Combining Heterogenous Social and Geographical Information for Event Recommendation

Zhi Qiao[1,3], Peng Zhang[2], Yanan Cao[2], Chuan Zhou[2], Li Guo[2], Bingxing Fang[1]
[1]Institute of Computing Technology of the Chinese Academy of Sciences
[2]Institute of Information Engineering, Chinese Academy of Sciences
[3]University of the Chinese Academy of Sciences
qiaozhi@nelmail.iie.ac.cn, {zhangpeng, caoyanan, zhouchuan, guoli, fangbinxing}@iie.ac.cn

Abstract

With the rapid growth of event-based social networks (EBSNs) like *Meetup*, the demand for event recommendation becomes increasingly urgent. In EBSNs, event recommendation plays a central role in recommending the most relevant events to users who are likely to participate in. Different from traditional recommendation problems, event recommendation encounters three new types of information, *i.e.*, heterogenous online+offline social relationships, geographical features of events and implicit rating data from users. Yet combining the three types of data for offline event recommendation has not been considered. Therefore, we present a Bayesian latent factor model that can unify these data for event recommendation. Experimental results on real-world data sets show the performance of our method.

Introduction

Recent years have witnessed increased development and popularity of event-based social networks (EBSNs), such as *Meetup*, *Plancast* and *Douban Event*. The services allow users to organize, participate, comment and share events such as cocktail parties, seminars and concerts. To enhance user experience on these services, event recommendation systems are studied lately that aims to recommend the most relevant events to users who are likely to participate. The recommendation provides convenience to both event organizers and participants. For the participants, they can easily find the events they are interested in. For the organizers, their events can attract potential users who share similar interests.

	e_1	e_2	e_3	e_4	...	$e_{\|E\|-1}$	$e_{\|E\|}$
u_1	0	0	0	0	...	1	1
u_2	1	0	0	1	...	1	0
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
$u_{\|U\|-1}$	0	0	1	1	...	0	1
$u_{\|U\|}$	0	1	0	0	...	0	0

Figure 1: An illustration of user-event ratings.

Different from existing recommendation problems, event recommendation in EBSNs meets new characteristics.

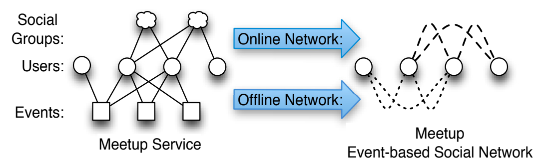


Figure 2: An illustration of Event-based Social Networks (Liu et al. 2012).

- Heterogenous social relationships. The event-based social networks (EBSNs) are proposed in a recent work (Liu et al. 2012), as shown in Fig. 2. There are both online and offline data in EBSNs. The online social service data reflect online social relationships, while the offline social interaction data are represented by event-based services.
- Geographical information. According to our observation, a user tend to participate in nearby events, and the possibility of the user participating in an event decreases with the distance increasing. Moreover, most events locate at areas with dense entertainment facilities, such as shopping malls, gymnasiums, theaters and bars.
- Implicit Rating. User rating data reveal the presence/absence of an event. In *Meetup*, the rating data are reflected by the RSVP choice ("yes", "no"). [1]. As shown in Fig. 1, the rating data are different from those in traditional recommendation systems where explicit ratings are provided with values $0 \sim 5$ or above.

The above new properties in EBSNs need to be carefully considered for offline event recommendation. In this paper, we present a Bayesian latent factor model that combines **He**terogenous **S**ocial Information and **G**eographical information (***HeSig*** for short) for event recommendation. Experiments on real-world data verify the performance of the method.

Related work

We briefly survey three aspects of work that are technically related to this work.

[1]http://meetupblog.meetup.com/post/20064732882/no-more-maybe-option-for-rsvps

Table 1: Statistics of the Data sets.

	NYC	LA	Houston	Chicago	SF
users	338144	124040	36199	89796	119569
events	108170	54538	16694	36009	45213

Recommender system with social information. How to wisely utilize social network information for recommendation has been extensively studied in recent years. There are some pioneer work in the literature (Xin et al. 2009; Ma et al. 2011; 2008; Liua and Lee 2010) which studied the social recommendation problems. In order to predict review quality, Lu et al. (Lu et al. 2010) proposed a generic framework to incorporate social context information (author identity and social networks) by adding regularization constraints. In the paper, we apply the social information to improve the recommendation performance.

