Recommender Systems with Characterized Social Regularization

Tzu-Heng Lin, Chen Gao, Yong Li
Beijing National Research Center for Information Science and Technology
Department of Electronic Engineering
Tsinghua University
lzhbrian@gmail.com,gc16@mails.tsinghua.edu.cn,liyong07@tsinghua.edu.cn

ABSTRACT

Social recommendation, which utilizes social relations to enhance recommender systems, has been gaining increasing attention recently with the rapid development of online social network. Existing social recommendation methods are based on the fact that users preference or decision is influenced by their social friends' behaviors. However, they assume that the influences of social relation are always the same, which violates the fact that users are likely to share preference on diverse products with different friends. In this paper, we present a novel CSR (short for **C**haracterized **S**ocial **R**egularization) model by designing a universal regularization term for modeling variable social influence. Our proposed model can be applied to both explicit and implicit iteration. Extensive experiments on a real-world dataset demonstrate that CSR significantly outperforms state-of-the-art social recommendation methods.

CCS CONCEPTS

• **Information systems → Social recommendation**; **Recommender systems**;

KEYWORDS

Recommender Systems, Social Recommendation, Characterized Social Regularization, Matrix Factorization

1 INTRODUCTION

Traditional recommender systems that learn user's preference from single domain's data often suffer from data sparsity and cold start problem. To overcome them, social recommendation utilizes user's social relations as auxiliary information to aid learning preference. It relies on the fact that users preference is similar or influenced by their social friends. Over the past recent years, extensive researches

have been working on how to utilize social information to improve recommendation performance [1–6, 8–10].

Despite their effectiveness, we argue that all the existing social recommendation models suffer from a limitation: the social relation between users influences their preference on all items, which violates the fact that a user shares preference on an item with only a part of friends in real life. For example, a user may have similar preferences on book with her schoolmates or workmates, while have similar tastes on food with her family members.

To solve this shortcoming of all existing methods for social recommendation, we propose a Characterized Social Regularization (CSR) model that considers user's various similarities with different friends. Briefly, our method models the item-specific preference similarity of friends to efficiently and accurately learn the social influence on preferences. Specifically, we design a general item-specific regularization term in objective function, which can be easily adapted to most existing recommendation models. Through this way, we address the limitation of existing methods, and make models more effective and expressive in social recommendation.

In summary, the contributions of this paper are as follows:

- To the best of our knowledge, we are the first to model the characteristics of social relations in the field of social recommendation.
- We introduce a novel CSR method that effectively models the characteristics of social relations' influence, and prove that existing social recommendation models can be interpreted as a simplest special case of our CSR model.
- Extensive experiments on a real-world dataset show that our method CSR outperforms the state-of-the-art solutions.

2 PRELIMINARIES

We first formulate the task of social recommendation, and then introduce the existing social regularization models by highlighting the limitations to motivate our work.

2.1 Problem Formation

In social recommendation, social relations between users are considered as auxiliary information to help learning users' preferences. The target is the same with traditional recommendation tasks, *i.e.*, estimating the unobserved values of the interaction matrix $\mathbf{Y}_{M \times N} = \{y_{ui}\}$ with M and N denoting the number of users and items. Specifically, for explicit interactions, y_{ui} of user u and item i is a continuous score; while for implicit interactions, it is a binary value. We further use $\mathbf{U}$ and $\mathbf{V}$ to denote set of user and item, respectively, and denote the social relations between users as $\mathbf{S} = \{\mathbf{s}_{uu'} | u, u' \in \mathbf{U}\}$. It is a vector representing various relationship between two users u and u', such as friends, following, followed, and etc. In the existing works, this vector is 1-dimension, which is

a constant value 1 or a continuous value between 0 and 1 standing for social strength. Then, the task of social recommendation is to estimate the unobserved values in interaction matrix $\mathbf{Y}$ with the help of social relation $\mathbf{S}$.

2.2 Social Regularization

Regularization, as a widely used technique in matrix factorization approach, sets some limitations to latent vectors through a added regularization term in objective function. Social regularization is a regularization term considering social relation $\mathbf{S}$. We denote the added term in social regularization as $Social(\Omega, \mathbf{S})$, where Ω denotes parameters of latent model. Now, we discuss two mainstream methods utilizing social regularization.

