Joint Modeling of Participant Influence and Latent Topics for Recommendation in Event-based Social Networks

YI LIAO, The Chinese University of Hong Kong
WAI LAM, The Chinese University of Hong Kong
LIDONG BING, Tencent AI Lab
XIN SHEN, The Chinese University of Hong Kong

Event-based social networks (EBSNs) are becoming popular in recent years. Users can publish a planned event on an EBSN website, calling for other users to participate in the event. When a user is making a decision on whether to participate in an event in EBSNs, one aspect for consideration is existing participants defined as users who have agreed to join this event. Existing participants of the event may affect the decision of the user, to which we refer as participant influence. However, participant influence is not well studied by previous works. In this paper, we propose an event recommendation model that considers participant influence, and exploits the influence of existing participants on the decisions of new participants based on Poisson factorization. The effect of participant influence is associated with the target event, the host group of the event, and the location of the event. Furthermore, our proposed model can extract latent event topics from event text descriptions, and characterize events, groups, and locations by distributions of event topics. Associations between latent event topics and participant influence are exploited for improving event recommendation. Besides making event recommendation, the proposed model is able to reveal the semantic properties of the participant influence between two users semantically. We have conducted extensive experiments on some datasets extracted from a real-world EBSN. Our proposed model achieves superior event recommendation performance over several state-of-the-art models. The results demonstrate that the consideration of participant influence can improve event recommendation.

ACM Reference format:
Yi Liao, Wai Lam, Lidong Bing, and Xin Shen. 2017. Joint Modeling of Participant Influence and Latent Topics for Recommendation in Event-based Social Networks. 1, 1, Article 1 (January 2017), 31 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Event-based social networks (EBSNs), such as Meetup[1] and Facebook Events[2], are activity-oriented social networks that have drawn lots of research attention in recent years. EBSNs link people's online interactions to offline activities by providing a platform for users to gather online and participate planned events together. An event organizer can publish an offline planned event, also known as activity, such as hiking on a mountain, on the EBSN platform, calling for potential participants to join the proposed event. Typically, the published event is also associated with a

[1] https://www.meetup.com
[2] https://events.fb.com/

The work described in this paper is substantially supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region, China (Project Code: 14203414).

DOI: 10.1145/nnnnnnn.nnnnnnn

homepage on the EBSN site, showing the basic information about the event and a list of RSVPs for some users. The term RSVP stands for a French phrase "répondez, s'il vous plaît", which means "please reply". In an EBSN such as Meetup.com, a RSVP of a user for the event refers to the fact that the user has expressed his/her willingness to participate in the event. The basic information about the event includes the description of the event, the location, and so on. When users browse the information displayed on the homepage of the event, they can find out whether the event is interesting and send a RSVP to the event if the event appeals to them, indicating that they will attend this event. Due to the popularity of EBSNs where a large number of events are published every day, a user is not able to view all the events in detail. Consequently, an event recommendation system is essential to help a user discover events that he/she is probably interested in. Such kind of service also enhances user satisfaction.

Several models have exploited various information to solve the problem of event recommendation. Generally, the information utilized by these models can be categorized into two types. The first type is the context of the event, such as the text description of the event, the host of the event, and the location where the event will be held. Event context has been investigated by the majority of existing models [26, 43]. The context of the event forms the basic information of the event. The second type is the social relationship of the users since an EBSN is a typical online social network. Some EBSN sites allow users to build social relations with each other via group memberships or friendships, which can be exploited based on the observation that friends tend to have similar interest. Some models [22, 23, 31, 40, 43, 46] leverage social relations to enhance event recommendation performance.

However, most existing works neglect the influence of existing participants when making event recommendation. Indeed, potential participants are sometimes affected by the existing participants reflected from the RSVPs listed on the homepage of events. The RSVP information indicates who have already shown the willingness to participate in the events. Often, when a user browses an event, he/she not only checks whether the basic context information of the event, such as description and location, is interesting and convenient, but also considers whether the existing participants, who have decided to participate in the event, are good companions or not. This is a rather common phenomenon because an offline event may involve lots of interactions between participants. For example, when a user decides to participate a party, some particular existing participants may attract him/her, rather than just the party itself.