Location-aware Recommendation. Recommendation with geographical information has been extensively studied in recent years (Cheng et al. 2012; Ye et al. 2011; Zheng et al. 2010; Qiao et al. 2014). Several simple yet practical methods were applied in location-aware recommendation. For example, the KNN technique (Chaudhuri and Gravano 1999; Bruno, Gravano, and Marian 2004) that retrieve the top k objects nearest to a user, the preference based method skylines (Takeuchi and Sugimoto 2005) and location-based methods (Borzsonyil, Kossmann, and Stocker 2011) that require explicit preference constraints. However, the existing methods didn't directly consider the impacts and correlation between geographical information and preference.

Latent Factor Models. The matrix factorization method as a popular latent factor models is widely used in recommendation: (Salakhutdinov and Mnih 2008; Koren, Bell, and Volinsky 2009; Somekh, Aizenberg, and Koren 2014; Wang and Blei 2011; Zhang et al. 2014). Koren at al. (Koren, Bell, and Volinsky 2009) conducted a line of work on matrix factorization by considering time factor, bias influence factor, etc. In (Somekh, Aizenberg, and Koren 2014), Aizenberg et al. incorporated artist-enhanced latent factor into matrix factorization to alleviate the sparsity problem in recommending music. Due to its practicability and flexibility, we extend the latent factor model to combine geographical information in the paper.

In conclusion, the existing studies cover only a portion of available information in event recommendation tasks. In this paper, we propose a Bayesian model that can combine geographical and heterogenous social impact data through an extended latent matrix model.

Preliminary

We select for analysis the data from New York City(NYC), Los Angeles(LA), Houston, Chicago and San Francisco(SF), as these cities have a large number of group/event pairs. The data sets are extracted from the work (Liu et al. 2012). The details of the data sets are shown in Table 1.

Geographical characteristics *Regional concentration of events*. Based on the data statistics, we found that most events were located around centers and concentration areas. Commonly, these concentration areas have more entertainment facilities that are essential to supply comfortable spaces for participants. For example, we use the Houston data and set k to 20. The results in Fig. 3 show the phenomena of regional concentration of events, where we use k-Means clustering to find these concentration areas. Note that different k leads to different cluster granularity. Similar results have been observed in other four data sets.

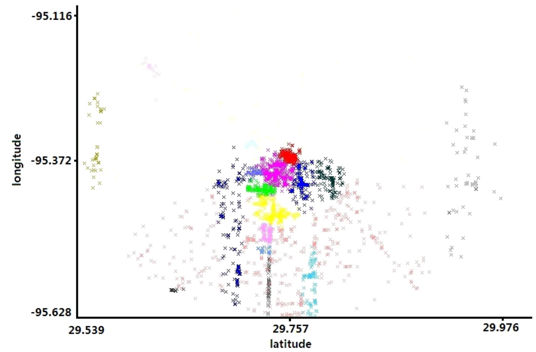


Figure 3: Geographical analysis for social events. Each node represents an event's location, a node set with the same color represents a geographical cluster.

Regional preference of users. We also observe that the personal preference of a user may vary from one area to another. Take Houston data for example and let number of areas is 20. In Fig. 4 over 80% users visit less than 6 areas in the whole 20 candidate areas. Fig. 5 shows the statistical results on average visit number of area for each user, where the average visit number of area is higher than 1 for most of users on each data sets. From Figs.4 and 5, we can observe that *users are likely to participate in the events that will happen in their favorite regions*.

Heterogeneous social relationships *Online social relationship*. In event-based social services, *social group* reflects social relationship. In *Meetup*, users in the same social group often share comments, photos and event plans. We use the edge weight w_{ij}^{on}, as in Eq. (1), to represent online social relationship between users i and j.

$$w_{ij}^{on} := \frac{|G(u_i) \cap G(u_j)|}{|G(u_i) \cup G(u_j)|} \tag{1}$$

where $G(u_i)$ represents all social groups that contain user i, $|G(u_i) \cap G(u_j)|$ denotes the number of nodes in the set $G(u_i) \cap G(u_j)$, and $|G(u_i) \cup G(u_j)|$ denotes the number of nodes in the set $|G(u_i) \cup G(u_j)|$.

