

Personalized Recommendation of Classification based on Social Relationship and Time Information

Mengzi Tang^{1, a}, Li Li^{1, b}

¹Department of Computer and Information Science, Southwest University, Chongqing, 400715, China

^aemail: tmzsummer@163.com, ^bemail: lily@swu.edu.cn

Keywords: Recommender performance, Social relationship, Time information, Classification

Abstract. Improving recommender performance is beneficial. However, sparsity and scalability of data is a problem faced by recent recommender systems. Customer interests and available products are changing constantly. Social interactions among users are highly influential on the effectiveness of the recommendation. Following these intuitions, this paper proposed a recommendation method incorporating with social and temporal information with probabilistic matrix factorization, which is called PMFST (Probabilistic Matrix Factorization with Social and Temporal information), to solve the problem of data sparsity and achieve real and dynamic recommender performance. The experiment on two real data sets shows that the proposed method outperforms the state-of-the-art methods in terms of minimal error and recommender performance.

Introduction

In recommender systems, sparsity and scalability are two common difficulties. Much work addresses the challenge of cold start problem by using social relationship. People usually turn to their friendship/neighbors for help to obtain our needed information in real life. Massa et al. [1][2] proposed a CF frame based on trust and a trust degree algorithm—MoleTrust. In fact, people's choices are influenced by not only social relationship but also time factor. The popularity of products on Internet may change with time, as well as people's interests. Time information is crucial, especially when comments and ratings are related to products within a specific period to model recommendation. Koren et al. [3] study Netflix's dynamic properties and succeed in enhancing predicting ratings with time information.

However, few attempts have taken full advantages of social relationship and time information in recommender systems. In this paper, a probabilistic graph model is proposed to combine social relationship and time to achieve precise and comprehensive recommendation. The probabilistic graph model shows superior performance of inference. It achieves better performance than the state-of-the-art models. In recommender systems, large data sets are hard to handle due to limited resources. To illustrate the problem, Fig.1 describes our approach by exploiting part data of Epinions. We recommend items for users from rating (or other feedback) data concerning items that can be classified into different categories. It is our basic idea to transform recommending items into recommending categories firstly. Then we consider recommending item for every user according to the category that belongs to, which can effectively avoid computing large data sets.

In this paper, we try to conduct a systematic research of improving recommendation performance with following contributions. We depend on tensor factorization model incorporating social and time information (PMFST) to realize real-time and dynamic recommendation, which overcomes the cold start and sparsity problem. We take some methods to recommend items according to the corresponding category, which realize particular and personal recommendation. We conduct our experiments on two different data sets: Epinions and Ciao. Experimental results indicate different data sets applied with the PMFST model, which can achieve good performance.

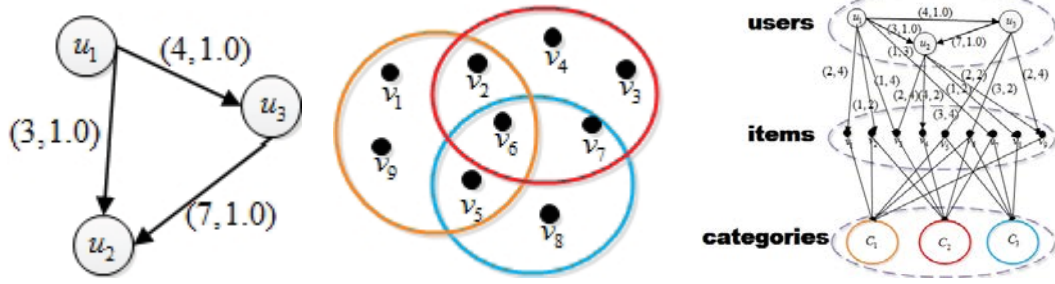


Fig. 1: An example of a classification network: (a) A social network graph generated by three users. For example, (4, 1.0) at the edge of u_1 and u_3 means u_1 trusts u_3 at t_4 and trust value is 1.0 in the social network. (b) Each item belonging to some categories, so it has overlap; (c) The graph of our model by adding the categories level between users and items.

Problem Definition

Problem Formulation

In this section, we present a formal definition of the problem. At first, we give several definitions concerning the problem.

