TCRec: Product Recommendation via Exploiting Social-trust Network and Product Category Information

Yu Jiang, Jing Liu, Xi Zhang, Zechao Li, Hanqing Lu
National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences, Beijing, China
{yjiang, jliu, xi.zhang, zcli, luhq}@nlpr.ia.ac.cn

ABSTRACT

In this paper, we develop a novel product recommendation method called TCRec, which takes advantage of consumer rating history record, social-trust network and product category information simultaneously. Compared experiments are conducted on two real-world datasets and outstanding performance is achieved, which demonstrates the effectiveness of TCRec.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information filtering; H.3.5 [**Online Information Services**]: Commercial services

General Terms

Algorithms, Experimentation

Keywords

Recommendation, Social-trust Network, Product Category

1. INTRODUCTION

Nowadays, e-commerce is becoming dramatically popular, since it brings a kind of convenient, fast, and cheap shopping experience to users. Product recommender systems, suggesting to customers the products that are potentially of their interests, are playing a critical role in e-commerce, especially in the condition that massive amounts of products are on the e-commerce websites and the problem of information overload becomes so severe.

Most of the traditional recommender systems are based on Collaborative Filtering (CF), which focus on utilizing customer rating history data. Thanks to the booming online social networking websites, recently, there are a surge of research interests on mining the social-trust network to improve the prediction quality [1, 2, 4]. Additionally, products on the e-commerce websites are well categorized, such as "Books", "Electronic & Computer", etc. Such kind of structured information is also beneficial to product recommendation, since interests of customers on products are varying according to different category . As a consequence, this paper presents a novel product recommendation method, named TCRec for short, which expands the typical Probabilistic Matrix Factorization (PMF) [3] model by jointly leveraging social-trust network and product category information as well as rating history record. On one hand, TCRec integrates the social-trust network by a graph Laplacian regularization, the rationale behind which is that customers are likely to have similar tastes with their linked friends in social-trust network. On the other hand, the product category information is utilized through a regularized regression model, which assumes that the latent features of products in one category can be discriminated from others. Experiments carried out on two real-world datasets verify the effectiveness of TCRec.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

2. TCREC

To achieve the ultimate target, TCRec involves three components to exploit the consumer rating history record, the social-trust network and the product category information, respectively.

2.1 Rating History Record

Suppose we have a customer-product rating matrix $\mathbf{X} \in \mathbb{R}^{N\times M}$. Let $\mathbf{U} \in \mathbb{R}^{K\times N}$ and $\mathbf{V} \in \mathbb{R}^{K\times M}$ be the latent customer and product feature matrices, with column vectors $\mathbf{U}_n$ and $\mathbf{V}_m$ representing customer-specific and product-specific latent feature vectors, respectively. Normally, there are a large number of products, and a customer may only rates a small portion of the whole item set practically. To take into account the large number of missing elements in $\mathbf{X}$, we solve the following constrained minimization problem:

$$\begin{aligned}\mathcal{L}_1(\mathbf{U},\mathbf{V}) = &\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbf{I}_{nm}(\mathbf{X}_{nm} - \mathbf{U}_n^T\mathbf{V}_m)^2 \\ &+ \lambda(||\mathbf{U}||_F^2 + ||\mathbf{V}||_F^2)\end{aligned} \tag{1}$$

where $\mathbf{I}_{nm}$ is indicator function that is equal to 1 if customer n rated product m and equal to 0 otherwise, $||.||_F^2$ denotes the Frobenius norm, λ is the regularization coefficient.

2.2 Social-trust Network

Consider we have a undirected social-trust network graph, where a vertex represents a customer, a weighted edge represents the relation strength between customers. $\mathbf{W} \in \mathbb{R}^{N\times N}$ is used to represent the weight of edge. $\mathbf{W}_{ij}$ is set to 1 if the trust relation is bilateral, set to 0.5 if the trust relation is unilateral, and set to 0 if there is no trust relation. We assume that customer's interest to products is influenced by the other linked customers in social-trust network. Thus, we integrate the trust relation information by minimizing the

Table 1: Statistics of the Datasets

	Epinions	Ciao
# of Customers	7475	4137
# of Products	11783	7828
# of Rates	170056	101331
# of Trust Links	225462	130238
# of Categories	8	10
Rating Sparsity	99.81%	99.66%
Trust Network Sparsity	99.60%	99.24%

following objective function:

$$\mathcal{L}_2(\mathbf{U}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}||\mathbf{U}_i - \mathbf{U}_j||^2\mathbf{W}_{ij} = Tr(\mathbf{U}^T\mathbf{L}\mathbf{U}) \quad (2)$$

where $\mathbf{L}$ is the graph Laplacian of $\mathbf{W}$.

