International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 2 Issue: 12

ISSN: 2321-8169
4150 - 4155

Improving the Performance of Recommendation on Social Network by Investigating Interactions of Trust and Interest Similarity

Priyanka Dhananjay Chavan
PG Scholar, Dept of Computer Engineering,
KJ College Of Engineering and Management Research,
Pune, India
priya.chavan9@gmail.com

Vikas Maral
Assistant Professor, Dept of Computer Engineering,
KJ College Of Engineering and Management Research,
Pune, India
vikasmaral@gmail.com

Abstract— On the social media, lots of people share their experiences through various factors like blogs, online ratings, reviews, online polling and tweets. Study shows that the factors such as interpersonal interest and interpersonal influence from the social media which is based on the circles as well as groups of friends leads to opportunities and challenges in solving the problems on datasets. This challenge is for the Recommender System (RS) to find the solution on cold start and sparsity problems. In this paper, on the basis of the probabilistic matrix factorization, the social factors like personal interest, interpersonal influence and interpersonal interest similarity are combined into a unified personalized recommendation model. These factors can improve the associating linkage in latent space. Various datasets are used to conduct the experiments to get the results that show that the proposed model performs better than the existing approaches.

*****

I. INTRODUCTION

Recommender system (RS) has been successfully exploited to solve information overload. In Ecommerce, like Amazon, it is important to handling mass scale of information, such as recommending user preferred items and products [64]. A survey shows that at least 30 percent of the sales in Amazon come from the work of the RS. It can be viewed as the first generation of RSes [7] with traditional collaborative filtering algorithms to predict interest of user. However, with the fast increase in number of registered users and various products, the problem of cold start for users such as new users into the RS with little historical behavior and the sparsity of datasets that is the proportion of rated user-item pairs in all the user-item pairs of RS have been increasingly intractable.

Exponential growth of data in the form of information generated by online social networks demands effective recommender systems. So often the social context has not been fully considered as traditional techniques though they are qualified as they ignore social relation data. The challenging task is to fuse social contextual factors which are derived from user's motivation of social behaviors into social recommendation. We are studying and analyzing social recommendation based on psychology and sociology studies exhibiting two important factors that are, individual preference and interpersonal influence. First, we are presenting the particular importance of these two factors in online item adoption as well as recommendation. Probabilistic matrix factorization method is proposed to fuse them in latent spaces. The empirical results on large datasets demonstrate that our method is significant over existing approaches.

Appearance of web2.0 greatly improves user's initiative on the Internet brings volume of social networks such as Facebook, Orkut, Twitter, Yelp, Douban, Epinions, etc. The interpersonal relationship, especially the friend circle social networks makes it helpful to solve the cold start and sparsity problem. Social media give us some valuable clues to recommend user favorite items like music, preferred brand/products user's approach of tagging when sharing a photo to social media networks and user interested travel places by exploring social community contributed.

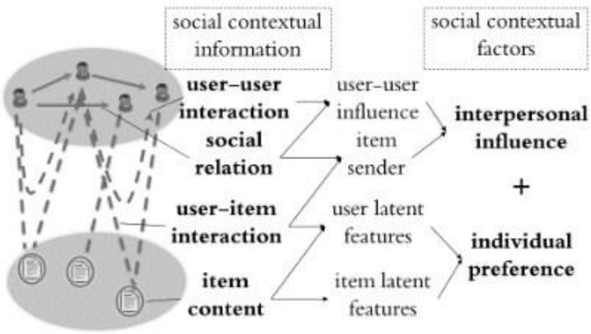

Figure 1. Framework for social recommendation:
It understands the mechanism of user behaviour on social networks that also utilizes contextual information and summarizes the knowledge as two social contextual factors.

