Multi-Category Item Recommendation Using Neighborhood Associations in Trust Networks

Feng Xia, Haifeng Liu, Nana Yaw Asabere, Wei Wang, Zhuo Yang
School of Software, Dalian University of Technology
Dalian 116620, China
f.xia@acm.org

ABSTRACT

This paper proposes a novel recommendation method called RecDI. In the multi-category item recommendation domain, RecDI is designed to combine user ratings with information involving user's direct and indirect neighborhood associations. Through relevant benchmarking experiments on two real-world datasets, we show that RecDI achieves better performance than other traditional recommendation methods, which demonstrates the effectiveness of RecDI.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval - *Information filtering*

Keywords

Recommendation, Neighborhood Relations, Trust Networks

1. INTRODUCTION

The trust between users in social networks can be used to improve recommendation performance. The common concept of trust-based recommender systems (RSs) is that users' interests can be influenced by their trusted friends in social networks. However, users behave differently across different domains. This means that users often express different trust relations in different domains. For example, a user u may trust user v in terms of books but the same user u may not trust user v in terms of movies. Traditional trust-based recommendation methods use trust information on all categories, which is not consistent due to the fact that trust isn't applicable in different domains. It is therefore important to develop appropriate methods that utilize trust relations of users for recommendation in different domains.

In this work, we propose a novel recommendation method called RecDI, which combines users' ratings with the information about direct and indirect neighborhood associations in trust networks for multi-category item recommendation. According to items' categories, we first partition users and items into different subsets and obtain rating and trust relation sets on each category. We then use direct neighborhood association degrees to compute indirect neighborhood association degrees by exploiting trust propagation. Finally, we fuse these association degrees into a traditional matrix factorization model for item recommendation on each category.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577326.

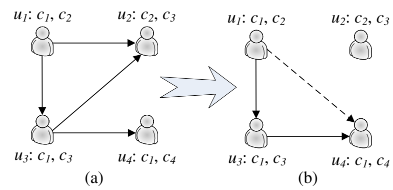

Figure 1: (a) Original trust network on all categories; (b) Generated trust network on category c_1

2. PROPOSED SCHEME

2.1 Partition Rating and Trust Relation Sets

As mentioned above, users often behave differently across multiple domains of interest. A simple way to obtain these domains is to partition items into different sets according to their categories. Users are also grouped into a relevant category according to whether they have rated items pertaining to that category. Direct neighborhood associations related to each category are easily extracted from original trust networks by removing users who do not belong to the category. Indirect neighborhood associations are obtained through the process of trust propagation. Consequently, a trust network on each category is generated, as shown in Figure 1.

2.2 Build Unified Objective Function

Matrix Factorization (MF) is a model-based collaborative filtering method. In recommender systems there are a set of users $U = \{u_1, ..., u_N\}$ and a set of items $I = \{i_1, ...i_M\}$. The ratings expressed by users on items are given in a rating matrix $R = [R_{u,i}]_{N\times M}$.In this matrix $R_{u,i}$ denotes the rating of user u on item i. $R_{u,i}$ can be any real number, but the ratings are integers in the range (1-5). The objective function of MF method on each category is denoted as:

$$L_1^c(P^c, Q^c) = \sum_{u\in U, i\in I} W_{u,i}^c(R_{u,i}^c - P_u^c Q_i^c) + \lambda(\|P^c\|_F^2 + \|Q^c\|_F^2) \tag{1}$$

Table 1: Epinions: Top-5 Category Statistics

Category	User count	Item count	Rating count	Trust count	Rating sparsity	Trust sparsity
Movies	14178	28601	166195	153916	0.9996	0.9992
Books	10729	59109	102600	104728	0.9998	0.9991
Music	9008	34524	85113	61508	0.9997	0.9992
Travel	10620	15412	56942	99607	0.9997	0.9991
Family	8601	24120	84769	68492	0.9996	0.9991

