Collaborative Social Group Influence for Event Recommendation

Li Gao[†], Jia Wu[⋆], Zhi Qiao[‡], Chuan Zhou[†], Hong Yang[§], Yue Hu[†]
[†] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100093, China
[⋆] Centre for Quantum Computation and Intelligent Systems (QCIS), University of Technology Sydney, NSW 2007, Australia
[‡] IBM Research - China, Beijing, 100193, China
[§] MathWorks, Beijing, 100190, China
{gaoli, huyue, zhouchuan}@iie.ac.cn; jia.wu@uts.edu.au; qzbj@cn.ibm.com; hong.yang@mathworks.cn

ABSTRACT

In event-based social networks, such as *Meetup*, social groups refer to self-organized communities that consist of users who share the same interests. In many real-world scenarios, users usually have social group preference and join interested social groups to attend events. It is therefore necessary to consider the influence of social groups to improve the event recommendation performance; however, existing event recommendation models generally consider users' individual preferences and neglect the influence of social groups. To this end, we propose a new Bayesian latent factor model *SogBmf* that combines *social group influence* and *individual preference* for event recommendation. Experiments on real-world data sets demonstrate the effectiveness of the proposed method.

CCS Concepts

•Information systems → Data mining; •Human-centered computing → *Social recommendation;*

Keywords

Event recommendation; Social group influence

1. INTRODUCTION

Recent years have witnessed the rapid growth and popularity of event-based social networks (EBSNs), which are convenient tools with which users can organize and participate in social events, *e.g.,* cocktail parties, playing sports and attending technical conferences. To enhance user experience, given the sheer volume of available events in EBSNs, event recommender systems [2, 6, 10] have been developed to suggest the events that users are most likely to attend.

A typical event-based social network, such as *Meetup*, integrates a new module called *social groups* which refer to self-organized communities that consist of users who share the same interests. By categorizing users' social relationships into several distinct social groups, social events can be promoted within related social groups to attract more users. Users of EBSNs join social groups and often share comments to promote events, which indicates that social groups are necessary tools for connecting users and events. Based on this observation, it is clearly necessary to consider the influence of social groups when attempting to improve the performance of event recommender systems.

The recent work in [5] presents a group preference-based Bayesian personalized ranking method for recommendation; however, the formulated social groups in this method depend on the interactions and collaborations between users who give positive feedback on a specific item. The study in [3] presents a context-aware approach, in which the information from social groups is used to regularize the user and event latent factor matrices. In deed, the group-collaborative event recommendation encounters the following challenges: 1) how to properly formulate the social group influence, and 2) how to strategically embed social group influence to achieve accurate event rating scores.

Motivated by the above observations, our work aims to build a more realistic model of social group influence on the rating scores generated by the interactions between user-group preferences and group-event correlation, where user-group preference is regarded as latent factor (Challenge #1). We propose a Bayesian latent factor model (denoted as *SogBmf*) for event recommendation, based on the matrix factorization framework, to integrate social group influence with individual preference (Challenge #2). A gradient descent algorithm is applied to learn the parameters. Experimental results on real-world data sets empirically demonstrate that our method consistently achieves an improvement in accuracy.

CIKM'16 , October 24-28, 2016, Indianapolis, IN, USA

DOI: http://dx.doi.org/10.1145/2983323.2983879

2. PROBLEM STATEMENT

The problem of event recommendation is formulated as follows: suppose that we have a set of users, $S = \{u_1, u_2, \dots, u_n\}$ with the ith user denoted as u_i, and a set of events, $I = \{v_1, v_2, \dots, v_m\}$ with the jth event denoted as v_j. We denote the user-event interaction matrix as $R \in \{0,1\}^{n\times m}$. If user u_i joins event v_j, the entry R_{ij} is 1; otherwise, it is 0. Then, the problem turns to predict the unobserved entries in R and rank the event v_j for each user u_i by score $r(u_i, v_j)$ which indicates user u_i's preference for event v_j.