Distance to weighted sum of friends [2, 3]. This method aims to minimize the distance of a user's latent vector with the weighted sum of the connected users' vector as possible. The added social regularization term is expressed as follows:

$$Social(\Omega, \mathbf{S}) = \sum_{u \in \mathbf{U}} ||\mathbf{p}_u - \sum_{s_{uu'} \in \mathbf{S}} s_{uu'} \mathbf{p}_{u'}||_2^2, \qquad (1)$$

where $s_{uu'}$ is a continuous value between 0 and 1 standing for social strength and $\mathbf{p}$ denotes users' latent vector. This method is used in [3] for explicit interactions and in [2] for implicit interactions, respectively.

Sum of weighted distance to friends [1, 6]. This method aims at minimizing the sum of weighted distance between latent vector of a user and her connected users' vector, in which the weight depends on the social strength. The added social regularization term is expressed as follows,

$$Social(\Omega, \mathbf{S}) = \sum_{u \in \mathbf{U}} \sum_{s_{uu'} \in \mathbf{S}} s_{uu'} ||\mathbf{p}_u - \mathbf{p}_{u'}||_2^2, \qquad (2)$$

where symbols are consistent with (1). This method is used in [6] for explicit interactions and in [1] for implicit interactions, respectively.

As discussed earlier in the Introduction, all these existing methods assume $\mathbf{s}$ is constant for a given pair (u, u'), ignoring the fact that the connection $\mathbf{s}$ may vary for interactions with different products. Specifically, preference of u and u' may be very similar on some products, while may be totally different on other products. As a result, social relations' influences on users' preference should be variable on diverse products. To address the shortcoming of existing methods, our proposed solution takes the variable influences into account by modeling different impact on diverse products.

3 OUR CSR SOLUTION

3.1 Learning from Interactions

Since social relation is regarded as side information to help improving recommendation, competitive collaborative filtering model to learn from historical interactions is a basic solution. Here, we rely on Matrix Factorization (MF), a widely used latent factor model in recommender system. Let $\mathbf{P} \in \mathbb{R}^{K \times M}$ and $\mathbf{Q} \in \mathbb{R}^{K \times N}$ be latent user and item features matrices, with column vectors $\mathbf{p}_u$ and $\mathbf{q}_i$ representing K-dimensional user-specific and item-specific latent feature vectors of users u and item i, respectively. Since the target of social recommendation is to estimate the unobserved values in interaction matrix $\mathbf{Y}$, MF model tries to decompose $\mathbf{Y}$ as $\mathbf{Y} = \mathbf{P}^\mathsf{T} \mathbf{Q}$.

Then the objective function of MF model can be denoted as:

$$\min_{\mathbf{P}, \mathbf{Q}} \mathcal{L}(\mathbf{Y}, \mathbf{P}, \mathbf{Q}) = \sum_{u=1}^{M} \sum_{i=1}^{N} (\mathbf{Y}_{ui} - \mathbf{p}_u^\mathsf{T} \mathbf{q}_i)^2, \qquad (3)$$

where L_2 regularization terms of latent matrices $\mathbf{P}$ and $\mathbf{Q}$ are omitted for simplification.

3.2 Characterized Social Regularization

As introduced in Section 2.2, regularization is extensively used in social recommendation. In our model, we also utilize regularization technique to constrain the learning of MF model. The motivation of our CSR model is to use social relations to constrain learning from interactions, which is the same with existing methods [1–3, 6]. Then, the objective function of MF models considering social regularization can be denoted as:

$$\min_{\mathbf{P}, \mathbf{Q}} \mathcal{L}(\mathbf{Y}, \mathbf{P}, \mathbf{Q}, \mathbf{S}) = \sum_{u=1}^{M} \sum_{i=1}^{N} (\mathbf{Y}_{ui} - \mathbf{p}_u^\mathsf{T} \mathbf{q}_i)^2 + \lambda_s Social(\mathbf{P}, \mathbf{Q}, \mathbf{S}) \quad (4)$$

where λ_s is a positive hyper-parameter controlling the weight of social information compared with interaction data.

Dimension-weighted Distance. In latent factor model for MF, embedding vector for a user encodes his/her latent interest and embedding vector for an item encodes its features. Considering social influence, a user's interest is similar with his/her friends, which means the embedding vector of user is similar with that of friends. This motivates the two mainstream social regularization methods introduced in Section 2.2, which simply utilize Euclidean distance to build $Social(\mathbf{P}, \mathbf{Q}, \mathbf{S})$ to limit similarity.