A **rating network** can be represented as $G_r = (U, V, E_r)$, where the vertex set $|U| = n$ users in a rating network, the vertex set $V = \bigcup_{j=1, c=1}^{m, cm} V_j^c$ represents m items and cm categories (we divide the items into cm category), and the edge set E_r represents the links between users and items. $|V_j^c|$ describes item j belonging to the category c . $R = \{r_{ij}^t\}$ is called rating tensor, which is a $m \times n \times t$ matrix, $t \in \{1, \dots, T\}$ and $r_{ij}^t \in [1, 5]$. In order to match the range of predictions our model, we map the ratings $1 \dots 5$ to the interval $[0, 1]$ by normalizing the ratings. r_{ij}^t denotes the ratings of user i rates item j at time t .

A **directed social network** can be represented as $G_s = (U, E_s)$, where the vertex set $U = \{u_i\}_{i=1}^n$ denotes all users in a social network and the edge set E_s represents the trust relation between users. $S = \{s_{ik}^t\}$ is a $n \times n \times t$ matrix of G_s , which is called the social network tensor, where $t \in \{1, \dots, T\}$ and weighted social relationship of u_i with u_k at time t (e.g. u_i trusts/knows u_k) in a social network is depicted by a positive value $s_{ik}^t \in (0, 1]$. In fact, $s_{ik}^t = 0$ reflects an unknown or unobserved social relationship.

In order to consider the social influence, we further define the influence of friend u_k on u_i as following:

$$s_{ik}^{t*} = \alpha s_{ik}^t + (1 - \alpha) \frac{In(u_k)}{In(u_k) + Out(u_i)} \quad (1)$$

In which, α and $1 - \alpha$ represent the u_i and u_k 's weight parameters from real and social influence respectively. $In(u_k)$ means the in-degree of u_k and $Out(u_i)$ is the out-degree of u_i .

The category rating tensor: r_{ij}^t is the rating of item j belongs to c ($c \in \{1, \dots, cm\}$), so we consider the ratings of all items belong to category c and define the rating of category r_{ic}^t as follows:

$$r_{ic}^t = \frac{\sum_{j \in c} r_{ij}^t}{\sum_{j \in c} 1} \quad (2)$$

Specifically, $R_c = \{r_{ic}^t \in [0, 5]\} (t \in \{1, \dots, T\})$ is called category rating tensor.

Probabilistic Matrix Factorization with Social and Time Information (PMFST)

The idea of our approach is to draw real rating with probabilistic matrix factorization. And it incorporates with social and time information. Therefore, the problem transforms 2-dimension matrix factorization into 3-dimension matrix factorization called tensor factorization. Assuming $U \in R^{D \times n}$, $V \in R^{D \times m}$, $T \in R^{D \times l}$, $Z \in R^{D \times n}$ respectively denote the latent user feature, item features (category features), time features in rating tensor and factor feature in social tensor, where R represents ratings of users on items and calculating ratings of users on categories as well. In fact, $C = \{s_{ik}^{t*}\}$ reflects the set of social relationship. This paper consider time factor T as one of factors in matrix factorization. Therefore, U_i, V_j, T_l, Z_k respectively indicates user-specific, item-specific (category-specific), time-specific and factor-specific. Fig.2 shows the graphical representation of our approach (PMFST).

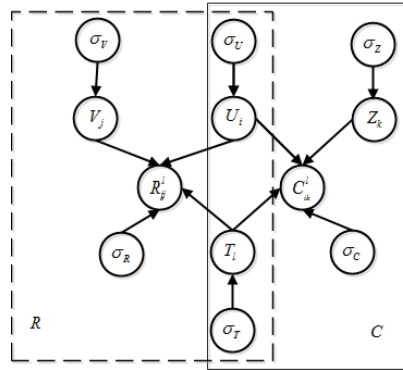


Fig. 2: The graphical representation of PMFST model

We define user latent feature vector U_i , item (category) latent feature vector V_j , time information latent feature vector T_l and social factor latent feature vector Z_k following zero-mean spherical Gaussian distribution. They can be defined as