Predicting the rating score $r(u_i, v_j)$ plays a key role. In this work, there are three types of entities: users, groups and events, and their relationships, including 1) user/group relationships where users join the social groups, 2) group/event relationships where social groups post events, and 3) user/event relationships where users participate in events. We **aim** to embed the influence of the social groups into the matrix factorization framework [4] for event recommendation, based on these data.

3. METHODOLOGY

We propose a Bayesian latent factor model, denoted as *SogBmf*, which embeds social group influence for event recommendation.

3.1 Mixture Rating for Recommendation

As shown in Figure 1, we propose a mixture rating framework for the score $r(u_i, v_j)$ for event recommendation, which factorizes the user-event interaction/rating matrix into the social group influence rating and individual preference rating matrices. The matrices given in Figure 1 are defined as follows.

DEFINITION 1. *(User-group preference matrix). $F \in \mathbb{R}^{g\times n}$ denotes the user-group preference matrix, where each entry F_{ij} denotes the degree of preference for the jth user toward the ith group. Generally, $F_{ij} > 0$ if the jth user joins the ith group.*

DEFINITION 2. *(Group-event matrix). The group-event matrix is denoted by $C \in \{0,1\}^{g\times m}$. $C_{ij} = 1$ means that the jth event is posted in the ith social group.*

DEFINITION 3. *(Latent factor matrix). The user latent factor matrix is denoted by $U \in \mathbb{R}^{k\times n}$, where user i is determined by a latent factor vector $U_i \in R^k$. Similarly, the event latent factor matrix is denoted by $V \in \mathbb{R}^{k\times m}$, where event j is described by a latent factor vector $V_j \in R^k$.*

Thus, the rating score $r(u_i, v_j)$ is calculated as:

$$r(u_i, v_j) = r_{sg}(u_i, v_j) \oplus r_{ip}(u_i, v_j) \tag{1}$$

where $r(u_i, v_j)$ is the ultimate rating score by combining the social group influence rating $r_{sg}(u_i, v_j)$ and the individual preference rating $r_{ip}(u_i, v_j)$, where the new operator $r_1 \oplus r_2 = \alpha r_1 + (1-\alpha) r_2$ with $0 \le \alpha \le 1$ is a tradeoff parameter.

Social group influence rating. Let F_i be a social group preference vector of user u_i, C_j be an event v_j vector posted by groups. Then $r_{sg}(u_i, v_j)$ is denoted as:

$$r_{sg}(u_i, v_j) = F_i^T C_j \tag{2}$$

Individual preference rating [1]. Let $U_i \in R^k$ be user u_i's latent factor vector, $V_j \in R^k$ be event v_j's latent factor vector. $r_{ip}(u_i, v_j)$ can be denoted as:

$$r_{ip}(u_i, v_j) = U_i^T V_j \tag{3}$$

Thus, the problem of event recommendation needs to solve **F**, **U** and **V**, so that the mixture rating can be calculated as follows:

$$R = \alpha F^T C + (1-\alpha) U^T V \tag{4}$$

3.2 Our SogBmf Model

The conditional distribution with Gaussian observation noise over the observed entries in R is defined as:

$$P(R|F,U,V,\sigma_R^2) = \prod_{i=1}^{n}\prod_{j=1}^{m}[\mathcal{N}(R_{ij}|\alpha F_i^T C_j + (1-\alpha) U_i^T V_j, \sigma_R^2)]^{I_{ij}} \tag{5}$$

where I_{ij} is the indicator function that is equal to 1 if user u_i joins event v_j and equal to 0 otherwise.