Many social network based models have been proposed to improve the performance of the RS. Recently, Yang et al. [1] propose to use the concept of 'inferred trust circle' that is based on the domain-obvious circles of friends on social networks to recommend user favorite items. Their approach not only refines the interpersonal trust in the complex networks, but also reduces the load of big data. Meanwhile, besides the interpersonal influence, Jiang et al. [2] demonstrate that individual preference is also a significant factor in social network. Just like the idea of interpersonal influence, due to the preference similarity, user latent features should be similar to his/her friends' based on the probabilistic matrix factorization model. However, do all users actually need the relationship on the social networks to recommend items? Does the relationship submerge user's personality, especially for the experienced users? It is still a great challenge to embody users per personality in RS, and it is still an open issue that how to make the social factors be effectively integrated in recommendation model to improve the accuracy of RS.

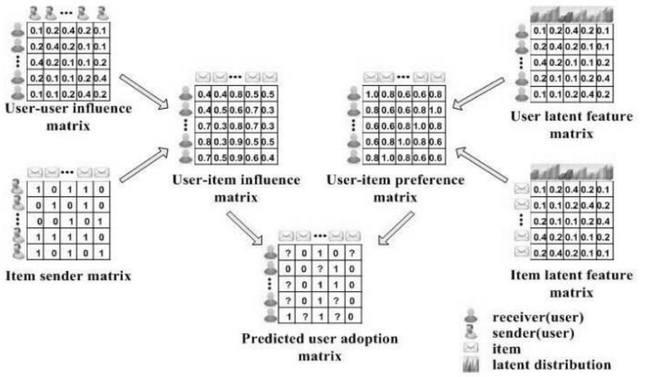

Figure 2. Illustrator on social contextual recommendation framework. Here the receiver is the user who receives the item such as post, reply, retweet, etc. The sender is the receiver's friend or followed user who generates the item in the form of post, tweet etc.

The recommendation qualities, usability of six online recommender systems were examined. Study results prove that the user's friends consistently provided better recommendations. For example almost 90% of users believe the book recommended is good from friends, 75% of such users believe in this recommendation that it is useful from friends. This research shows that the interpersonal influence is important in social media. Java et al. had analyzed that a widespread social network in a new form of social media known as micro-blogging. It has a high degree correlation and reciprocity, indicating close mutual acquaintances among users. They had identified different types of user intentions and studied the community structures. Categorizing friends into groups (e.g. family, co-workers) would come handy in the adoption of micro-blogging platforms to analyze user intentions. That is, user intentions or interests should be reflected by those of its friends. Rahman and Hailes provided a model for supporting trust in virtual communities based on experience and reputation. The significance of user's information can be seen such as the number of ratings in the category of their reputation or reliability. Yuan et al. have explored a kind of social relation, the membership with combined effect on friendship.

A. Related Works

Heterogeneous social relations of two types are fused in the Collaborative Filtering based recommender via a factorization process along with the distinguished effectiveness of social relationships in the sparse data condition was demonstrated. "Moves as one desires, decides as you like." As if the logo says, user's choice is always closely related to his/her personal interest. It is very popular for users to share, upload, and comment their favorite contents. Users' personal interests can be depicted by their historical rating records in social rating networks [39], [40], [41]. There are three social factors in this paper, personal interest, interpersonal interest similarity, and interpersonal influence are combined into a unified personalized recommendation model based on probabilistic matrix factorization.

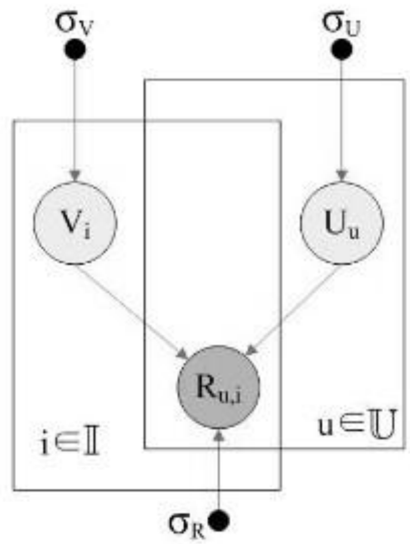

Figure 3. Graphical Model of the baseline factorization of user-item rating matrix.