Offline social relationship. The offline network is constructed based on the co-participation of social events. If users u_i and u_j co-participate in a social event, they share a connection. Let w_{ij}^{off} represent offline social relation between users i and j, the offline social relation can be described as in Eq. (2),

$$w_{ij}^{off} := \frac{|E(u_i) \cap E(u_j)|}{|E(u_i) \cup E(u_j)|}, \tag{2}$$

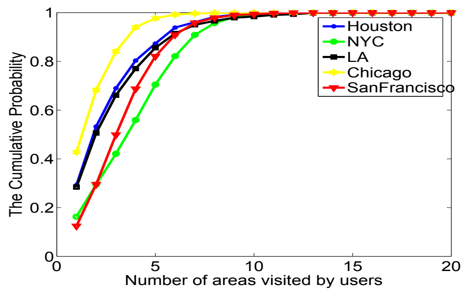


Figure 4: The Cumulative Probability Distribution of number of areas visited by users.

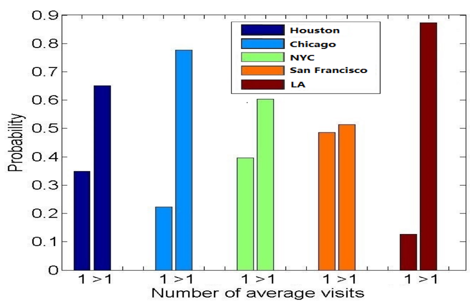


Figure 5: Distribution of average visit number of area.

where $E(u_i)$ represents the set of events user i has participated, $|E(u_i) \cap E(u_j)|$ denotes the number of nodes in the set $E(u_i) \cap E(u_j)$, and $|E(u_i) \cup E(u_j)|$ denotes the number of nodes in the set $|E(u_i) \cup E(u_j)|$.

Problem Description

For the task of event recommendation, we have four types of entities: S(user), L^S(user location), E(event) and L^E(event location), and two additional social networks: G^{On}(online social networks) and G^{Off}(offline social networks).

Let $S = \{u_1, u_2, ..., u_n\}$ denote the set of users. For each user $u_i \in S$, we use a unique location L_i^u to indicate the geographical information. Let the set of events be $V = \{v_1, v_2, ..., v_n\}$. For each event $e_j \in V$, it has a location L_j^V . The Online social network G^{On} represents social relationship among users and the offline social network G^{Off} represents offline relationship among users.

Formally, the problem is defined as ranking all events for each user, according to the dyadic rating score $r(u, v)$, which indicates user u's preference to event v. Hence, predicting $r(u, v)$ plays the central in event recommendation.

While, there exists geographical preference on the events by analysing data. Additionally, latent factor model is practical and flexible. Hence, we propose a mixture rating based latent factor model which can combine personal and geographical preference. In order to tackle with the implicit rating characteristic, a Bayesian personal ranking framework is introduced to learn the mixture rating. Heterogeneous social relations are applied into improve the performance by popular social regularization method.

Mixture Rating

In order to model the geographical preference and improve model performance, we use the *regional preference-aware geographical rating* to extend the *personal rating*,

$$r(u_i, v_j) = \alpha \dot{r}(u_i, v_j) + (1 - \alpha)\ddot{r}(u_i, v_j) \tag{3}$$

where the relative weight α is the fusion parameter that controls the contribution of the two parts.

Personal rating. The personal rating is similar to the matrix factorization latent factor model, which is proved efficient in recommendations. The basic idea is to embody user i and event j with the low-dimensional latent factor vectors $U_i \in R^k$ and $V_j \in R^k$. Then, the dyadic rating $\dot{r}(u_i, v_j)$ of user i to item j is approximated as follows,

$$\dot{r}(u_i, v_j) = U_i^T V_j. \tag{4}$$

Regional preference-aware geographical rating. In the last section, we have observed the reality of regional concentration for events. Hence, we consider the regional preference-aware geographical rating, $\ddot{r}(u_i, v_j)$. The regional preference-aware geographical rating is derived by integrating the weighted user rating *w.r.t.* each region. During the integrating, the weight value is the probability of event v_j belonging to a specific region. Consider a region space of k elements where all events are located, i.e., $D = d_1, d_2, ..., d_k$. Then, the rating can be derived by using the equation as follows,

$$\ddot{r}(u_i, v_j) = \sum_{t=1}^{k} \Pi_i M_t C_{jt} \tag{5}$$

where $\Pi_i \in R^l$ is a low-dimensional latent row vector associated with the user u_i, $M_t \in R^l$ is a low-dimensional latent column vector associated with the region d_t, and C_{jt} represents the probability of event v_j belonging to region d_t.