2.3 Product Category Information

The products in different categories should be discriminated by exploring the product-specific latent representations $\mathbf{V}$. To this end, we formulate the discriminative learning problem into a regularized regression model as follows.

$$\mathcal{L}_3(\mathbf{B}, \mathbf{V}) = ||\mathbf{B}^T\mathbf{V} - \mathbf{Y}||_F^2 + \gamma||\mathbf{B}||_{2,1} \quad (3)$$

where $\mathbf{Y} \in \mathbb{R}^{C\times M}$ is the product-category matrix. $\mathbf{Y}_{cm}$ is equal to 1 if the m-th product belongs to the c-th category, and equal to 0 otherwise. $\mathbf{B} \in \mathbb{R}^{K\times C}$ is the regression coefficient matrix and

$$||\mathbf{B}||_{2,1} = \sum_{k=1}^{K}\sqrt{\sum_{c=1}^{C}\mathbf{B}_{kc}^2} \quad (4)$$

The $l_{2,1}$-norm regularization term is introduced to ensure $\mathbf{B}$ sparse in rows. In that way, $\mathbf{B}$ does a feature selection during the fitting process that leads to a better reconstruction of the product category matrix $\mathbf{Y}$.

2.4 Unified Objective Function

The unified object function for TCRec is defined as

$$\mathcal{L}(\mathbf{U}, \mathbf{B}, \mathbf{V}) = \mathcal{L}_1(\mathbf{U}, \mathbf{V}) + \alpha\mathcal{L}_2(\mathbf{U}) + \beta\mathcal{L}_3(\mathbf{B}, \mathbf{V}) \quad (5)$$

where α and β are nonnegative parameters to trade off these objectives. Requiring that the derivative of $\mathcal{L}$ with respect to $\mathbf{B}$ vanish, we have the update formula of $\mathbf{B}$ and substitute it into Eq.(5). Then, a local minimum of the objective function can be found by performing gradient descent in $\mathbf{U}_n$, $\mathbf{V}_m$.

3. EXPERIMENTS

In this section, we evaluate the effectiveness of TCRec. Epinions and Ciao [4] are employed as the data source for experiments. Both of them are well-known consumer opinion websites where customers can not only give reviews to their familiar products but also maintain a trust list which presents a social-trust network. What's more, products on these websites are categorized well. We remove products rated less than 5 as well as customers without trust relations or rating less than 10. The detailed statistics of the two datasets are showed in Table 1. To validate our approach, we compare TCRec with the following methods: **PMF** [3]: only the rating information is utilized. **SoRec** [2]: a PMF-based recommendation algorithm with social-trust network. **CRec**: our proposed method with α set to 0, that is only product category information utilized. **TRec**: our proposed method with β set to 0, that is only social-trust network utilized. In our approach, λ is set to 5, α and β are both set to 1, γ is set to 10 and the rank K is set to 10. Both the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are employed as the performance measures to evaluate the prediction quality. 80% of the review data are picked randomly to form the training set and the rest for the test set. 5 random train-test splits are applied, and 10 test runs are conducted for each split. The mean of the performance is reported.

Table 2: Performance Comparisons

	Epinions		Ciao	
	RMSE	MAE	RMSE	MAE
PMF	1.0793	0.9223	1.0454	0.9084
SoRec	1.0596	0.9155	1.0221	0.8955
CRec	1.0677	0.9162	1.0262	0.8934
TRec	1.0480	0.9115	1.0053	0.8878
TCRec	**1.0241**	**0.9006**	**0.9813**	**0.8706**

Table 2 summarizes the performance comparisons. From the experimental results, it can be clearly observed that our approach outperforms the other methods. We also observe that all of SoRec, CRec and TRec improve the performance to some extent compared with PMF. All these observation reveals that both social-trust network and product category information are helpful for product recommendation. The performance of TRec is better than that of SoRec. The reason may be that the effect of graph Laplacian regularization is more directly.

4. CONCLUSION

We present an extended PMF model named TCRec, to improve product recommendation accuracy by jointly exploring both social-trust network and product category information. The central idea of TCRec is that the tastes of customers who trust each other are similar and the latent features of products in one category can be discriminated from others. The experimental results on two real-world datasets indicate the effectiveness of our approach.

5. ACKNOWLEDGMENTS

This work was supported by 973 Program (2012CB316304) and National Natural Science Foundation of China (61272329, 61170127 and 61070104).

6. REFERENCES

[1] B.-K. Bao, W. Min, J. Sang, and C. Xu. Multimedia news digger on emerging topics from social streams. In *MM*, pages 1357–1358, 2012.

[2] H. Ma, H. Yang, M. R. Lyu, and I. King. SoRec: social recommendation using probabilistic matrix factorization. In *CIKM*, pages 931–940, 2008.

[3] R. Salakhutdinov and A. Mnih. Probabilistic Matrix Factorization. In *NIPS*, pages 1257–1264, 2008.

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