The personality is denoted by user's interest to the topic of item. To embody the effect of user's personality, we mine the topic of item based on the naturalitem category tags of rating datasets. Hence, each item is shown by a category distribution or topic distribution vector that reflects the characteristic of the rating datasets. Moreover, we get user interest based on his/her rating behavior. We then assign to the effect of user's personality in our personalized recommendation model proportional to their expertise levels. The relationship of user-user in the social network contains two factors: interpersonal influence and interpersonal interest similarity. We apply the inferred trust circle of Circlebased Recommendation (CircleCon) model to enforce the factor of interpersonal influence. Also we infer interest circle to enhance the intrinsic link of user latent feature for the interpersonal interest similarity.

This paper contributes to following factors:

1. Proposal of a personalized recommendation system combining user's personal interest and interpersonal interest similarity as well as interpersonal influence. The user personal interest makes direct connections between the user and the item latent feature vectors. Other two social factors make connections between user and his/her friends' latent feature vectors.

2. Propose a personalized recommendation approach by enforcing user personal interests, which is category related and represented by a multi-level tree structure. To get an accurate model for the cold start user and those with very few friends and rated items personal unique interest is modeled. Also the impacts of those three factors to the recommendation performances are systematically compared. 3. Extensive experiments based on three datasets including Yelp, MovieLens show the effect of the proposed model to solve the user cold start and sparsity problem. 4. We also will be sharing our datasets for researchers in social recommendation area. The most salient feature of the shared datasets is that the objective social recommendation performance evaluation can be carried out.

ISSN: 2321-8169
4150 - 4155

II. PRELIMINARIES AND PROBLEM FORMULATION

Symbols and notations utilized in this paper are given in Table 1. For personalized RS, the system aims at recommending user interested items based on their historical behavior and interpersonal relationship of social networks. Moreover, we predict the ratings of user u on an unknown item i to measure how much user u interested in item i in social rating networks (like Netflix4, Yelp, Epinions). In RS, we have a set of users U {u1,...,uM} and a set of items P {i1,...,iM}.

TABLE 1
SYMBOLS AND THEIR DESCRIPTIONS UTILIZED IN THIS PAPER

Symbol	Description	Symbol	Description
U	a set of users	u	a user in the set of users
P	a set of items	i	an item in the set of items
$R_{M\times N}$	the rating matrix expressed by users on items	$\hat{R}_{M\times N}$	the predicted rating matrix based on the latent feature space
M	the number of users in the test social network	N	the number of items in the test social network
c	the category of the items (Table 3)	v	a friend of user u
F_u^c	the set of user u's friends in c	H_u^c	the set of items rated by user u in c
I	the indicator function	D	the topic distribution vector
$Q_{M\times N}$	the relevance matrix of user u's interest to the topic of item i	$W_{M\times M}$	the interest similarity matrix of user u to user v
$S_{M\times M}$	the matrix of user u trust on v	$U_{M\times k}$	the user latent feature matrix
$P_{N\times k}$	the item latent feature matrix	k	the dimension of the latent space
r	users' average rating value in the training dataset	λ, β, γ, η	the tradeoff parameters in the objective function
Ω	zero-mean spherical Gaussian priors	$\mathcal{N}$	Gaussian observation noise
Ψ	the objective function of the recommendation model	X^c	the variable corresponding to X in category c
X^{c*}	the normalized matrix of X in category c	X^T	the transposition of matrix X

The ratings expressed by users on items are given in a rating matrix R=[Ru,i]m*n. In this matrix, Ru,i denotes the rating of user u on item i. It can be any real number, but often ratings are integers in the range of 1 to 5. In a social rating network, each user u has a set of friends, and denotes the value of user u trust on user v or the influence of user v to user u. The trust values are given in a matrix S=[Su,p]m*n . Note that S is asymmetric in general, because the influence of user v to user u maybe different from the influence of user u to user v. Meanwhile, Wu,v denotes the interest similarity of user u to user v. The interest similarity values are given in a matrix , which is symmetric in general. And denotes the relevance of user u's interest to the topic of item i. The relevance values are given in a matrix, which represents users' personal interests.