$$p(U | \sigma_U^2) = \prod_{i=1}^n [N(U_i | 0, \sigma_U^2 I)]^{I_{ij}^k} \quad p(V | \sigma_V^2) = \prod_{j=1}^m N(V_j | 0, \sigma_V^2 I) \quad (3)$$

$$p(T | \sigma_T^2) = \prod_{l=1}^T N(T_l | T_{l-1}, \sigma_T^2 I) \quad p(T_0 | \sigma_0^2) = \prod_{l=1}^T N(T_l | \mu_T, \sigma_0^2 I) \quad p(Z | \sigma_Z^2) = \prod_{ci=1}^n N(Z_{ci} | 0, \sigma_Z^2 I)$$

where $N(x|\mu, \sigma^2)$ is the probability density function of the Gaussian distribution with mean μ and variance σ^2 . We define the conditional distribution over the observed ratings as:

$$p(R|U, V, T, T_0, \sigma^2) = \prod_{i=1}^N \prod_{j=1}^M \prod_{l=1}^T [N(r_{ij}^l | \langle U_i, V_j, T_l \rangle, \sigma_R^2)]^{I_{ij}^{l,r}} \quad (4)$$

where $I_{ij}^{l,r}$ is the indicator function that is equal to 1 if user i rates item or category j at time l and equal to 0 otherwise. The posterior distribution over the user, items' category and time features is given as

$$\begin{aligned} p(U, V, T, T_0 | R, \sigma_R^2, \sigma_U^2, \sigma_V^2, \sigma_T^2, \sigma_0^2) &\propto p(R|U, V, T, T_0) p(U | \sigma_U^2) p(V | \sigma_V^2) p(T | \sigma_T^2) p(T_0 | \sigma_0^2) \\ &= \prod_{i=1}^N \prod_{j=1}^M \prod_{l=1}^T [N(r_{ij}^l | \langle U_i, V_j, T_l \rangle, \sigma_R^2)]^{I_{ij}^{l,r}} * \prod_{i=1}^n [N(U_i | 0, \sigma_U^2 I)] * \prod_{j=1}^m N(V_j | 0, \sigma_V^2 I) \\ &* \prod_{l=1}^T N(T_l | 0, \sigma_T^2 I) * N(T_0 | 0, \sigma_0^2 I) \end{aligned} \quad (5)$$

Similarly, we define the conditional distribution over social relationship as

$$p(C|U, Z, T, T_0, \sigma^2) = \prod_{i=1}^N \prod_{k=1}^N \prod_{l=1}^T [N(c_{ik}^l | \langle U_i, Z_k, T_l \rangle, \sigma_C^2)]^{I_{ik}^{l,c}} \quad (6)$$

Combining the social and rating tensor, the posterior distribution over all features is written as

$$p(U, V, Z, T, T_0 | C, R, \sigma_C^2, \sigma_R^2, \sigma_U^2, \sigma_Z^2, \sigma_T^2, \sigma_0^2) \propto p(U, V, T, T_0 | R, \sigma_R^2, \sigma_U^2, \sigma_V^2, \sigma_T^2, \sigma_0^2) * p(U, Z, T, T_0 | C, \sigma_C^2, \sigma_U^2, \sigma_Z^2, \sigma_T^2, \sigma_0^2) \quad (7)$$

The log of the posterior distribution is equivalent to minimizing the sum-of-squared-errors objective function with quadratic regularization terms, and simplifying as

$$E_r = \sum_{l=1}^T \sum_{i=1}^n \sum_{j=1}^m I_{ij}^{l,r} (R_{ij}^l - \langle U_i, V_j, T_l \rangle^2) + \frac{\lambda_C}{2} \sum_{l=1}^T \sum_{i=1}^n \sum_{k=1}^n I_{ik}^{l,c} (C_{ik}^l - \langle U_i, Z_k, T_l \rangle^2) + \sum_{i=1}^n \frac{\lambda_U \|U_i\|^2}{2} + \sum_{j=1}^m \frac{\lambda_V \|V_j\|^2}{2} + \sum_{k=1}^n \frac{\lambda_Z \|Z_k\|^2}{2} + \sum_{l=1}^T \frac{\lambda_T \|T_l - T_{l-1}\|^2}{2} + \frac{\lambda_0 \|T_0 - \mu_T\|^2}{2} \quad (8)$$

where $\lambda_C = \sigma_R^2 / \sigma_C^2$, $\lambda_U = \sigma_R^2 / \sigma_U^2$, $\lambda_V = \sigma_R^2 / \sigma_V^2$, $\lambda_Z = \sigma_C^2 / \sigma_Z^2$ and $\lambda_T = \sigma_T^2 * \sigma_T^2 / (\sigma_Z^2 * \sigma_R^2)$.