We first use the K-Means algorithm to cluster all events to obtain the k regions, where the geographical feature of each event is denoted as the binary variant (latitude, longitude). We use Gaussian distribution to model the relationship between the events and the regions. Thus, each region d_i has a parameter pair (μ_i, Σ_i) where μ_i is the expected value of events in the region d_i and Σ_i is the variance matrix of latitude and longitude. The parameter pairs can be obtained by computing the events located in d_i. Then, we can compute the probability of a event belonging to a region, given the region set D. Here, for each event v_i, we denote a vector C_i which is used to represent the probability of an event belonging to each region,

$$C_{ij} := \frac{N(L^i|\mu_j, \Sigma_j)}{\sum_{t=1}^{k} N(L^i|\mu_t, \Sigma_t)} \tag{6}$$

MODEL TRAINING

The feedback of an event is often implicit, with presence/absence represented by a binary value 1/0. Bayesian Personalized Ranking (BPR)(Rendle et al. 2009) emphases on predicting the dyadic rating $r(u, v)$ and top-ranking items with high scores, which can be used to solve the implicit feedback recommendation problem. Based on the BPR optimization criterion, we regard events involving user u as a positive event set (denoted by P_u^I), while the remaining events as the negative set (denoted by N_u^I). Then, we expect to maximize the *objective function* that ranks P_u^I higher than N_u^I, as in Eq. (7),

$$\max_{\Theta} \prod_{(u_i,v_j,v_k)\in(U,P_u^I,N_u^I)} P(r(u_i,v_j) > r(u_i,v_k)|\Theta) \tag{7}$$

where $\Theta = (U, V)$ is the parameter set in the model.

Heterogenous social regularization. It has been widely admitted that social relationship can improve the recommendation performance. In the paper, we use social regularization term, which is based on the assumption that the preference of a user is close to the weighted average preference of his friends. We also add a Gaussian prior in the model as follows,

$$\begin{aligned} &P(U_i - \sum_{j=1}^{n} \frac{w_{ij}}{\sum_{k=1}^{n} w_{ik}} U_j|W^{on}) \propto \\ &N(U_i - \sum_{j=1}^{n} \frac{w_{ij}}{\sum_{k=1}^{n} w_{ik}} U_j|\mathbf{0}, \sigma^2 I) \end{aligned} \tag{8}$$

In the problem, there are two different social relationships. Event-based social services usually publish events in some related social groups in order to attract attentions. Hence, there will be hidden relationships between groups and events. Generally, a social group may organize many events. Such a group/event pair can be defined by the confidence weight in heterogenous social relationships as follows,

$$f_{ij}^H := \frac{|\{G(u_i) \cap G(u_j)\} \oplus \{E(u_i) \cap E(u_j)\}|}{|\{G(u_i) \cap G(u_j)\} \otimes \{E(u_i) \cap E(u_j)\}|} \tag{9}$$

where $|\{G(u_i) \cap G(u_j)\} \oplus \{E(u_i) \cap E(u_j)\}|$ denotes the size of the shared groups/events between user i and j, and $\{G(u_i) \cap G(u_j)\} \otimes \{E(u_i) \cap E(u_j)\}|$ denotes the size of all possible group/event pairs. Combining heterogenous social relationships and auxiliary heterogenous confidence weights, we extend former single social regularization term into multiple social regularization terms as follows,

$$\begin{aligned} &g(U_i|W^{on}, W^{off}, F^H) \\ &\propto N(U_i - \sum_{j=1}^{n} \frac{f_{ij}^H(w_{ij}^{on} + w_{ij}^{off})}{\sum_{k=1}^{n}[f_{ik}^H(w_{ik}^{on} + w_{ik}^{off})]} U_j|\mathbf{0}, \sigma^2 I) \end{aligned} \tag{10}$$

Here, we denote $\Upsilon_{ij} = \frac{f_{ij}^H(w_{ij}^{on}+w_{ij}^{off})}{\sum_{k=1}^{n}[f_{ik}^H(w_{ik}^{on}+w_{ik}^{off})]}$ for short. Hence, $g(U_i|W^{on}, W^{off}, F^H, \sigma) \sim N(U_i - \sum_{j=1}^{n} \Upsilon_{ij} U_j : \mathbf{0}, \sigma^2 I)$.