The task of personalized recommender is as follows: Given a user and an item for which Ru,i is unknown, predict the rating for u on item i using R, S, W and Q. In this paper, we employ matrix factorization techniques [1], [2], [3], [4], [5] to learn the latent features of users and items, and predict the unknown ratings using these latent features. Let and be latent user and item feature matrices, with row vectors Uu and Pi represents k-dimensional user specific and item specific latent feature vectors of user u and item i, where k is very less than M and N, and it is the rank of the latent matrices U and P. Moreover, Uu and Pi can be seen as the brief characterization of user u and item i. Matrix factorization is to learn these latent variables and exploit them for recommendation.

III. MODEL

In this paper, we focus on probabilistic matrix factorization with consideration of factors of social network. Here we review some relevant works to this paper, including the basic matrix factorization model [4] without any social factors, the CircleCon model [1] with the factors such as interpersonal trust values and the Social Contextual (ContextMF) model [2] with interpersonal influence and individual preference.

A. Basic Matrix Factorization:

To introduce various sophisticated approaches [1], [2], [3], [5], we first briefly review the basic probabilistic matrix factorization (BaseMF) approach [4], which does not take any social factors into consideration. RS is to decrease the error of predicted value using R to the real rating value. The BaseMF model is trained on the observed rating data by minimizing the objective function

$$\Psi(R,U,P)=\frac{1}{2}\sum_{u,i}(R_{u,i}-\hat{R}_{u,i})^2+\frac{\lambda}{2}(\|U\|_F^2+\|P\|_F^2)$$

(1)where M is the number of users, N is the number of items, Ru,i is the real rating values in the training data for item i from user u, U and P are the user and item latent feature matrices which need to be learn from the training data,F X is the Frobenius norm of matrix X, and

$$\|X\|_F=\left(\sum_{i,j}x_{i,j}^2\right)^{1/2}$$

The second term is used to avoid over fitting[3].

$$\hat{R}=r+UP^T$$

This objective function can be minimized efficiently using gradient descent method in [3]. (2) where r is an offset value [1], which is empirically set as user's average rating value in the training data. Once U and P are learned by the gradient decent approach which are the low-rank matrices, which we detail in Section IV. And then, rating values can be predicted according to (2) for any user-item pairs.

B. CircleCon Model

The CircleCon model [1] has been found to outperform BaseMF and SocialMF [3] with respect to accuracy of the RS. The approach focuses on the factor such as interpersonal trust in social network as well as infers the trust circle. The user-user trust value is represented by the matrix S. Furthermore, the whole trust relationship in the social network is divided into several sub-networks Sc, called inferred circle [1], and each circle is related to a single category c of items. in the example item The Dakota Bar of New York that belongs to the category Night Life in Yelp as shown in Table 3. If user u rated the item, then user u is in the circle of category Night Life. In category c, the directed and weighted social relationship of user u with user v (the value of u trusts v or the influence of v to u) is represented by a positive value . Here c u F is the set of user u's friends in c. Four variants of defining interpersonal trust value are systematically compared in this model: 1) CircleCon1, which means each user v gets assigned the same trust value to user u in c; 2) CircleCon2a, , here is the set of items rated by user v in c and | c v H |is the total number of items in ; 3) CircleCon2b, c= c*, where c is the voting value in c from all followers of user v. It is a good indication that v is an expert in c, if most of v's followers have many ratings in c. 4) CircleCon3, trust splitting. Assume user u1 and user u2 both belong to category c1 and c2, u1 is a friend of u2 and the number of ratings u1 issued in category c1 and c2 are 7 and 3 respectively. The trust value in original social network is. To decrease the predicted error, the CircleCon model combines interpersonal trust value S with the rating matrix R, and the objective function is just like SocialMF [3], but the difference is that the CircleCon model is trained in each category. And the basic idea is that a user in social network may be influenced by other users, especially his/her friends in the same category. Thus their objective function is as follows:

$$\Psi(\boldsymbol{R},\boldsymbol{U},\boldsymbol{P},\boldsymbol{S}^*,\boldsymbol{W}^*)$$
$$=\frac{1}{2}\sum_{u,i}(R_{u,i}-\hat{R}_{u,i})^2+\frac{\lambda}{2}(\|\boldsymbol{U}\|_F^2+\|\boldsymbol{P}\|_F^2)$$
$$+\frac{\beta}{2}\sum_{u}((\boldsymbol{U}_u-\sum_{v}S^*_{u,v}\boldsymbol{U}_v)(\boldsymbol{U}_u-\sum_{v}S^*_{u,v}\boldsymbol{U}_v)^{\mathrm{T}})$$
$$+\frac{\gamma}{2}\sum_{u,v}(W^*_{u,v}-\boldsymbol{U}_u\boldsymbol{U}_v^{\mathrm{T}})^2$$

where the estimated ratings for a user is expressed as follows:

$$\hat{R}^c_{u,i}=r^c+\boldsymbol{U}^c_u\boldsymbol{P}^{c\mathrm{T}}_i$$

where is empirically set as user's average rating valuein category c. In (3), the factor of interpersonal trust is enforced by the last term in the objective function that says that user latent feature Uu should be similar to the average of their friend's latent features with weight ofin category c. Ratings only in category c are used to train user and item latent feature matrices U and P in this model. And once the model is trained in c, the rating value in c can be predicted according to (4).

C. ContextMF Model

Jiang et al. [2] demonstrate the significance of social contextual factors (including individual preference and interpersonal influence) for item adopting on real Facebook and Twitter style datasets. The task of ContextMF model in [2] is to recommend acceptable items from sender u to receiver v. The factor of interpersonal influence is identicle to the trust values in CircleCon model [1]. Moreover, individual preference is mined from receiver's historical adopted items. And matrix W represents the interpersonal preference similarity values. Each of the rows is normalized to unity. The objective function of this model is

$$\Psi(\boldsymbol{R},\boldsymbol{U},\boldsymbol{P},\boldsymbol{S}^*,\boldsymbol{W}^*)$$
$$=\frac{1}{2}\sum_{u,i}(R_{u,i}-\hat{R}_{u,i})^2+\frac{\lambda}{2}(\|\boldsymbol{U}\|_F^2+\|\boldsymbol{P}\|_F^2)$$
$$+\frac{\beta}{2}\sum_{u}((\boldsymbol{U}_u-\sum_{v}S^*_{u,v}\boldsymbol{U}_v)(\boldsymbol{U}_u-\sum_{v}S^*_{u,v}\boldsymbol{U}_v)^{\mathrm{T}})$$
$$+\frac{\gamma}{2}\sum_{u,v}(W^*_{u,v}-\boldsymbol{U}_u\boldsymbol{U}_v^{\mathrm{T}})^2$$

where the factor of individual preference is enforced by the last term in (5), which means that user latent feature Uu should be similar to their friend's latent feature with weight of their preference similarity in social network, and the rating values is predicted by (1). Once the model is trained, the rating values can be predicted according to (1) for any user-item pairs. Besides the interpersonal influence (similar to the trust values in CircleCon model [1]), individual preference is a novel factor in ContextMF model, and is enforced by the last term in (5). Note that we still execute the interpersonal influence as CircleCon model [1] and omit the topic relevance of items, as we also predict ratings of items in Epinions style datasets and use the circle based idea in our experiments. Although individual preference is proposed in this model, user u's latent feature is still connected with user's friends rather than their characteristic. The factor of individual preference of this model is enforced by interpersonal preference similarity. Comparing ContextMF model, the proposed personalized recommendation model has three differences: 1) Our model's task is to recommend user, regardless of sender-receiver, interested-unknown items. 2) user personal interest is directly related to his/her rated items rather than connect with his/her friends. 3) the factor of user interest in our model mined from user rated items has more influence than individual preference in ContextMF model, because it easier for the recommended items of our model to be transformed into purchase rate than the adopted items social networks like Facebok, Twitter etc.