However, from the perspective of representation learning, $\mathbf{q}_{ik}$, the k-th dimension in the K-dimensional item latent vector, may represent k-th feature of item i; similarly $\mathbf{p}_{uk}$ may users' preference on the k-th item feature. As mentioned in the Introduction, a user may share different preference on various products with different friends. Then we can assume two users have similar preference on an item i is due to some dimensions' of item features. In other words, embedding vector of this item has a larger value in those dimensions. Therefore, we design a dimension-weighted distance rather than a simple Euclidean distance to guarantee the similarity of social-connected users. For k-th dimension of relationship between user u and one of his/her friends u', we use $s_{uu'}^k$ to denote the k-th weight, then we obtain a weight vector: $\mathbf{s}_{uu'}$. Based on this vector, the characterized regularization term can be denoted as:

$$Social(\mathbf{P}, \mathbf{Q}, \mathbf{S}) = \sum_{(u, u') \in \mathbf{S}} ||(\mathbf{p}_u - \mathbf{p}_u') \circ \mathbf{s}_{uu'}||_2^2, \qquad (5)$$

where $\circ$ means a Hadamard (element-wise) product.

Product-sharing Based Social Relations Although the designed social relations, weight vector $\mathbf{s}_{uu'}$, can help modeling social influence theoretically, it is not possible for them to be learned automatically, directly optimizing the above objective function will result in a trivial all zero solution for $\mathbf{s}_{uu'}$. In practice, users exchange opinions on various kinds of products, such as movies, news, etc., through sharing with friends. Therefore, to some extent, a sharing behavior or discussion on item i between user u and a friend u' is a social relation accompanied by item i. Thus, we can infer that u and u' share tastes on item i and those items similar with i. Since the embedding vector of shared item $\mathbf{q}_i$ represents

(a) Existing models   (b) Our CSR model
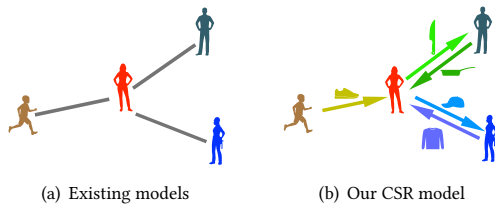
Figure 1: Comparison of existing models and our CSR model

item features, it can be straightly used as weight vector $\mathbf{s_{uu'}}$. The weight vector can also be built from other data or information of social domain certainly. This paper only focuses on the general framework, and leaves the specific design as future work.

Thus, we formulate our proposed social regularization term as:

$$Social(\mathbf{P}, \mathbf{Q}, \mathbf{S}) = \sum_{(u,u',i) \in \mathcal{D}} ||(\mathbf{p}_u - \mathbf{p}'_u) \circ \mathbf{q}_i||_2^2, \qquad (6)$$

where $\mathcal{D}$ is the triplets of product-sharing logs, which is utilized to build social relations $\mathbf{S}$, and $(u, u', i) \in \mathcal{D}$ means user u shows preference on item i with his/her friend u'. Then the objective function of our CSR model is:

$$\min_{\mathbf{P}, \mathbf{Q}} \mathcal{L}(\mathbf{Y}, \mathbf{P}, \mathbf{Q}, \mathbf{S}) = \sum_{u=1}^{M} \sum_{i=1}^{N} (\mathbf{Y}_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 \\ + \lambda_s \sum_{(u,u',i) \in \mathcal{D}} ||(\mathbf{p}_u - \mathbf{p}'_u) \circ \mathbf{q}_i||_2^2. \qquad (7)$$

3.3 Training

We optimize parameters with stochastic gradient descent (SGD), and implement it on Tensorflow[1], which provides the function of automatic differentiation. So here we omit the derivation of the objective function. We introduce our model in a point-wise manner for explicit interaction data, and for implicit interaction data, in fact, our CSR model can also be trained in a pair-wise manner [7]. Therefore, our proposed model can be adapted to these two kinds of interaction data just by switching the training manner.

3.4 Discussions

Here, we summary some desirable properties of CSR. Firstly, it is obviously that if we set the value of every dimension of $\mathbf{s}$ to be 1 in (5), then our proposed CSR model is transferred to the methods discussed in Section 2.2. Thus, existing social regularization methods can be taken as a special case of our CSR model. Secondly, the regularization term in our CSR model can be adapted to the objective function of most social recommendation models, not only for the MF model, which means the model learning from historical interaction can be diverse. Furthermore, this makes our CSR model able to handle complicated tasks and various interaction data.