Generally, users' feedback scores on events are implicit [6], thus Bayesian personalized ranking (BPR) [8] can be used to effectively

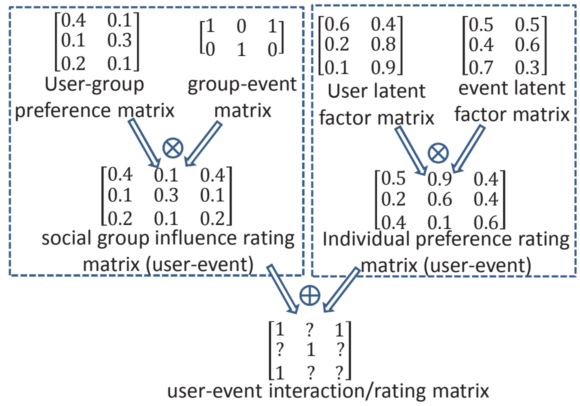

Figure 1: The mixture rating for event recommendation, which incorporates social group influence and individual preference to factorize the user-event interaction/rating matrix into a social group influence rating matrix and an individual preference rating matrix.

solve the implicit feedback recommendation problem. Let P_u^I be the positive event set for which the entry gets positive feedback from user u, and let N_u^I denote the negative event set for which the entry gets negative feedback. Based on the BPR optimization criterion [8], Eq. (5) can be extended as Eq. (6), where $R_>$ denotes that events in P_u^I are ranked higher than events in N_u^I.

$$P(R_>|F,U,V) = \prod_{(u_i,v_j,v_k)\in(S,P_u^I,N_u^I)} P(r(u_i,v_j) > r(u_i,v_k)|F,U,V) \tag{6}$$

In EBSNs, users join social groups and often share comments to promote social events. Social groups reflect social relationships. We use $H \in \mathbb{R}^{n\times n}$ to denote social relationships between users that are defined as follows:

$$H_{ij} = \frac{|G(u_i)\cap G(u_j)|}{|G(u_i)\cup G(u_j)|} \tag{7}$$

where $G(u_i)$ denotes the social groups joined by user u_i, and $|\mathcal{X}|$ denotes the cardinality of the set $\mathcal{X}$.

In terms of events v_k, if users u_i and u_j co-participate in v_k, they share the same interest. We use $Q \in \mathbb{R}^{n\times n}$ to represent the interest similarity between users

$$Q_{ij} = \frac{|E(u_i)\cap E(u_j)|}{|E(u_i)\cup E(u_j)|} \tag{8}$$

where $E(u_i)$ denotes the social events attended by user u_i.

We hypothesize that the similarity in the observed space is consistent with the latent space. Thus, we regularize the latent space by the observed matrices H and Q in that:

- Users that are similar in the user latent space U have similar event preferences which can be derived from the interest similarity matrix Q;

- Users that are similar in the social group preference latent space F have tight social connections which can be derived from the social relations matrix H.

The posterior distribution is formulated as follows:

$$
\begin{aligned}
&P(F,U,V|R_>,C,Q,H,\Omega)\\
&=\frac{P(R_>,C,Q,H|F,U,V,\Omega)P(F,U,V|\Omega)}{P(R_>,C,Q,H,\Omega)}\\
&\propto P(R_>|F,U,V)P(Q|U,\Omega)P(H|F,\Omega)P(F|\Omega)P(U|\Omega)P(V|\Omega)\\
&=\prod_{(u_i,v_j,v_k)\in(S,P_u^I,N_u^I)} P(r(u_i,v_j)>r(u_i,v_k)|F,U,V)\\
&\quad\prod_{s,t}\mathcal{N}(Q_{st}|U_s^TU_t,\sigma_Q^2)\prod_{p,q}\mathcal{N}(H_{pq}|F_p^TF_q,\sigma_H^2)\\
&\prod_x\mathcal{N}(F_x|0,\sigma_F^2)\prod_y\mathcal{N}(U_y|0,\sigma_U^2)\prod_z\mathcal{N}(V_z|0,\sigma_V^2) \qquad (9)
\end{aligned}
$$

where Ω denotes that the zero-mean spherical Gaussian prior is placed on latent factor vectors and observed matrices.