Then we utilize Gradient descent to update the latent feature vectors and acquire the min-error and final latent features to achieve predicting ratings.

$$\begin{aligned} \frac{\partial E}{\partial U_i} &= \sum_{l=1}^T \sum_{j=1}^m I_{ij}^{l,r} (R_{ij}^l - \langle U_i, V_j, T_l \rangle) V_j T_l + \lambda_C \sum_{l=1}^T \sum_{k=1}^n I_{ik}^{l,c} (C_{ik}^l - \langle U_i, Z_k, T_l \rangle) Z_k^T T_l + \lambda_U U_i \\ \frac{\partial E}{\partial V_j} &= \sum_{l=1}^T \sum_{i=1}^n I_{ij}^{l,r} (R_{ij}^l - \langle U_i, V_j, T_l \rangle) U_i^T T_l + \lambda_V V_j \\ \frac{\partial E}{\partial T_l} &= \sum_{i=1}^n \sum_{j=1}^m I_{ij}^{l,r} (R_{ij}^l - \langle U_i, V_j, T_l \rangle) U_i^T V_j + \lambda_C \sum_{i=1}^n \sum_{k=1}^n I_{ik}^{l,c} (C_{ik}^l - \langle U_i, Z_k, T_l \rangle) U_i^T Z_k + \lambda_T T_l \\ \frac{\partial E}{\partial Z_k} &= \lambda_C \sum_{l=1}^T \sum_{i=1}^n I_{ik}^{l,c} (C_{ik}^l - \langle U_i, Z_k, T_l \rangle) U_i^T T_l + \lambda_Z Z_k \end{aligned} \quad (9)$$

Recommendation

Firstly, we make recommendation for users with our proposed method — PMFST. In order to recommend personally and improve the performance of recommendation, we directly select the items among same category and apply PMFST model during this phrase. It can effectively calculate the ratings of items among the same category instead of items in whole data set. It not only reduces the calculating but also solves time and space expansion.

Experiment

In this section, in order to evaluate the effectiveness and efficiency of our approach, we conduct two different data sets on experiment.

Experimental Settings

Data Sets. In our paper, experiments are performed on two data sets: Epinions and Ciao both consist of rating and trust information. The data sets contain information about users, items, categories, time and social relationships. We divide the collected data into different time slices. Table 1 is the statistical summary of the data sets.

Baseline Algorithms. We compare our approach with the following five methods.

Table 1: The statistical information of two data sets

	Users	items	Categories	Time slices	Relationships
Ciao	2378	78352	28	11	57544
Epinions	8519	153596	27	11	300548

PMF [4]: It is a basic matrix model regarding the user preference matrix as a product of two lower-rank user and movie matrices. It is only based on rating matrix.

SoRec[5]: Ma et al. presented a social recommendation framework by employing both users' social network information and rating records.

SVD [6]: As the Netflix Prize competition has demonstrated, matrix factorization models are superior to classic nearest-neighbor techniques for products recommendation.

SVD++ [7]: It improves the basic matrix factorization model which incorporating implicit

feedback information. Making use of implicit feedback can help indicate users' preference, but time information and real social relationship may be very important for recommendation.

TimeSVD++ [3]: Koren [3] proposed TimeSVD++ method to improve the performance, which assumes that time information can be a latent feature. But it lacks some real social relationship (such as trust information).