4 EXPERIMENTS

In this section, we conducted extensive experiments on a real-world dataset to answer the following three research questions:

- **RQ1:** How does our proposed CSR model perform as compared with state-of-the-art social recommendation methods?
- **RQ2:** How do the key hyper-parameters affect performance of our CSR model?

[1]https://www.tensorflow.org/

Table 1: Statistics of our utilized Beibei dataset

# Users	# Items	# Purchases	# Shares
337	553	2,572	476

- **RQ3:** How does number of users' social relations affect recommendation performance of our CSR model?

4.1 Experimental Settings

Dataset   We experimented with a dataset collected from Beibei (https://www.beibei.com), which is the largest E-commerce platform for maternal and infant products in China. This dataset includes user logs from 2017/06/01 to 2017/06/30. This dataset is made up of two parts: purchase record containing users' purchase log and sharing record containing users' item-sharing logs. The statistics of the dataset are summarized in Table 1. Note that, here we learn CSR model based on pair-wise loss [7] for better performance. Due to the need for sharing data, no publicly available dataset are suitable for our methods. An interesting future work would be discovering more possible approaches to build explicit social relations.

Evaluation Metrics   To evaluate the performance, we adopted the *leave-one-out* evaluation method with the following metrics.

- **HR@K:** Hit Rate
- **NDCG@K:** Normalized Discounted Cumulative Gain

Baselines   We compared our CSR model with six baseline methods can be used in our implicit dataset. The first group contains methods only utilizing interaction data.

- **Random (Rand)** This method randomly orders the items.
- **ItemPop** This method ranks items base on their popularity, as judged by the number of historical interactions.
- **BPR [7]** This is a basic method based on pairwise learning.

The second group contains social recommendation methods.

- **SocialBPR [2]** This method optimizes a pairwise loss with a distance-to-weighted-sum-of-friends social regularization term.
- **UGPMF [1]** This method optimizes a pairwise loss with a sum-of-weighted-distance-to-friends social regularization term.
- **SBPR [11]** This method assumes that users tend to assign higher ranks to items that their friends prefer.

Parameter Settings   Following [3, 5, 6], we set the weights of regularization terms λ_P and λ_Q for P and Q to a trivial value 0.01. To make the experiments reasonable, we also use grid search on learning rate α, dimensionality K and the weight of social regularization term λ_s to find the best performance for each models. In the following sections, we report the best parameter settings for each model.

4.2 Performance Comparison (RQ1)

The experiment results are shown in Table 2. Following findings are observed.

- Rand and ItemPop achieve the worst performance, which indicates the power of MF model in modeling personalized behaviors.
- Generally, models with social information (SBPR, SocialBPR, CSR) outperforms the one without it (BPR) in terms of HR@10, NDCG@10. However, due to the sparsity of our social data, some methods (e.g. UGPMF) fail to utilize social information well.
- We can observe that CSR significantly and consistently outperforms the best baseline methods in terms of HR@5, NDCG@5, HR@10, NDCG@10 with a relative improvement of 11.19%, 16.21%, 2.85%, 8.63%, respectively. This justifies the expressiveness and effectiveness of our model.

Table 2: Performance comparison

Method	HR@5	NDCG@5	HR@10	NDCG@10
Rand	0.0119	0.0082	0.0297	0.0143
ItemPop	0.0623	0.0353	0.0920	0.0450
BPR [7]	0.0742	0.0401	0.1246	0.0556
SBPR [11]	0.0677	0.0357	0.1294	0.0562
UGPMF [1]	0.0682	0.0371	0.1231	0.0537
SocialBPR [2]	0.0742	0.0397	0.1403	0.0626
Our CSR	**0.0825**	**0.0466**	**0.1443**	**0.0680**
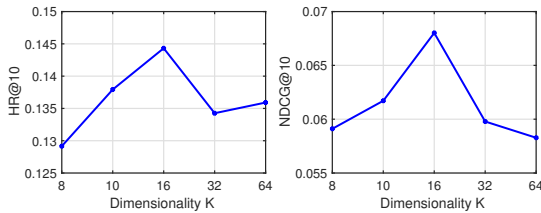
Figure 2: Performance with different dimensionality K
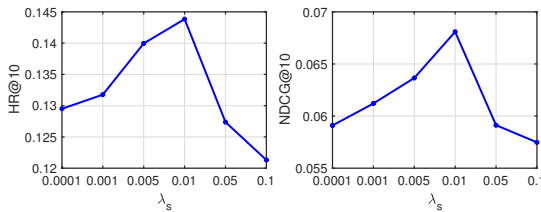
Figure 3: Performance with different weight λ_s

4.3 Parameter Study of CSR (RQ2)

Impact of Dimensionality K Intuitively, increasing the dimensionality of latent feature vectors can make models more expressive. However, on the other hand, training models with a larger K needs more data. As seen in Figure 2, in our experiment, we find that a middle size $K = 16$ can have the best performance.