Based on the BPR optimization criterion [8], we set $P(r(u_i,v_j)>r(u_i,v_k)|F,U,V)=l(r(u_i,v_j)-r(u_i,v_k))$, where $l(x)=\frac{1}{1+e^{-x}}$ is the logistic function. We expect to maximize the log-posterior distribution $\ln P(F,U,V|R_>,C,Q,H,\Omega)$, which can be converted to find F, U and V to minimize the function with hybrid quadratic regularization terms as follows:

$$
\begin{aligned}
\min_{F,U,V}\ \mathcal{O}=&\sum_{(u_i,v_j,v_k)\in(S,P_u^I,N_u^I)}\ln(1+e^{r(u_i,v_k)-r(u_i,v_j)})\\
&+\lambda_Q\|Q-U^TU\|_F^2+\lambda_H\|H-F^TF\|_F^2\\
&+\lambda_F\|F\|_F^2+\lambda_U\|U\|_F^2+\lambda_V\|V\|_F^2 \qquad (10)
\end{aligned}
$$

where λ_Q, λ_H, λ_F, λ_U and λ_V are non-negative regularization parameters, and $\|\cdot\|_F$ is the Frobenius norm.

3.3 Parameter Learning

Let the function $I(v_j\in P_{u_i}^I, v_k\in N_{u_i}^I)$ ($I(i,j,k)$ for short) denote the indicator function, which equals 1 if both $v_j\in P_{u_i}^I$ and $v_k\in N_{u_i}^I$ stand; otherwise, it equals 0. Let function $J(i,j,k)=\frac{e^{r(u_i,v_k)-r(u_i,v_j)}}{1+e^{r(u_i,v_k)-r(u_i,v_j)}}$. We can then obtain the gradient descent of the optimization function given in Eq. (10) with respect to the latent parameters F, U and V as follows:

$$
\begin{aligned}
\frac{\partial\mathcal{O}}{\partial U_i}=&\sum_{v_j\in P_{u_i}^I}\sum_{v_k\in N_{u_i}^I}(1-\alpha)J(i,j,k)I(i,j,k)(V_k-V_j)\\
&-2\lambda_Q\sum_{s=1}^{n}U_s(Q_{is}-U_i^TU_s)+2\lambda_UU_i \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\mathcal{O}}{\partial V_p}=&\sum_{u_i\in S}\sum_{q=1}^{m}-(1-\alpha)\{J(i,p,q)I(i,p,q)U_i-J(i,q,p)\\
&I(i,q,p)U_i\}+2\lambda_VV_p \qquad (12)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\mathcal{O}}{\partial F_i}=&\sum_{v_j\in P_{u_i}^I}\sum_{v_k\in N_{u_i}^I}\alpha J(i,j,k)I(i,j,k)(C_k-C_j)\\
&-2\lambda_H\sum_{t=1}^{n}F_t(H_{it}-F_i^TF_t)+2\lambda_FF_i \qquad (13)
\end{aligned}
$$

The parameters update rules are thus as follows:

$$
U_i=U_i-\gamma_u\frac{\partial\mathcal{O}}{\partial U_i},V_j=V_j-\gamma_v\frac{\partial\mathcal{O}}{\partial V_j},F_i=F_i-\gamma_f\frac{\partial\mathcal{O}}{\partial F_i} \qquad (14)
$$

Algorithm 1 shows *SogBmf* model for event recommendation.

Algorithm 1 The *SogBmf* algorithm.

Input: User-event interaction R, group-event matrix C, users social relations H, users interest similarity Q. Parameters α, γ_u, γ_v, γ_f, ξ and the regularization parameters.

Output: User-group preference F, latent factor matrices U, V.

1: Initialize U, V and F with randomly generated values
2: Set index $b=0$
3: Compute the objective function $\mathcal{O}^b=\mathcal{O}(U,V,F)$
4: **repeat**
5: Compute the gradients as in Eq. (11), Eq. (12), Eq. (13)
6: Update the parameters U, V and F as in Eq. (14)
7: $b=b+1$
8: Compute the objective function $\mathcal{O}^b=\mathcal{O}(U,V,F)$
9: **until** $\frac{|\mathcal{O}^{b-1}-\mathcal{O}^b|}{\mathcal{O}^{b-1}}\le\xi$
10: **return** the parameters F, U and V.