The notion of participant influence is different from various social relations among users which have been studied in previous works [26, 40, 43, 46]. Social relations are utilized by existing works as auxiliary information which conveys the message that users sharing the social relations tend to have similar interest. For example, Zhang and Wang [43] utilize friendships between users to infer the interests of each user more precisely based on the assumption that friends tend to have similar interest. Interestingly, they also report that the method making use of friendship only improves event recommendation performance slightly. In the framework proposed by Macedo et al. [26], users are linked via the groups to which they belong. Such information is utilized to facilitate recommendation assuming that users affiliated to the same or similar groups tend to the same event created by these groups. Xu et al. [40] introduce the concept of mutual influence to investigate which members of a group will participate in the event once the event is published. Their model assumes that users strengthen the likelihood of each other to participate in the event via the mutual influence which is modeled as a weighted link. However, the above models still rely on the basic information of the events for calculating the preference of the user, with social relations acting as auxiliary information. In contrast, the participant influence investigated in this paper is different from the explicit social relations between two users. It can be viewed as an implicit social relation between two users and such relations may directly affect the decision of a new participant. When a user influences another user to some extent in participating in various events, it can be treated as a kind of implicit social relation between them. However, such kind of social relation between two users is not necessarily reflected from the follower relationship.

The existing participants can be regarded as a kind of historical data of the events. Some recommendation models consider the influence of historical data when making recommendation in the future, such as next-basket recommender [4, 27, 33, 38] which recommends the next item given the history, usually a sequence of transaction data. However, event recommendation considering the participant influence is different from next-basket recommendation. When making a recommendation for a specific user, the history of the user encompasses item sequence in the task of next-basket recommendation. In contrast, in event recommendation considering participant influence, we consider the historical participants of the events, rather than the historical event sequences of users. Moreover, participant influence varies with different events since each event is associated with different existing participants.

We model an instance of participant influence by three components, namely, the user who exerts the influence, the user who is affected by the influence, and the target event. One characteristic is that different users exert different participant influence to others. For example, some users are influential in *Sport* event while some are influential in *Shopping* event. Likewise, the second characteristic is that different users are influenced in different ways. The third characteristic is that when the target event changes, the participant influence between two users should not stay the same. For example, the participant influence between a soccer expert and a soccer learner for a Soccer event will be definitely larger than that for a Basketball event. We employ the Poisson distribution to model the observed data with participant influence, resulting in a probabilistic Poisson factorization model. The observed data could be, for example, the number of times that one user influences another user for a particular event. The Poisson factorization model enjoys two advantages over traditional factorization models such as Gaussian factorization model [8]. First, it implicitly learns the budget of a user. For example, how much spare time the user has in total for spending on participating in various events. With such properties, a Poisson factorization model can capture the fact that a user does not participate in an event, or an unobserved record, in two ways; either the user dislikes the event or the user does not have sufficient spare time to participate in the event. Such consideration is realistic in real-world. Second, Poisson factorization provides different treatment for zeros, or unobserved records, from traditional factorization models. Poisson factorization automatically adapts the effect of unobserved records, as a user may not dislike the event as mentioned previously. In contrast, traditional factorization models impose equal treatment for zeros and non-zeros, which may be unrealistic in practical. To overcome such drawback, additional treatment such as confidence [13] is patched in traditional factorization models. As another benefit of its treatment for zeros, Poisson factorization can easily handle massive sparse data for event recommendation, where most user-event records are unobserved. As a consequence of the advantages, Poisson factorization has shown superior performance over traditional factorization models in various recommendation tasks [8].

Our proposed model also extracts latent event topics from event text descriptions with a latent Poisson topic model. The text description is written by the event organizer introducing the details of the event. The extracted latent event topics can be treated as a kind of semantic class of the events. The event topics are incorporated into participant influence modeling so as to model the event latent factor more effectively. Besides, by associating participant influence with event topics, the proposed model is able to recommend both existing events and new events. Moreover, the proposed model is also able to provide semantic explanations on the participant influence when recommending events to users.

We further exploit two types of context information of events, namely, host groups and locations, for improving recommendation performance. Each event is hosted by a group and is held at a location. The participant influence between two users for an event is affected by the host group and location because a user may exert different influence in different groups or different locations. For example, a soccer expert is more influential in a sport field held at a soccer court than in one on a basketball court. We model the participant influence between users regarding each group and each location in a similar manner with the participant influence regarding each event. We characterize events, groups, and locations by distributions of event topics. We also develop methods to learn the topics of groups and the topics of locations by analyzing the relationships with events.

In our recent work [20], we have investigated an event recommendation model taking into account of participant influence. This previous model can be regarded as a preliminary model for participant influence that only considers the interactions among two users and an event. The proposed recommendation framework in this paper is a more comprehensive and sophisticated model. Specifically, this new framework learns the latent topics from the event text descriptions and associate the topics with groups and locations. Moreover, it incorporates such latent topics into participant influence modeling to further improve the recommendation quality. By associating latent topics with participant influence, this new framework is able to reveal and interpret the semantic properties of the participant influence between two users.