 ISSN: 2321-8169
4150 - 4155

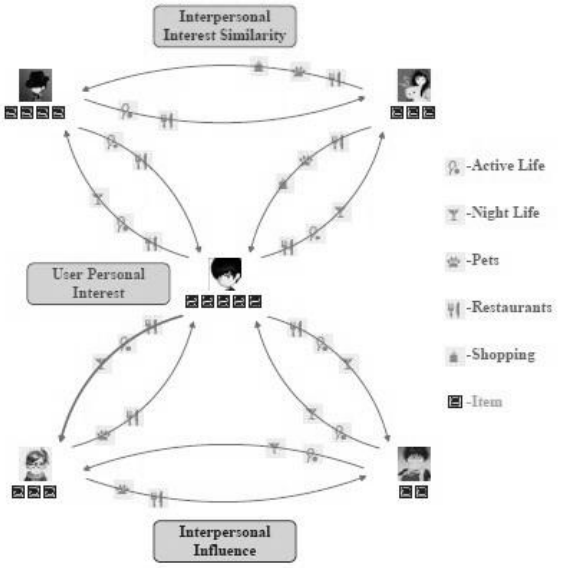

Figure 4. Interpersonal interest similarity and interpersonal influence of the user on social media.

IV. THE APPROACH

The proposed approach of the personalized recommendation fuses three social factors: user personal interest, interpersonal influence and interpersonal interest similarity to recommend user interested items. Our approaches are illustrated in Fig. 1. Out of the three factors the user personal interest and interpersonal interest similarity are important contributions of the approach and all related to user interest. Thus, we introduce user interest factor firstly. Later we infer the objective function of proposed personalized recommendation model. At last, we give the training approach of the model. Now we turn to the details of our approach.

A. User Interest Factor

Besides values of the trust between friends in the same category [1], user interest is another significant factor to affect their decision making process, which has been proved by psychology and sociology studies. Moreover, Jiang et al. [2] demonstrated the effect of ContextMF model considering both individual preference and interpersonal influence. Still there are two important differences of the user interest factor in our model to individual preference in ContextMF [2]: 1) The independence of user interest. It means we can items can be recommended based on user interest at a certain extent. In other words, we utilize user's connection with the items to train the latent feature vectors for the some experienced users. 2) Interest circle inference. Like CircleCon model [1], we divide the tested social network into several sub-networks, and each of them correspond to a signal category of items. As cold start users who are having less rating records friend's interest is used in the same category to link user latent feature vector.

User Interest Description: With respect to natural item category tags of rating datasets, we can get category distribution of the item, can be seen as the naive topic distribution of the item Di. For example, each item has the tags of its category in Yelp. Just like the item The Dakota Bar of New York belongs to the category Night Life. So Night Life is one of the category tags of the item. We summarize all the rated items to measure user interest from user's historical rating data in category c:

$$D_u^c = \frac{1}{\left|H_u^c\right|}\sum_{i \in H_u^c} D_i$$

where D is the set of items rated by user u in c. Here method we use to get the topic distribution Di is also different from ContextMF [2]. Here we apply the tree structure of categories of items shown in Fig. 2, to extract the topic distributions of items similar to that we utilized in [17], [18]. The first level of the tree structure is the big category of items. Consider an example that we have 8 big categories: 1 Active Life, 2 Beauty and Spas, 3 Home Services, 4 Hotels & Travel, 5 Night Life, 6 Pets, 7 Restaurants, and 8 Shopping according to Yelp data shown as Table 3. Subcategory of each big category in the first level is the second level.

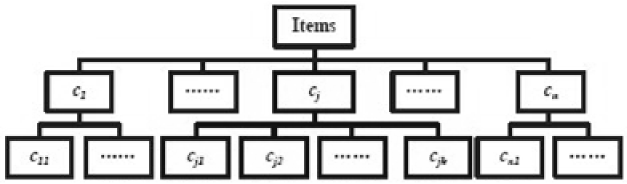

Figure 5. Structure of categories of items

Corresponding to the two level of the tree, we get two level topic distributions of each item in the datasets as following:

$$D1_i = [I_{c_1}, I_{c_2}, \ldots, I_{c_n}]$$
$$D2_i = [I_{c_{j1}}, I_{c_{j2}}, \ldots, I_{c_{jk}}], j \in [1, n]$$