Impact of weight λ_s The weight of social regularization term λ_s controls the influence of social data compared with the interaction data. A larger value of λ_s gives more contribution of the social sharing information. However, it may also cause social term dominates the total loss, and do harm to the performance with less utilization of the interaction data. We compare the different settings of λ_s in Figure 3 and find the optimal value for λ_s is 0.01 in our experiment.

4.4 Impact of Social Relation (RQ3)

Since social recommendation models utilize social information to improve recommendation, it is meaningful to study how our model outperforms in making use of social relations. As users with more relations are sparser, we divide users to four groups according to the number of sharing records: [0, 1, 2-3, >3]. This makes sure that the number of users in each group is very close. The recommendation performances for each group of our CSR and existing methods of social recommendation are shown in Figure 4.

Firstly, we can observe that all methods achieve better recommendation performance for users with more social relations, which demonstrates the importance of social information. Secondly, although all models achieve good performance for users with abundant social relations, our CSR model performs best for users with
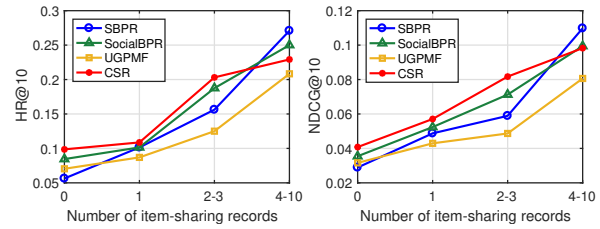
Figure 4: Performance on users with different number of social relation

sparse social relations. Even for those users without social relations, our CSR model can improve recommendation via better learned latent user/item vectors. Therefore, we can conclude that our CSR model can efficiently address the sparsity problem of social data.

5 CONCLUSION

In this work, we have proposed a novel model *Characterized Social Regularization* (CSR) to generalize the traditional social regularization methods. We have shown that existing mainstream social recommendation models utilizing regularization techniques can be regarded as special cases of our CSR model. To the best of our knowledge, this is by far the first approach to address and tackle the problem of modeling the variable characteristics of social relations in the field of social recommendation. We conducted sufficient experiments on a real-world E-commerce dataset, showing that our proposed CSR outperforms the existing methods. As future work, we would like to apply CSR to perform recommendation task on explicit interaction data, and we are also interested in exploring more possible approaches to build social relations besides item sharing.

REFERENCES

[1] Liang Du, Xuan Li, and Yi-Dong Shen. 2011. User graph regularized pairwise matrix factorization for item recommendation. In *ADMA*. 372–385.
[2] Lei Guo, Jun Ma, Hao-Ran Jiang, Zhu-Min Chen, and Chang-Ming Xing. 2015. Social trust aware item recommendation for implicit feedback. *JCST* 30, 5 (2015), 1039–1053.
[3] Mohsen Jamali and Martin Ester. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In *RecSys*. 135–142.
[4] Hao Ma, Irwin King, and Michael R Lyu. 2009. Learning to recommend with social trust ensemble. In *SIGIR*. 203–210.
[5] Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *CIKM*. 931–940.
[6] Hao Ma, Dengyong Zhou, Chao Liu, Michael R Lyu, and Irwin King. 2011. Recommender systems with social regularization. In *WSDM*. 287–296.
[7] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*. 452–461.
[8] Jiliang Tang, Huiji Gao, Xia Hu, and Huan Liu. 2013. Exploiting homophily effect for trust prediction. In *WSDM*. 53–62.
[9] Jiliang Tang, Huiji Gao, and Huan Liu. 2012. mTrust: discerning multi-faceted trust in a connected world. In *WSDM*. 93–102.
[10] Jiliang Tang, Xia Hu, and Huan Liu. 2013. Social recommendation: a review. *SNAM* 3, 4 (2013), 1113–1133.
[11] Tong Zhao, Julian McAuley, and Irwin King. 2014. Leveraging social connections to improve personalized ranking for collaborative filtering. In *CIKM*. 261–270.