2 RELATED WORK

2.1 Event Recommendation

In contrast with the ordinary item recommendation problem, event recommendation in EBSNs involves heterogeneous types of information. Event recommendation is related to: 1) Point of Interest (POI) recommendation [14, 16–19, 21, 48] in that each event is usually held at a place in reality with a geographical location. 2) Time-aware recommendation [6, 35, 41, 42] since each event is held at a certain time. 3) Content-aware recommendation [1, 28, 36, 39, 45] since each event usually contains text descriptions written by the organizer. 4) Social-aware recommendation [11, 25, 37, 44] since users in EBSNs are usually connected via social relations, such as group memberships and friendships. Several event recommendation models [23, 26, 30, 40, 43, 46] have been proposed to jointly analyze the flourish information. Liu et al. [23] employ data mining techniques to investigate properties of EBSNs and discover many unique characteristics, such as heavy-tailed degree distributions and strong locality of social interactions. Qiao et al. [30] extend their preliminary study and develop a Bayesian latent factor model that combines social relations and geographical information. They design a social regularization term on the factor representation of users based on the assumption that the preference of a user is close to the weighted average preference of his/her friends. Later, Macedo et al. [5] conduct a statistical study about the characteristics of events and RSVPs in EBSNs. They further design a learning-to-rank model for event recommendation [26]. The input features are derived from several aspects of the event and social relations. Recently, Zhang and Wang [43] propose a Poisson factorization model that collectively considers context information of the event and friendships of users. In their model, two users tend to have similar interest if they are friends. Rather than recommending events to individual users, Purushotham and Kuo [29] recently investigate recommending events to a group of users, i.e. group recommendation in EBSN. Xu et al. [40] study a slightly different event recommendation scenario from previous works, where the system tries to find out which members of a group will participate in an event when the event is published to the group. They introduce the concept of mutual influence, assuming users influence the likelihood of others to participate in the event within the group. They formulate the problem as a global decision-making process. In contrast with previous works, we study the participant influence in EBSNs, which directly affects users' decisions on events. The participant influence is further interacted with latent event topics to improve event recommendation performance.

2.2 Next-Basket Recommendation

Given a user's purchase history, next-basket recommendation predicts the next few items that the user most probably would like. Rendle et al. [33] propose Factorizing Personalized Markov Chains (FPMC) model to tackle the task. Recently, Wang et al. [38] propose a Hierarchical Representation Model (HRM) which explores sequential behavior and general taste of users. HRM subsumes FPMC. Recently, the idea of next-basket recommendation has been applied to many applications that have sequential patterns, such as successive POI recommendation [4, 10, 24, 47]. For example, Cheng et al. [4] add Localized Region Constraint to the FPMC model (FPMC-LR) to predict the next Point of Interest (POI) that the user will probably visit given the most recent visited POI. Liu et al. [24] propose a POI recommender system called WWO that exploits user check-in sequential patterns with temporal interval assessment, representing the temporal intervals of the soonest visit of a POI after each recent POI. The major difference between event recommendation considering participant influence and next-basket recommendation lies in what the type of history encompasses and its relation to the type of elements to be recommended. In next-basket recommendation, the history is composed of a sequence of items purchased by the user. However, in our proposed task, the history contains a sequence of RSVPs.

2.3 Poisson Factorization

Non-negative matrix factorization (NMF) [15] finds low dimensional representations/factors of users and items such that the square error function between the inner-products of latent representations/factors and the observed values is minimized. In contrast, Poisson factorization model is a generative version of NMF, where the observed values are generated from a Poisson distribution with parameters associated with the latent factors. Gaussian factorization [34] is another probabilistic derivative of the matrix factorization (MF), where the observed values are modeled with Gaussian distributions. Compared with Gaussian factorization or NMF, Poisson factorization enjoys the ability to handle massive sparse data [8].