4. EXPERIMENTS

4.1 Experimental Settings

4.1.1 Data Sets

We select the data from Chicago, Los Angeles, and New York City from *Meetup* for analysis, as these cities have a large number of group/event pairs. The data sets are extracted from the work [2]. Details of the data sets are summarized in Table 1.

Table 1: Data set statistics

	Chicago	Los Angeles	New York
#users	59504	89201	205592
#events	86988	167254	234160
#groups	8007	14657	21155

4.1.2 Setups

We randomly partition each of the three data sets into two non-overlapping sets for training and testing. For each user u_i, 70% of the historical events P_u^I are randomly sampled as the training data, and the remaining 30% are used for testing. We then randomly take one historical event of each user from the training data to construct the validation data. A similar setup was used in the previous work [5].

We repeat the above steps for five times and thus receive five copies of training data and test data. The reported experimental results are the average of the five runs.

4.1.3 Metrics

We use two metrics to evaluate the event recommendation performance, namely, AUC and NDCG@n. AUC in the evaluation can be formulated as

$$
AUC=\frac{1}{|S|}\sum_{u\in S}\frac{1}{|P_u^I|}\sum_{v_j\in P_u^I}\frac{1}{|N_u^I|}\sum_{v_k\in N_u^I}I(r(u,v_j)>r(u,v_k)),
$$

and NDCG@n is used to compare the top-n recommendation performance [5].

4.1.4 Benchmark Methods

We implement the following methods for comparison, including 1) *PMF* [4], a probabilistic matrix factorization method for recommendation; 2) *BPR-MF* [8], the recommendation model dealing with implicit feedback in event recommendation; 3) *GBPR* [5].

Table 2: Experimental results of the compared methods *w.r.t.* AUC and NDCG@5.

Data sets	Metrics	*PMF*	*BPR-MF*	*GBPR*	*SogBmf*
Chicago	AUC	$0.7142_{\pm 0.0004}$	$0.8453_{\pm 0.0009}$	$0.8645_{\pm 0.0011}$	$\mathbf{0.8894}_{\pm 0.0005}$
	NDCG@5	$0.1711_{\pm 0.0009}$	$0.2194_{\pm 0.0003}$	$0.2308_{\pm 0.0004}$	$\mathbf{0.2723}_{\pm 0.0013}$
Los Angeles	AUC	$0.7611_{\pm 0.0010}$	$0.8521_{\pm 0.0005}$	$0.8724_{\pm 0.0004}$	$\mathbf{0.8948}_{\pm 0.0004}$
	NDCG@5	$0.1916_{\pm 0.0010}$	$0.2231_{\pm 0.0005}$	$0.2417_{\pm 0.0005}$	$\mathbf{0.2744}_{\pm 0.0005}$
New York	AUC	$0.7384_{\pm 0.0005}$	$0.8536_{\pm 0.0017}$	$0.8713_{\pm 0.0005}$	$\mathbf{0.8967}_{\pm 0.0007}$
	NDCG@5	$0.1842_{\pm 0.0005}$	$0.2128_{\pm 0.0005}$	$0.2421_{\pm 0.0012}$	$\mathbf{0.2931}_{\pm 0.0006}$

We adapt *GBPR* for event recommendation, based on group preferences that indicate the interaction and collaboration between users.

4.1.5 Parameters Setting

The parameter α is searched from $\alpha \in \{0.2, 0.4, 0.6, 0.8, 1\}$ and the regularization parameters in Eq. (10) are selected from $\lambda_Q = \lambda_H = \lambda_F = \lambda_U = \lambda_V \in \{0.001, 0.01, 0.1\}$. The learning rate for the gradient descent method is selected from $\gamma_u = \gamma_v = \gamma_f \in \{0.001, 0.01, 0.1\}$. The AUC performance on the validation data for the three data sets is used to select the best parameters. As a result, parameter α is set to $\alpha = 0.2$, the regularization parameters are set to 0.1, and the learning rate for the gradient descent is set to 0.01. The dimension of the latent factor matrix is set to $k = 100$. We use the above tuned parameters for evaluation.