V. CONCLUSION

We proposed ContextMF which is a novel social recommendation model utilizing social contextual factors such as individual preference and interpersonal influence. Extensive study on two large real-world social network datasets showed that social contextual information can greatly boost the performance of recommendation on social network datasets. Also, the proposed algorithm is general and can be easily adapted according to different real-world recommendation scenarios. A personalized recommendation approach was proposed by combining social network factors. The personal interest denotes user's individuality of rating items, especially for the experienced users, and these factors were fused together to improve the accuracy and applicability of recommender system. We conducted extensive experiments on three large real-world social rating datasets, and showed significant improvements over existing approaches that use mixed social network information. At present, the personalized recommendation model only takes user historical rating records and interpersonal relationship of social network into consideration. In our future works, we will take user location information to recommend more personalized and real-time items.

REFERENCES

[1] X. -W. Yang, H. Steck, and Y. Liu. Circle-based recommendation in online social networks. KDD'12, pp. 1267-1275, Aug.2012.

[2] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. -W. Zhu and S. -Q. Yang. Social contextual recommendation. CIKM'12, pp. 45-54,2012..

[3] M. Jamali and M. Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In Proc. ACM conference on Recommender systems (RecSys), 2010.

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

[5] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. Computer, pp. 30-37, Aug.2009.

[6] R. Bell, Y. Koren, and C. Volinsky. Modeling relationships at multiple scales to improve accuracy of large recommender systems. KDD'07, pp. 95-104, 2007.

[7] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. In WWW, pp. 285-295, 2001.

[8] G. -R. Xue, C. Lin, Q. Yang, W. Xi, H. -J. Zeng, Y. Yu, and Z. Chen.Scalable collaborative filtering using cluster-based smoothing. SIGIR'05, pp. 114-121, 2005.

[9] J. Chen, R. Nairn, L. Nelson, M. Bernstein, and E. Chi. "Short and tweet: experiments on recommending content from information streams." In Proceedings of the 28th international conference on Human factors in computing systems, pp. 1185-1194. ACM, 2010.

[10] R. Sinha, and K. Swearingen, Comparing recommendations made by online systems and friends[C]: Dublin, Ireland, 2001.

[11] A. Java, X. Song, T. Finin, and B. Tseng, Why we twitter: understanding microblogging usage and communities. In WebKDD/SNAKDD'07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, pages 56–65, NY, USA, 2007. ACM.

[12] A. Rahman, and S. Hailes, Supporting trust in virtual communities. In Proceedings of the 33rd Hawaii International Conference on System Sciences. Maui, HW, USA (2000).

[13] Q. Yuan, Li Chen and Shiwan Zhao. Factorization vs. Regularization: Fusing Heterogeneous Social Relationships in Top-N Recommendation. Chicago, Illinois, USA, 2011, ACM.

[14] M. Ou, P. Cui, F. Wang, J. Wang, W. Zhu, S. Yang, "Comparing Apples to Oranges: A Scalable Solution with Heterogeneous Hashing," ACM SIGKDD, 2013.

[15] P. Cui, F. Wang, S. Liu, M. Ou, S. Yang, "Who Should Share What? Item-level Social Influence Prediction for Users and Posts Ranking," International ACM SIGIR Conference (SIGIR), 2011.

[16] J. Shen, D. Tao, X. Li, "Modality Mixture Projections for Semantic Video Event Detection," IEEE Trans. Circuits Syst. Video Techn. 18(11): 1587-1596 (2008).

[17] H. Feng, and X. Qian, "Mining User-Contributed Photos for Personalized Product Recommendation", Neurocomputing, 2013.

[18] H. Feng, and X. Qian, "Recommend Social Network Users Favorite Brands", PCM 2013.

[19] S. Jiang, X. Qian, T. Mei, K. Lan, and L. Zhang, "Mobile multimedia travelogue generation by exploring geo-locations and image tags", In Proc. ISCAS, 2013, pp.881-884.