Objective posterior probability. Let $\Phi = (\sigma, \sigma_u, \sigma_v)$ be the prior parameters. In order to avoid over-fitting in the learning process, we also enforce Gaussian priors on the latent factor vectors U_i, V_j, Π_i, M_t. Then, given the user/event feedback matrix R, online social relation W^{on}, offline social relationship W^{off} and heterogenous weight matrix F^H, we can obtain that the probability $P(\Theta|R, W^{on}, W^{off}, F^H, \Phi)$ is proportional to Eq. (10),

$$\begin{aligned} &\prod_{(u_i,v_j,v_k)\in(U,P_u^I,N_u^I)} P(r(u,v_j) > r(u,v_k)|\Theta) \\ &\cdot \prod_i g(U_i|W^{on}, W^{off}, F^H, \sigma) \cdot \prod_i P(U_i|\sigma_u) P(\Pi_i|\sigma_\pi) \\ &\cdot \prod_j P(V_j|\sigma_v) \cdot \prod_t P(M_t|\sigma_m) \end{aligned} \tag{11}$$

where $P(r(u,v_j) > r(u,v_k)|\Theta) := l(r(u,v_j) - r(u,v_k))$, and function l represents the logistic function. Then, we can get the target optimization function, commonly represented as a log-likelihood function.

Parameter Learning

The parameters can be learned by maximizing the above objective function by using the stochastic gradient descent (SGD) algorithm (Pan, Xiang, and Yang 2012). SGD has fast speed to convergence and high scalability to large-scale data sets. The main process of SGD is to randomly scan all training instances and iteratively update parameters.

We firstly use the function $I(v_j \in P_{u_i}^I, v_k \in N_{u_i}^I)$ ($I(i,j,k)$ for shorten) as an indicator function. The function equals to 1 if both $v_j \in P_{u_i}^I$ and $v_k \in N_{u_i}^I$ stand, otherwise 0. Then, based on the objective function in Eq. (11), we have the gradients as follows,

$$\begin{aligned} \frac{\partial F}{\partial U_i} = &- \sum_{v_j \in P_{u_i}^I} \sum_{v_k \in N_{u_i}^I} \alpha \frac{e^{r(u_i,v_k)-r(u_i,v_j)}}{1+e^{r(u_i,v_k)-r(u_i,v_j)}} I(i,j,k)(V_k - \\ &V_j) - \frac{1}{\sigma_u^2}(2U_i - \sum_{j=1}^{|U|} \Upsilon_{ij} U_j) + \frac{\Upsilon_{ti}}{\sigma_u^2} \sum_{t=1 \wedge t\neq i}^{|U|} (U_t - \sum_{j=1}^{n} \Upsilon_{tj} U_j) \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial F}{\partial V_j} = &\sum_{i=1}^{|U|} \sum_{k=1}^{|V|} \alpha\{[\frac{e^{r(u_i,v_k)-r(u_i,v_j)}}{1+e^{r(u_i,v_k)-r(u_i,v_j)}} \cdot U_i] I(i,j,k) \\ &-[\frac{e^{r(u_i,v_j)-r(u_i,v_k)}}{1+e^{r(u_i,v_j)-r(u_i,v_k)}} \cdot U_i] I(i,k,j)\} \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial F}{\partial \Pi_i} = &- \sum_{j=1}^{|V|} \sum_{k=1}^{|V|} (1-\alpha)\{I(i,j,k) \frac{e^{r(u_i,v_k)-r(u_i,v_j)}}{1+e^{r(u_i,v_k)-r(u_i,v_j)}} \\ &\cdot [\sum_{t=1}^{K} M_t(C_{kt} - C_{jt})]\} \end{aligned} \tag{14}$$

$$\begin{aligned} \frac{\partial F}{\partial M_t} = &- \sum_{i=1}^{|U|} \sum_{j=1}^{|V|} \sum_{k=1}^{|V|} (1-\alpha)\{I(i,j,k) \frac{e^{r(u_i,v_k)-r(u_i,v_j)}}{1+e^{r(u_i,v_k)-r(u_i,v_j)}} \\ &\cdot [\Pi_i(C_{kt} - C_{jt})]\} \end{aligned} \tag{15}$$

We thus have the update rules used in the SGD algorithm framework,

$$U_i = U_i + \gamma\frac{\partial F}{\partial U_i};\ \ V_j = V_j + \gamma\frac{\partial F}{\partial V_j};$$
$$\Pi_i = \Pi_i + \gamma\frac{\partial F}{\partial \Pi_i};\ \ M_t = M_t + \gamma\frac{\partial F}{\partial M_t} \tag{16}$$

where γ is the predefined step size.

Algorithm 1 gives the algorithm of parameter learning.