4.2 Experimental Results

The experimental results are shown in Table 2, from which we make the following observations: 1) On all three data sets, the proposed *SogBmf* method always obtains better results than the other baseline methods, measured by the evaluation metrics. The improvements over the baselines are statistically significant with $p < 0.005$ using the Wilcoxon signed rank significance test. 2) *SogBmf* and *GBPR* obtain better performance than *PMF* and *BPR-MF*, demonstrating that embedding social group information can improve the performance of event recommendation. 3) *SogBmf* obtains better results than *GBPR*. The social groups in *GBPR* are virtual, while the groups in EBSNs are self-organized communities consisting of users who share the same interests, which demonstrates the greater effect of social group influence on event recommendation.

5. RELATED WORK

Several recommendation methods based on social group information and users' preferences for social groups [7, 9] have been proposed, but they do not target event recommendation. In contrast, this paper focuses on applying the influence of social groups to event recommender systems.

The work [2] proposed to use online and offline patterns of interaction for recommending new events. Later studies [6, 10] extended this work to improve recommendation performance, but they did not incorporate the effects of social group influence on event recommendation. Based on the interactions and collaborations between users who give positive feedback on a specific item, [5] proposed a recommendation method that used group preferences. The literature [3] exploited social-aware group information for event recommendation, in which group information was used to regularize the user and event latent factor matrices. However, they have not properly formulated social group influence on the event rating scores. In this paper, we embed social group influence into the matrix factorization framework to further reflect the interactions between users and events.

6. CONCLUSIONS

In this paper, we study a new problem of event recommendation embedding with social group influence in EBSNs. We present a new Bayesian latent factor model *SogBmf* to consider social group influence and individual preference for improving the performance of event recommender systems. Experiments on real-world data demonstrate the effectiveness of the proposed method.

Acknowledgements

This work was supported by the 973 Program (No. 2013CB329605), NSFC (No. 61502479 and 61370025), Strategic Leading Science and Technology Projects of CAS (No. XDA06030200), Xinjiang Uygur Autonomous Region Science and Technology Project (No. 201230123), Australian Research Council (ARC) Discovery Projects (No. DP140100545), and China Scholarship Council Foundation (No. 201206410056). C. Zhou is the corresponding author.

7. REFERENCES

[1] M. Jiang, P. Cui, F. Wang, W. Zhu, and S. Yang. Scalable recommendation with social contextual information. *IEEE Trans. Knowl. Data Eng.*, 26(11):2789–2802, 2014.

[2] X. Liu, Q. He, Y. Tian, W.-C. Lee, J. McPherson, and J. Han. Event-based social networks: linking the online and offline social worlds. In *SIGKDD*, pages 1032–1040, 2012.

[3] A. Q. Macedo, L. B. Marinho, and R. L. Santos. Context-aware event recommendation in event-based social networks. In *RecSys*, pages 123–130, 2015.

[4] A. Mnih and R. Salakhutdinov. Probabilistic matrix factorization. In *NIPS*, pages 1257–1264, 2007.

[5] W. Pan and L. Chen. GBPR: Group preference based Bayesian personalized ranking for one-class collaborative filtering. In *IJCAI*, pages 2691–2697, 2013.

[6] Z. Qiao, P. Zhang, Y. Cao, C. Zhou, L. Guo, and B. Fang. Combining heterogenous social and geographical information for event recommendation. In *AAAI*, pages 145–151, 2014.

[7] L. Quijano-Sanchez, J. A. Recio-Garcia, B. Diaz-Agudo, and G. Jimenez-Diaz. Social factors in group recommender systems. *ACM Trans. Intell. Syst. Technol.*, 4(1):1–30, 2013.

[8] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*, pages 452–461, 2009.

[9] A. Salehi-Abari and C. Boutilier. Preference-oriented social networks: Group recommendation and inference. In *RecSys*, pages 35–42, 2015.

[10] W. Zhang and J. Wang. A collective Bayesian poisson factorization model for cold-start local event recommendation. In *SIGKDD*, pages 1455–1464, 2015.