Gopalan et al. [8] conduct several empirical studies to show that Poisson factorization model has superior recommendation performance over traditional matrix factorization models such as Gaussian factorization. This can be attributed to two advantages of Poisson factorization. First, Poisson factorization implicitly captures the limited budget of users to participate in events. We present a brief illustration as follows. Let $y_{u_i,e_j} \sim Poisson(u_i e_j)$ denote the number of times that the user u_i participate in the event e_j. Then $Y_{u_i} = \sum_{j \in |E|} y_{u_i,e_j}$ denotes the frequencies that the user u_i participates in events, which can be treated as the budget of the user u_i given his/her limited spare time. An interesting property of Poisson distribution is that a sum of Poisson variable is itself a Poisson variable with the rate equal to the sum of the rates, that is, $Y_{u_i} \sim Poisson(\sum_{j \in E} u_i e_j)$. Specifically, a Poisson variable is commonly known as naturally expressing the number of events occurring in a fixed interval of time or space. In other words, Y_{u_i} can naturally capture the budget of events in which the user u_i can participate given his/her limited spare time. As a consequence, unobserved data can be partially explained by Poisson factorization as that the user does not have sufficient time to participate in an event rather than that he/she dislikes the event. Second, Poisson factorization automatically adapts the effect of zero entries. Such property captures a more realistic consideration that no participation of an event by a user may not indicate dissatisfaction. An illustration is given as follows. The probability that y_{u_i,e_j} is generated is denoted as $p(y_{u_i,e_j}|u_i, e_j) =$

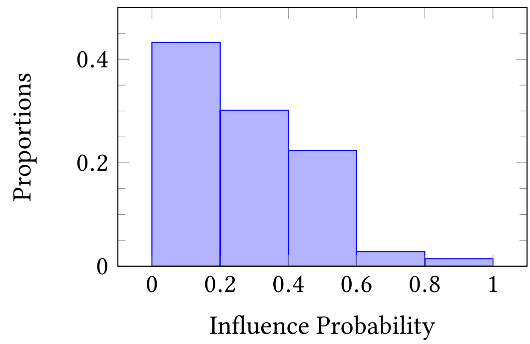

Fig. 1. Statistics on influence probability

$(u_i e_j)^{y_{u_i,e_j}} exp\{-u_i e_j\}/y_{u_i,e_j}!$. Recall the facts that $0! = 1$ and $(u_i e_j)^0 = 1$. When y_{u_i,e_j} equals to zero, $p(y_{u_i,e_j} = 0|u_i, e_j) = exp\{-u_i e_j\}$, which implies that the effect of y_{u_i,e_j} is automatically reduced.

Besides making recommendation, Poisson factorization can be employed as an alternative text topic modeling to Latent Dirichlet Allocation (LDA) [3] . Similarly to LDA, it can also capture latent topics of documents, which are represented by a distribution over words. Gapalan et al. [9] build a Poisson factorization model for capturing the topics of documents and make recommendations simultaneously.

3 BACKGROUND AND PROBLEM DEFINITION

We investigate the problem of event recommendation in EBSNs with the consideration of existing participants of events. We conduct a preliminary investigation in our real-world EBSN dataset corresponding to Los Angeles obtaining some insights on the influence of existing participants. We extract all the pairs of users $< u_a, u_b >$, where the user u_a and the user u_b ever participate at least one event together, and u_a indicates his/her willingness via RSVP to participate before u_b does. We examine the influence probability calculated as the probability that the user u_a influences the user u_b to participate in an event. Then the influence probability can be estimated as the proportions of the number of events that the user u_b participated after the user u_a did, among the number of events that u_b had participated. Figure 1 depicts the proportion of the user-pairs with influence probability in different ranges. It can be observed from the figure that 26.6% of the user-pairs have influence probability larger than 0.4, which implies that existing participants can be exploited to improve the performance of an event recommendation model. Note that we remove the cases that some users only participate one event. Moreover, our proposed model does not employ influence probability defined by this simple statistics. Instead, our model considers participant influence via a more advanced paradigm based on Poisson factorization.

We present a formal problem definition as follows. The problem involves the following elements: event set E, user set U, location set L, and group set G. Each event $e \in E$ is associated with the context information including the event location $l \in L$, the event group $g \in G$, and the textual description of the event. In many EBSNs, users sharing similar interest form a group, and events can be published under the group. Consequently, each event e is always associated with a group g. Each event e is also associated with a list of existing RSVPs, indicating who has expressed his/her willingness to participate in the event and the time of the RSVP. All the users can view the context information and current RSVPs of an event by which they are potentially affected when deciding whether to participate. We refer to such consideration as participant influence. Similar to the