Algorithm 1: The algorithm of HeSig

Input: user-event feedback matrix R, online social relationship W^{on}, Offline social relationship W^{off}, confidence weight f^H, users' geographical data L^S, events' geographical data L^E, parameter ξ

Output: user's personal latent factor U_u, user's geographical latent factor Π_u, event's latent factor V_v, region's geographical latent factor M_r

```
Initialize U_u, V_v, Π_r and M_r with randomly generated vectors,
Cluster events based on L^E to obtain K regions by using K-means,
Compute the parameter pair (μ_i, σ_i) for each region,
Compute C_ij, for all event-region pairs,
Initialize parameter pQ = 0,
Compute AUC value Q for train data,
while Q − pQ > ξ
    pQ = Q,
    Calculate the gradients as in Eqs.(13∼16);
    Update the parameters as in Eq.(17).
    Compute AUC value Q for train data,
end while
return parameters
```

Experiments

We implement the proposed recommendation model and test on several real-world data sets to demonstrate its effectiveness. We later give experimental results with discussions.

Data

We target users in the same city. We select five cities as described in section `Preliminary`, as they are the largest cities in the USA and have more users and groups than other cities. For all the five data sets, we use 5-fold cross-validation for performance assessment, and we report the average results. In order to verify our proposed model, we transform the original test data sets into *Pair test sets* and List test sets.

The *pair test sets* consist of pair events (participated, not participated) of each user from the original test data. The test sets are mainly used for comparing on pair-wise data.

In our problem, the central task is to predict a personalized list of the events the user wants to participate, which can be seen as a ranking problem. In order to evaluate the effectiveness of event recommendation in a real scenario, we generate the *list test sets*. Then for each user in the test data, we regard the events participated by the user as the target events, and expect to give them higher ratings than other remaining events in the test data. Finally, we get the test sets for the five cities. We call them as list test sets because the test results are evaluated on the event ranking lists.

Evaluation Measures

To evaluate the event recommendation results, we adopt three standard evaluation metrics: AUC, $P@k$ (Precision at Position k), and MAP (Mean Average Precision).

AUC measures the overall results of classification. It is suitable for highly imbalanced data set, as in our case where the negative events take a high proportion. In this work, we use AUC in the pair test sets to measure the results.

$$AUC = \frac{\sum_{i=1}^{|U|}\sum_{v_i\in P_{u_i}^I}\sum_{v_j\in N_{u_i}^I} I(r(u_i,v_j) > r(u_i,v_k))}{\sum_{i=1}^{|U|}|P_{u_i}^I|\cdot|N_{u_i}^I|} \tag{17}$$

where $I(.)$ is also an indicator function that equals to 1 if $r(u_i, v_j) > r(u_i, v_k)$ is true, otherwise, 0.

P@k and MAP are mainly used in ranking problems. For each user u, the average precision (AP) is defined as follows,

$$AP_u = \frac{\sum_{k=1}^{m} P@k\cdot I(L^k(u)\in P_u^I)}{|P_u^I|} \tag{18}$$

where $I(.)$ is given in Eq. (17), m is the number of events, $L^k(u)$ denotes the kth event in the ranking event list $L(u)$ and $|P_u^I|$ represents the number of events joined by u in the test sets. Finally, we can obtain MAP by averaging AP_u for all users. In this paper, we use $P@k$ and MAP in the list test sets.

Parameter Setup

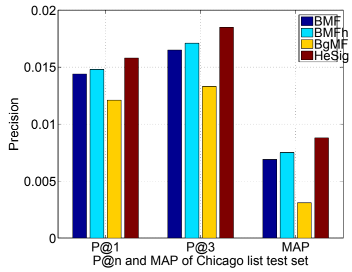
Learning rate and regularization parameters. Learning rate controls the speed of model training. However, it may not be able to converge if it is set too large. In this work, the learning rate is set to 0.001 for both matrix factorizations: geographical and personal. On the other hand, regularization parameters are also empirically set to 0.001 for all.

Relative weight α. α is the fusion coefficient of Eq. (3). We tune α by evaluating how the AUC changes in pair-test sets. As the results shown in Fig. 6(f), we get stable and better performance when $\alpha \in [0.94, 0.99]$. Taking the results of $P@k$ and MAP into consideration, we set α = 0.97

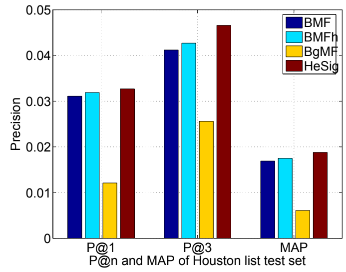
Dimension of latent factors and the number of reginal clusters. In addition to the above parameters, we also conducted sensitivity analysis in terms of the dimensionality of the latent factors. As we vary the number of dimensions, we found that it is not very sensitive. Empirically, we set the number of dimensions to be 10 for the latent factors in our model. Similar situation occurs in setting the number of reginal clusters. Empirically, we set the number to be 20 for each city's data.