Performance Metrics. Our model includes two evaluation methods, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

We use MAE and RMSE to measure the prediction of our proposed approach (PMFST) in comparison with other methods. MAE and RMSE are given as

$$MAE = \frac{\sum_{i,j} |r_{ij} - r_{ij}^*|}{T_N} \quad (11)$$

$$RMSE = \sqrt{\frac{1}{T_N} \sum_{i,j} (r_{ij} - r_{ij}^*)^2} \quad (12)$$

where r_{ij} denotes the rating which user i gave to item j , r_{ij}^* denotes the rating which user i gave to item j as predicted by our approach, and denotes the number of the testing ratings.

Result Analysis

The model based on probabilistic matrix factorization has several parameters. The parameter settings of this paper are $\lambda_u = \lambda_v = \lambda_z = \lambda_T = 0.001$, $\lambda_c = 0.5$ referenced to the empirically settings.

Table 2: The MAE and RMSE performance of different approaches on the Ciao-class, Ciao, Epinions-class, and Epinions data sets.

Data Set	Method	PMF	SoRec	SVD	SVD++	TimeSVD++	PMFST
Ciao_class	MAE	0.703	0.638	0.603	0.589	0.457	0.362
	RMSE	0.981	0.804	0.779	0.671	0.561	0.494
Ciao	MAE	0.863	0.685	0.663	0.629	0.606	0.451
	RMSE	0.981	0.804	0.917	0.671	0.631	0.579
Epinions_class	MAE	0.763	0.743	0.795	0.704	0.436	0.366
	RMSE	0.875	0.855	0.966	0.852	0.682	0.617
Epinions	MAE	1.199	0.802	0.927	0.763	0.796	0.489
	RMSE	1.433	1.136	1.177	0.998	0.828	0.711

Performance comparison. To demonstrate the better generalization ability of our model, we make comparison with methods shown above on Ciao data set of categories clustered called Ciao-class, Ciao, Epinions data set of categories clustered called Epinions-class and Epinions.

As the data scale of Ciao is smaller than Epinions, and Ciao of clustering by category is smaller as well. Therefore, it is suitable to testify our model perform well in comparatively small data sets. Table 2 lists the MAE and RMSE of the different approaches on the Ciao-class, Ciao, Epinions and Epinions-class with the following observations.

Our method PMFST consistently performs better performance comparing to the baseline methods. No matter what Ciao-class or Epinions-class obtained with aggregation, the model PMFST outperforms almost 10% in MAE and RMSE measures comparing to PMF, SoRec, SVD, SVD++, TimeSVD++. TimeSVD++ performs better than SVD++ and SVD, SVD++ also performs better than Sorec and PMF. Therefore, the experiment concerning different approaches on the

Ciao-class, Ciao, Epinions-class and Epinions data sets demonstrates trust relationship or time information, combining these two have a great influence on predicting ratings.

Conclusion

In this paper, we propose a novel method for personalized recommendation. To leverage tensor factorization and rich social relationships along with time information, we extend the basic matrix factorization and put forward a novel probabilistic model incorporating social relationship and time information. And we make use of the idea of classification to recommend categories and items instead of items directly. Finally, the model achieves better performance comparing to other models.

Acknowledgement

The work was supported by (XDJK2016E069), (KJ1501405), NSFC (No.61170192) and National High-tech R&D Program of China (No.2013AA013801).

References

- [1] Massa P, Avesani P. Trust-Aware Collaborative Filtering for Recommender Systems[M]// On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Springer Berlin Heidelberg, 2004:492-508.
- [2] Avesani P, Massa P, Tiella R. A trust-enhanced recommender system application: Moleskiing[C]// ACM Symposium on Applied Computing. 2010:1589-1593.
- [3] Koren Y. Collaborative filtering with temporal dynamics[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2009:89-97.
- [4] Mnih A, Salakhutdinov R. Probabilistic matrix factorization[C]// International Conference on Machine Learning. 2012:880-887.
- [5] Ma H, Yang H, Lyu M R, et al. SoRec: Social Recommendation Using Probabilistic Matrix Factorization[C]// ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, Usa, October. 2008:931-940.
- [6] Vozalis M G, Margaritis K G. Using SVD and demographic data for the enhancement of generalized Collaborative Filtering [J]. Information Sciences, 2007, 177(15):3017-3037.
- [7] Koren B. Y. "Factorization meets the neighborhood", Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining KDD 08[J]. 2013.