Table 1. Description of notations

Notation	Description
$< u^a, u^p, e >$	A instance of participant influence, where the user u^p occurs after the user u^a in the event e
E, e_i	The set of events and the ith instance of event in E
L, l_j	The set of locations and the jth instance of location in L
G, g_m	The set of groups and the mth instance of group in G
W_e, w_{en}	The set of words in the text descriptions of event e and the nth instance of word in W_e
Z	Number of latent topics / Dimension of topic factor
K	Number of latent influence facets / Dimension of influence factor
$\theta^a_{u_1}$	The active influence factor of the user u_1
$\theta^p_{u_2}$	The passive influence factor of the user u_2
θ_e, θ_l, θ_g	The influence factors of the event e, the location l, and the group g respectively
η_e, τ_l, ψ_g	The topic distributions of the event e, the location l, and the group g respectively
η^b_e	The topic distribution of the background content for the event e
C_{ew}	The counts of the word w in the text descriptions of the event e
ϕ_z	The distribution over words for the topic z
$PI^{e_i}_{u_1,u_2}$	The strength of participant influence for the instance $< u_1, u_2, e_i >$
$PI^{l_j}_{u_1,u_2}$	The strength of participant influence for the instance $< u_1, u_2, l_j >$
$PI^{g_m}_{u_1,u_2}$	The strength of participant influence for the instance $< u_1, u_2, g_m >$
$D^e_{u_1,u_2}$	The number of times that the user u_1 participates the event e before the user u_2.
$D^g_{u_1,u_2}$	The number of times that the user u_1 participates the events held by the group g before the user u_2
$D^l_{u_1,u_2}$	The number of times that the user u_1 participates the events held at the location l before the user u_2
N_{ge}	The number of times that the event e is held by the group g
N_{le}	The number of times that the event e is held at the location l

common event recommendation task, the aim is to recommend events given a particular user taking into account of both the existing participants and the context information.

4 PARTICIPANT INFLUENCE

4.1 Basic Model

In our proposed model, we represent an instance of participant influence with three components as a tuple $< u^a, u^p, e >$[3]. This tuple denotes the observation that the user u^a influences the user u^p's decision to participate in the event e. In other words, the user u^p is aware of the fact that u^a is already in the participant list of this event e when the user u^p decides to participate in the event e. To model the participant influence, it is not effective to employ a fully parameterized model by which different specific parameters are designed for different tuples, i.e. different combinations of the user u^a, the user u^p and the event e. Such fully parameterized model suffers from the data sparsity issue in practice. Moreover, it cannot be generalized to infer the participant influence between two users who have never appeared simultaneously in an observed historical set of tuples.

[3]The superscript "a" denotes "active" and "p" denotes "passive".

Hence we design a latent factor model to capture the participant influence regarding a specific tuple. Latent factor models [13] have been developed to tackle various recommendation problems.

We design a latent factor for u^a, u^p, and e, denoted as θ_u^a, θ_u^p and θ_e respectively. Note that each user is associated with two latent factors θ_u^a and θ_u^p. θ_u^a captures how the user u influences others while θ_u^p captures how he/she is influenced by others. The intuition is that how much users influence others and get influenced by others are not exactly the same. Each latent factor is a K-dimensional vector. We decompose the full participant influence into K latent influence facets, each corresponding to one dimension. As a consequence, each dimension in the latent factor θ_u^a captures the strength that the user u influences others for the corresponding facet, denoted as latent active influence $\theta_{u,1:K}^a$. Similarly, each dimension in the latent factor θ_u^p captures the likelihood that the user u is influenced by others, denoted as latent passive influence $\theta_{u,1:K}^p$. Besides the factors associated with each user, we also design a latent factor for each event, denoted as a K-dimensional vector θ_e. Each dimension in θ_e captures the degree that participant influence between two users will take effect regarding this event e for this particular facet. For example, a soccer expert is more likely to influence a soccer learner on a *Soccer* event, than other types of events such as *Basketball* events.

Let PI_{u^a,u^p}^e denote the strength of participant influence regarding the tuple $< u^a, u^p, e >$. Hence PI_{U^a,U^p}^E can be treated as a 3-way tensor. Each entry is jointly decided by three latent factors $< \theta_u^a, \theta_u^p, \theta_e >$. PI_{u^a,u^p}^e captures how likely the participation of the user u^a will result in the participation of the user u^p when the target event is e. With tensor factorization technique, the entry PI_{u_1,u_2}^e in this tensor can be calculated as $PI_{u_1,u_2}^e = \theta_{u_1}^a \otimes \theta_{u_2}^p \otimes \theta_e$. Interestingly, since the tensor in our EBSN datasets is sparse, we can simplify the tensor factorization as pair-wise interactions between these involved factors [33] as shown in Equation 1.

$$PI_{u_1,u_2}^e = \theta_{u_1}^a \theta_{u_2}^p + \theta_e \theta_{u_1}^a + \theta_e \theta_{u_2}^p \tag{1}$$

In this formulation, the direct interaction between the user u_2 and the event e is also modeled as $\theta_e \theta_{u_2}^p$ without θ_{u_1}. It implies that the proposed participant influence modeling can be regarded as a more general model than typical existing methods. Typical methods can be regarded as a special case that only considers the event e and the target user u_2.