Experimental Results

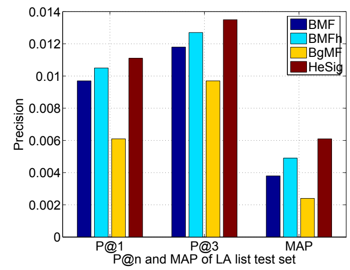
Results on Pair Test Sets. We first report the results on pair test sets. We use the AUC metric because it can reveal the overall results of all methods under the adopted pairwise learning framework, i.e., BPR. We compare our



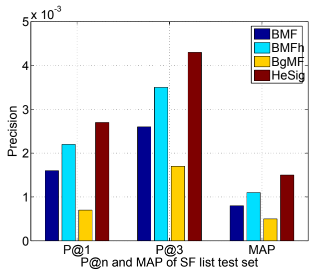
(a) Comparisons on the Chicago's list test sets



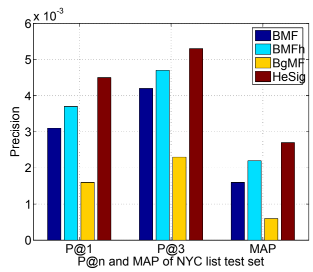
(b) Comparisons on the Houston's list test sets



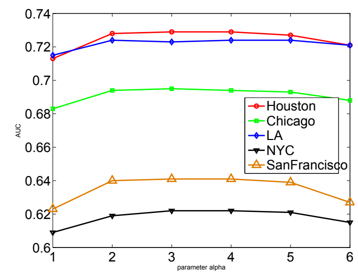
(c) Comparisons on the Los Angeles's list test sets



(d) Comparisons on the San Francisco's list test sets



(e) Comparisons on the New York's list test sets



(f) Test on the parameter α with pair data sets

Figure 6: The subfigures(a)-(e) are experimental results on different list test sets, and the subfigure(f) shows the AUC testing results on the parameter α.

method(**HeSig**) with Collaborative Filter based on Matrix Factorization(**MF**), Baysian MF for Implicit Rating(**BMF**), BMF with heterogenous social regularization(**BMFh**), and Baysian Regional preference-aware geographical MF for implicit rating(**BgMF**).

Table 2: AUC Results

Method	MF	BMF	BMFh	BgMF	HeSig
Houston	0.535	0.701	0.715	0.516	**0.729**
Chicago	0.456	0.686	0.691	0.433	**0.695**
LA	0.505	0.703	0.712	0.511	**0.724**
NYC	0.510	0.601	0.616	0.577	**0.622**
SF	0.534	0.621	0.633	0.589	**0.641**

Table 2 shows the experimental results under the standard evaluation metric **AUC**. From the results, we can observe that: 1)Models are better under considering imbalanced implicit rating characteristic; 2)The heterogenous social regularization can improve performance of all models; 3)BgMF can model users' regional preference, but its performance is not as good as BMF, that is due to what the regional preference just indirectly reflects a user's interests to the events; 4)The proposed model, $HeSig$, by combining geographically regional preference and heterogenous social information, achieves better performance than other methods.

Results on List Test Sets. We focus on the top-1 and top-3 recommendation results when all events in a city are considered. This is because, users tend to only focus on the top recommendation results while ignoring the rest. We also use MAP to measure the overall results of recommendation.

Since the methods considering implicit rating always perform better than the pure matrix factorization methods for our offline event recommendation problem, we compare our method **HeSig** with **BMF**, **BMFh** and **BgMF** *w.r.t.* $P@1$, $P@3$ and MAP.

Figure 6.(a)-(e) show that HeSig achieves the best results *w.r.t.* measures $P@1$, $P@3$ and MAP. For the top-1 and top-3 precision, it achieves significant improvements over the corresponding results in the five list test sets. It also achieves the best MAP among all methods. Thus, we can conclude that the proposed method can obtain better event recommendation results.

Conclusion

In this paper, we studied a new problem of event recommendation in event-based social networks (EBSNs). The technical challenge is to jointly model three sources of information, i.e., the *geographical features*, *heterogenous online+offline social relationships* and *user implicit rating data*. We presented a new Bayesian latent factor model to integrate these information for accurate event recommendation. We tested the model on real-world data sets, and the results demonstrated the performance of the method.

Acknowledgement

This work was supported by the NSFC (No. 61370025), 863 projects (No. 2012AA012502), 973 project (No. 2013CB329606), and the Strategic Leading Science and Technology Projects of Chinese Academy of Sciences(No.XDA06030200).

