negative social relations such as distrust and dislike, then we can investigate the effect of negative social relations on social recommender systems [Yang *et al.*, 2012].

Acknowledgments

We thank the anonymous reviewers for their useful comments. The work is, in part, supported by ARO (#025071) and NSF (#IIS-1217466).

References

[Agarwal *et al.*, 2008] N. Agarwal, H. Liu, L. Tang, and P.S. Yu. Identifying the influential bloggers in a community. In *WSDM*, 2008.

[Golbeck, 2006] J. Golbeck. Generating predictive movie recommendations from trust in social networks. *Trust Management*, pages 93–104, 2006.

[Jamali and Ester, 2009] M. Jamali and M. Ester. Trustwalker: a random walk model for combining trust-based and item-based recommendation. In *KDD*,2009.

[Jamali and Ester, 2010] M. Jamali and M. Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In *RecSys*, 2010.

[Jiang *et al.*, 2012] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang. Social contextual recommendation. In *CIKM*. 2012.

[Kleinberg, 1999] J.M. Kleinberg. Authoritative sources in a hyperlinked environment. *JACM*, 1999.

[Koren, 2008] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *KDD*,2008.

[Koren, 2009] Y. Koren. Collaborative filtering with temporal dynamics. In *KDD*, 2009.

[Lu *et al.*, 2010] Y. Lu, P. Tsaparas, A. Ntoulas, and L. Polanyi. Exploiting social context for review quality prediction. In *WWW*,2010.

[Ma *et al.*, 2008] H. Ma, H. Yang, M.R. Lyu, and I. King. Sorec: social recommendation using probabilistic matrix factorization. In *CIKM*, 2008.

[Ma *et al.*, 2011] H. Ma, D. Zhou, C. Liu, M.R. Lyu, and I. King. Recommender systems with social regularization. In *WSDM*, 2011.

[Marsden and Friedkin, 1993] P.V. Marsden and N.E. Friedkin. Network studies of social influence. *Sociological Methods and Research*, 22(1):127–151, 1993.

[Massa and Avesani, 2007] P. Massa and P. Avesani. Trust-aware recommender systems. In *RecSys*, 2007.

[Massa, 2007] P. Massa. A survey of trust use and modeling in real online systems. *Trust in E-services: Technologies, Practices and Challenges*, 2007.

[Miller McPherson and Cook, 2001] Lynn Smith-Lovin Miller McPherson and James M Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444, 2001.

[Nocedal and Wright, 1999] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer verlag, 1999.

[Page *et al.*, 1999] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. 1999.

[Zhang *et al.*, 2012] Q. Zhang, and B. Li. Mining discriminative components with low-rank and sparsity constraints for face recognition. *KDD*, 2012.

[Salakhutdinov and Mnih, 2008] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. *NIPS*, 2008.

[Sarwar *et al.*, 2001] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *WWW*, 2001.

[Seno and Lukas, 2007] D. Seno and B.A. Lukas. The equity effect of product endorsement by celebrities: A conceptual framework from a co-branding perspective. *European Journal of Marketing*, 2007.

[Tang *et al.*, 2012a] J. Tang, H. Gao, and H. Liu. mTrust: Discerning multi-faceted trust in a connected world. In *WSDM*, 2012.

[Tang *et al.*, 2012b] J. Tang, H. Gao, H. Liu, and A. Das Sarma. eTrust: Understanding trust evolution in an online world. In *KDD*, 2012.

[Weng *et al.*, 2010] J. Weng, E.P. Lim, J. Jiang, and Q. He. Twitterrank: finding topic-sensitive influential twitterers. In *WSDM*, 2010.

[Xiang *et al.*, 2010] Rongjing Xiang, Jennifer Neville, and Monica Rogati. Modeling relationship strength in online social networks. In *WWW*, 2010.

[Yang *et al.*, 2011] Zi Yang, Keke Cai, Jie Tang, Li Zhang, Zhong Su, and Juanzi Li. Social context summarization. In *SIGIR*, 2011.

[Yang *et al.*, 2011] Shuang-Hong Yang, Bo Long, Alex Smola, Narayanan Sadagopan, Zhaohui Zheng, and Hongyuan Zha. Like like alike: joint friendship and interest propagation in social networks. In *WWW*, 2011.

[Yang *et al.*, 2012] Shuang-Hong Yang, Alexander J Smola, Bo Long, Hongyuan Zha, and Yi Chang. Friend or frenemy?: predicting signed ties in social networks. In *SIGIR*,2012.