4.2 Incorporating Event Topics

Topic modeling algorithms [2] automatically discover a set of “topics" from a large collection of documents. Each topic is represented as a distribution over words that describe the corresponding topic. Topic models discover interpretable topics/themes for each document, which can be utilized for tasks such as document classification, document labeling, information retrieval, and so on. In our proposed model, we discover latent event topics from the text descriptions of events so as to provide a semantically interpretable latent structure for inferring participant influence between two users. Besides, by leveraging event topics, our proposed model is able to recommend both existing or newly created events. The topics can be further utilized to make explainable recommendations.

To incorporate event topics into our model, we augment the latent factors denoted in Equation 1 by encoding topic information into the event latent factors. Let η_e denote the topic distribution of the event e, which is represented by a weighted distribution over latent topics. For the event e, we concatenate the topic distribution η_e and the original event latent factor θ_e to obtain a new event factor $\tilde{\theta}_e$, denoted as a $(Z+K)$-dimensional vector in Equation 2.

$$\tilde{\theta}_e = (\underbrace{\eta_{e1}, \eta_{e2}, ..., \eta_{ez}, ..., \eta_{eZ}}_{topic\ distribution}, \underbrace{\theta_{e1}, \theta_{e2}, ..., \theta_{ek}, ..., \theta_{eK}}_{influence\ facets}) \tag{2}$$

where Z is the number of latent event topics and K is the number of latent influence facets. η_e is embedded into the first Z dimensions and the original event latent factor θ_e is embedded in the remaining K dimensions. We describe how the event topic distribution η_e is learned in Section 4.4. Accordingly, we augment the active participant influence factor θ_u^a and passive participant influence factor θ_u^p to (Z+K)-dimensional latent factor as well. The first Z dimensions of θ_u^a and θ_u^p capture the association between topics and participant influence. Specifically, the z-th dimension of θ_u^a captures the degree that the user u will influence others to participate in an event related to the z-th topic. Similarly, the entries in θ_u^p capture the degree that the user is influenced by others for the corresponding topics. Then Equation 1 is refined to Equation 3.

$$\begin{aligned} PI_{u_1,u_2}^e &= \theta_{u_1}^a \theta_{u_2}^p + \tilde{\theta}_e \theta_{u_1}^a + \tilde{\theta}_e \theta_{u_2}^p \\ &= \sum_{d=1}^{Z+K} (\theta_{u_1,d}^a \theta_{u_2,d}^p + \tilde{\theta}_{e,d} \theta_{u_1,d}^a + \tilde{\theta}_{e,d} \theta_{u_2,d}^p) \end{aligned} \tag{3}$$

We can also rewrite the result of Equation 3 into another form as shown in Equation 4. The first term in the right hand side is the strength of participant influence regarding event topic distribution, while the second term corresponds to event influence facets. Hence concatenating two factors has a similar effect as treating them as two separate factors with learning a joint representation in factorization models.

$$\begin{aligned} &\sum_{d=1}^{Z+K} (\theta_{u_1,d}^a \theta_{u_2,d}^p + \tilde{\theta}_{e,d} \theta_{u_1,d}^a + \tilde{\theta}_{e,d} \theta_{u_2,d}^p) \\ =& \sum_{d=1}^{Z} (\theta_{u_1,d}^a \theta_{u_2,d}^p + \eta_{e,d} \theta_{u_1,d}^a + \eta_{e,d} \theta_{u_2,d}^p) + \sum_{d=1}^{K} (\theta_{u_1,Z+d}^a \theta_{u_2,Z+d}^p + \theta_{e,d} \theta_{u_1,Z+d}^a + \theta_{e,d} \theta_{u_2,Z+d}^p) \end{aligned} \tag{4}$$

4.3 Modeling Observations with Participant Influence

Let D_{u_1,u_2}^e be the number of times that the user u_1 participates the event e before the user u_2. In our modeling framework, the observed data D_{u_1,u_2}^e can be viewed as the result of generating data from the participant influence PI_{u_1,u_2}^e using Equation 3. Intuitively, stronger participant influence from the user u_1 to the user u_2 regarding the event e will generate larger D_{u_1,u_2}^e.