References

Borzsonyil, S.; Kossmann, D.; and Stocker, K. 2011. The skyline operator. In *17th International Conference on Data Engineering*, 421–430. Heidelberg.

Bruno, N.; Gravano, L.; and Marian, A. 2004. Evaluating top-k queries over web-accessible databases. In *ACM Transactions on Database Systems (TODS)*, 319–362. ACM.

Chaudhuri, S., and Gravano, L. 1999. Evaluating top-k selection queries. In *the 25th International Conference on Very Large Data Bases(VLDB99)*, 397–410. Morgan Kaufmann.

Cheng, C.; Yang, H.; King, I.; and Lyu, M. R. 2012. Fused matrix factorization with geographical and social influence in location-based social networks. In *the Twenty-Sixth AAAI Conference on Artificial Intelligence(AAAI12)*. AAAI Press.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization technology for recommendation system. In *Journal of Computer*, 30–37. IEEE Computer Society Press.

Liu, X.; Hey, Q.; Tiany, Y.; Lee, W.-C.; McPhersony, J.; and Han, J. 2012. Event-based social networks linking the online and offline social worlds. In *the 18th ACM SIGKDD international conference on Knowledge discovery and data mining(KDD12)*. ACM.

Liua, F., and Lee, H. J. 2010. Use of social network information enhance collaborative filtering performance. In *Expert Systems with Applications*, 4772C4778. ScienceDirect.

Lu, Y.; Tsaparas, P.; Ntoulas, A.; and Polanyi, L. 2010. Exploiting social context for review quality prediction. In *the 19th international conference on World wide web(WWW10)*, 691–700. ACM.

Ma, H.; Yang, H.; Lyu, M. R.; and King, I. 2008. Sorec social recommendation using probabilistic matrix factorization. In *the 17th ACM conference on Information and knowledge managemen(CIKM08)*, 931–940. ACM.

Ma, H.; Zhou, D.; Liu, C.; Lyu, M. R.; and King, I. 2011. Recommender systems with social regularization. In *the fourth ACM international conference on Web search and data mining(WSDM11)*, 287–296. ACM.

Pan, W.; Xiang, E.; and Yang, Q. 2012. Transfer learning in collaborative filtering with uncertain ratings. In *Twenty-Sixth AAAI Conference on Artificial Intelligence(AAAI2013)*.

Qiao, Z.; Zhang, P.; He, J.; Cao, Y.; Zhou, C.; and Guo, L. 2014. Combining geographical information of users and content of items for accurate rating prediction. In *23rd International World Wide Web Conference, WWW '14, Seoul, Republic of Korea, April 7-11, 2014*, 361–362. ACM.

Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Lars, S.-T. 2009. Bpr: Bayesian personalized ranking from implicit feedback. In *the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI2009)*. ACM.

Salakhutdinov, R., and Mnih, A. 2008. Probabilistic matrix factorization. In *Neural Information Processing Systems(NIPS08)*.

Somekh, O.; Aizenberg, N.; and Koren, Y. 2014. Build your own music recommender by modeling internet radio streams. In *Journal of World Wide Web*. Springer.

Takeuchi, Y., and Sugimoto, M. 2005. Cityvoyager: An outdoor recommendation system based on user location. In *ubiPCMM 2005*. IEEE.

Wang, C., and Blei, D. M. 2011. Collaborative topic modeling for for recommending scientific articles. In *the 17th ACM SIGKDD international conference on Knowledge discovery and data mining(KDD11)*, 448–456. ACM.

Xin, X.; King, I.; Deng, H.; and Lyu, M. R. 2009. A social recommendation framework based on multi-scale continuous conditional random fields. In *the 18th ACM conference on Information and knowledge management(CIKM09)*, 1247–1256. ACM.

Ye, M.; Yin, P. Y.; Lee, W.-C. L.; and Lee, D.-L. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR2011)*, 325–334. ACM.

Zhang, P.; Zhou, C.; Wang, P.; Gao, B.; Zhu, X.; and Guo, L. 2014. E-tree: An efficient indexing structure for ensemble models on data streams. In *IEEE Transactions on Knowledge and Data Engineering*, doi.ieeecomputersociety.org/10.1109/TKDE.2013.185.

Zheng, V. W.; Zheng, Y.; Xie, X.; and Yang, Q. 2010. Collaborative location and activity recommendations with gps history data. In *the 19th international conference on World wide web(WWW10)*, 1029–1038. ACM.