Table 2: Ciao: Top-5 Category Statistics

Category	User count	Item count	Rating count	Trust count	Rating sparsity	Trust sparsity
DVDs	4357	11293	39523	59715	0.9992	0.9969
Books	3296	12602	21926	51397	0.9995	0.9953
Beauty	3281	9352	23852	50875	0.9992	0.9953
Travel	3675	12194	21773	57023	0.9995	0.9958
CiaoCafe	4378	2893	30090	78875	0.9976	0.9959

where $W^c_{u,i}$ is an indicator function that is equal to 1 if user u expressed rating on item i of category c and equal to 0 otherwise, $\|.\|^2_F$ denotes the Frobenius norm, λ is the regularization coefficient. In matrix factorization model, P_u denotes latent feature vector of user u. The theory of trust-based recommendation stipulates that the tastes of users existent in unidirectional and bidirectional trust associations have to be similar. Thus, we integrate the trust association information by minimizing the following objective function:

$$L^c_2(P^c) = \sum_{u\in U} \left\| (P^c_u - \sum_{v\in U} TD^c_{u,v} P^c_v - \sum_{v'\in U} TI^c_{u,v'} P^c_{v'}) \right\|^2 \quad (2)$$

where $TD^c_{u,v}$ denotes the trust association degree between user u and its direct neighbor v and equals the ratio of the number of items user u rated on category c to the total number of items of category c. $TI^c_{u,v'}$ denotes the association degree between user u and its indirect neighbor v' and can be computed by multiplying the association degrees among direct neighbors on the link from u to v' based on the process of trust propagation. The unified object function for RecDI is defined as:

$$L^c(P^c, Q^c) = L^c_1(P^c, Q^c) + \alpha L^c_2(P^c) \quad (3)$$

where α is a non-negative parameter that trades off the two objective functions. The minimum of the objective function can be found by performing gradient descent in P^c_u, Q^c_i.

3. EXPERIMENTS

We used the Epinions and Ciao datasets [1] for experiments. The distributions of users and items in top-5 categories of the two datasets are presented in Tables 1 and 2. We compared RecDI with the following methods to evaluate the effectiveness. **RecT**: is a traditional recommendation method (PMF [2]) which doesn't involve the use of trust information in the application of each category's ratings. **RecD**: is a direct trust relation-based recommendation method which involves the removal of the term $\sum_{v'\in U} TI^c_{u,v'} P^c_{v'}$ that is associated with indirect trust relation from Equation (2). In RecDI, λ is set to 0.01 and α is set to 20. The

Table 3: Epinions: Performance Comparisons

Category	RecT	RecD	RecDI	
Movies	1.225	1.2228	**1.2191**	RMSE
	0.9577	0.9576	**0.957**	MAE
Books	1.1019	1.0976	**1.088**	RMSE
	0.8862	0.8815	**0.8766**	MAE
Music	1.0987	1.0946	**1.0872**	RMSE
	0.8863	0.8829	**0.8791**	MAE
Travel	1.2222	1.2192	**1.2144**	RMSE
	0.9276	0.9264	**0.925**	MAE
Family	1.2115	1.2099	**1.2037**	RMSE
	0.9225	0.9207	**0.9182**	MAE
Average	1.1769	1.1739	**1.168**	RMSE
	0.9218	0.9198	**0.9175**	MAE

Table 4: Ciao: Performance Comparisons

Category	RecT	RecD	RecDI	
DVDs	1.1817	1.1761	**1.1687**	RMSE
	0.9142	0.9117	**0.9092**	MAE
Books	1.3431	1.3429	**1.3405**	RMSE
	0.9836	0.9834	**0.9825**	MAE
Beauty	1.1807	1.1685	**1.1571**	RMSE
	0.904	0.8972	**0.8918**	MAE
Travel	1.0251	1.0053	**0.9907**	RMSE
	0.8575	0.8451	**0.8369**	MAE
CiaoCafe	1.1531	1.1394	**1.1254**	RMSE
	0.8976	0.8888	**0.881**	MAE
Average	1.1915	1.1829	**1.1742**	RMSE
	0.9169	0.9117	**0.9074**	MAE

Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are employed as the performance measures to evaluate the prediction quality.

Tables 3 and 4 summarize the performance comparisons of the above recommendation methods on Epinions and Ciao. From the experimental results, it can be observed that RecD outperforms RecT and RecDI outperforms RecD. The first observation reveals that direct neighborhood associations can generate positive effects on prediction accuracy while the second reveals that indirect neighborhood associations can also generate positive effects in terms of prediction accuracy. This is due to the fact that trust relations express the similarities between users' interests and indirect trust relations complement the use of trust information which helps to accurately determine the similarities of users.

4. CONCLUDING REMARKS

The theory behind our method RecDI is that the tastes of users who make direct expressions on trust are similar and the tastes of users who make indirect expressions on trust are also similar. In future work, we will test this theory and the proposed method extensively.

5. REFERENCES

[1] J. Tang, H. Gao, H. Liu, and A. D. Sarma. eTrust: Understanding trust evolution in an online world. In *KDD*, pages 253–261, 2012.

[2] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *NIPS*, pages 1257–1264, 2008.