We employ Poisson distribution to model the generation of the observed data associated with the corresponding participant influence as shown in Equation 5.

$$\mathcal{P}(D_{u_1,u_2}^e; PI_{u_1,u_2}^e) = (PI_{u_1,u_2}^e)^{D_{u_1,u_2}^e} \exp(-PI_{u_1,u_2}^e)/(D_{u_1,u_2}^e)! \tag{5}$$

where $\mathcal{P}$ denotes Poisson distribution and PI_{u_1,u_2}^e refers to the participant influence defined in Equation 3. Participant influence acts as the shape parameters of the Poisson distribution governing the generation of the observed data D_{u_1,u_2}^e. To smooth the latent factors $\theta_{u_1}^a$, $\theta_{u_2}^p$, and θ_e, we apply a Gamma prior with parameters $\lambda_{a\cdot}$ and $\lambda_{b\cdot}$ to each latent factor described above to avoid over-fitting, which is also conjugate to Poisson distribution. The proposed model can be viewed as a Probabilistic non-negative matrix factorization model [34] with Poisson distribution, i.e., Poisson factorization model (PFM). As discussed in [8], PFM enjoys several advantages over the well-known Gaussian factorization model (GFM). Section 1 has presented the advantages.

4.4 Modeling Event Topics

Let V denote a predefined vocabulary set that is composed of the words in the event descriptions after some preprocessing. The size of the vocabulary set V is denoted as $|V|$. Let ϕ_z denote the distribution over the words in the set V for the topic z. The count of each word in the event distribution is generated by both the word distributions ϕ_z of the topics and the topic distribution η_e of the events. However, the text descriptions of the events may contain many background words that is unrelated to the topics of events. For example, the event organizer may write some content encouraging users to participate in the event. To better capture the topics of events while reducing the effect of background words, we treat the text descriptions of events as a combination of event topical content and background content. Note that the background words are different from stopwords which are usually removed during document preprocessing. For example, the word "invite" in the sentence "we are pleased to invite you for a special beach party" commonly appears in many event descriptions though this word seldom conveys any information about the topic of the corresponding event. Such kind of words are captured as background content. By reducing effect of background words, the topics of events can be captured more accurately.

Specifically, we design a variable η_e^b to capture the topic distribution for the background content in event descriptions on the EBSN. The sum of these two factors, represented by $\eta_e + \eta_e^b$, captures the topic distribution of the whole content in the event descriptions. Then the count C_{ew} for a word w in the description of the event e is generated based on Equation 6.

$$C_{ew} \sim \mathcal{P}(\phi_{1:Z,w}^T(\eta_e + \eta_e^b))) \tag{6}$$

4.5 Group Modeling for Participant Influence

As a feature of typical EBSNs, the group information provides additional clues to recommend events since each event is associated with a host group. Hence our proposed framework also models the group to which the event belongs to make better recommendations. We intend to capture the participant influence between users regarding groups using the same form as in Equation 3. Similar to event modeling as described in Sections 4.1 and 4.2, we design a group latent factor θ_g for each group g. Then the topics of group ψ_g is distilled to generate a new group factor $\tilde{\theta}_g$ using the same formulation as in Equation 2. We define the calculation of participant influence between the user u_1 and the user u_2 regarding the group g in Equation 7.

$$PI_{u_1,u_2}^g = \theta_{u_1}^g \theta_{u_2}^p + \tilde{\theta}_g \theta_{u_1}^a + \tilde{\theta}_g \theta_{u_2}^p \tag{7}$$

where $\theta_{u_1}^a$ refers to the active influence factor and $\theta_{u_2}^p$ refers to the passive influence factor that are described previously. We describe how the proposed model learns the topics of the group ψ_g below.

We characterize the topics of groups ψ_g as a distribution over latent event topics by analyzing the relations between groups and events. To fit ψ_g into the calculation of participant influence denoted in Equation 7, we define ψ_g as a Z-dimensional vector. Moreover, we employ collaborative Poisson matrix factorization to guarantee that the topics of the group ψ_g is in the same latent semantic space with the topic distribution η_e of events. Specifically, we fit the number of times that the event e is held by the group g using Equation 8.

$$N_{ge} \sim \mathcal{P}(\psi_g^T \eta_e) \tag{8}$$

where $\mathcal{P}$ is Poisson distribution with the shape parameter as the inner-product of ψ_g and η_e.

We employ Poisson distribution to generate the observed data for groups using the same form of Equation 5. Accordingly, Equation 9 models the generation of the observations that the total number of events hosted in this group, where the user u_1 participated before the user u_2 did.

$$\mathcal{P}(D^{g}_{u_1,u_2}; PI^{g}_{u_1,u_2}) = (PI^{g}_{u_1,u_2})^{D^{g}_{u_1,u_2}} \exp(-PI^{g}_{u_1,u_2})/(D^{g}_{u_1,u_2})! \quad (9)$$

4.6 Geographic Location Modeling for Participant Influence

In addition to the host group, we also exploit the locations of events for enhancing the modeling of participant influence. Similar to group modeling for participant influence, we design the formula in Equation 10 to capture the participant influence between users regarding a particular location l. The corresponding observations generated with Poisson distribution are the total number of events hosted in this location l where the user u_1 indicated his/her willingness to participate before the user u_2 did. The observations are modeled using Equation 11

$$PI^{l}_{u_1,u_2} = \theta^{l}_{u_1}\theta^{l}_{u_2} + \tilde{\theta}_l\theta^{a}_{u_1} + \tilde{\theta}_l\theta^{p}_{u_2} \quad (10)$$

$$\mathcal{P}(D^{l}_{u_1,u_2}; PI^{l}_{u_1,u_2}) = (PI^{l}_{u_1,u_2})^{D^{l}_{u_1,u_2}} \exp(-PI^{l}_{u_1,u_2})/(D^{l}_{u_1,u_2})! \quad (11)$$

To capture the topics τ_l of the location l, we model the number of times that the location l hosts the event e, denoted by N_{le}, using Equation 12.

$$N_{el} \sim \mathcal{P}(\tau_l^T \eta_e) \quad (12)$$

4.7 Generative Process

We depict the graphical model of the complete framework in Figure 2, with the generative process described as follows. The notations of the symbols in our proposed model are listed in Table 1. Steps (1), (2), (3), (6) and (7) refer to participant influence modeling and event topic modeling described in Section 4.1-4.4. Steps (4), (6) and (8) refer to the group modeling described in Section 4.5. Steps (5), (6) and (9) refer to the geographic location modeling in Section 4.6.

(1) For each user u,
   (a) draw latent active influence factor
   $\theta^{a}_{ud} \sim Gamma(\lambda^{a}_{ua}, \lambda^{a}_{ub})$ for the d-th dimension of θ^{a}_{u}
   (b) draw latent passive influence factor
   $\theta^{p}_{ud} \sim Gamma(\lambda^{p}_{ua}, \lambda^{p}_{ub})$ for the d-th dimension of θ^{p}_{u}
(2) For each word w, draw topic intensity $\phi_{wz} \sim Gamma(\lambda_{ta}, \lambda_{tb})$
(3) For each event e,
   (a) draw influence latent factor $\theta_{ek} \sim Gamma(\lambda_{ea}, \lambda_{eb})$
   (b) draw topic distribution $\eta_{ez} \sim Gamma(\lambda^{t}_{ea}, \lambda^{t}_{eb})$
   (c) Draw topic distribution of background content $\eta^{b}_{ez} \sim Gamma(\lambda^{b}_{ea}, \lambda^{b}_{eb})$
   (d) For each word w in the event description,
      (i) draw $C_{ew} \sim \mathcal{P}(\phi^T(\eta_e + \eta^{b}_{e}))$
(4) For each group g,
   (a) draw influence latent factor $\theta_{gk} \sim Gamma(\lambda_{ga}, \lambda_{gb})$
   (b) draw $\psi_{gz} \sim Gamma(\lambda_{ga}, \lambda_{gb})$
   (c) draw $N_{ge} \sim \mathcal{P}(\psi_g, \eta_e)$
(5) For each location l,
   (a) draw influence latent factor $\theta_{lk} \sim Gamma(\lambda_{la}, \lambda_{lb})$
   (b) draw $\tau_{lz} \sim Gamma(\lambda_{la}, \lambda_{lb})$
   (c) draw $N_{le} \sim \mathcal{P}(\tau_l, \eta_e)$
(6) Manipulate the factors $\tilde{\theta}_e$, $\tilde{\theta}_g$, and $\tilde{\theta}_l$ using the method in Equation 2
(7) For each identical tuple $< u_1, u_2, e >$

(a) calculate the associated participant influence $PI^{e}_{u_1,u_2}$ using Equation 3
(b) draw the count $D^{e}_{u_1,u_2} \sim \mathcal{P}(D^{e}_{u_1,u_2}; PI^{e}_{u_1,u_2})$, denoted in Equation 5

(8) For each identical tuple $< u_1, u_2, g >$
(a) calculate the associated participant influence $PI^{g}_{u_1,u_2}$ using Equation 3
(b) draw the count $D^{g}_{u_1,u_2} \sim \mathcal{P}(D^{g}_{u_1,u_2}; PI^{g}_{u_1,u_2})$, denoted in Equation 9

(9) For each identical tuple $< u_1, u_2, l >$
(a) calculate the associated participant influence $PI^{l}_{u_1,u_2}$ using Equation 3
(b) draw the count $D^{l}_{u_1,u_2} \sim \mathcal{P}(D^{l}_{u_1,u_2}; PI^{l}_{u_1,u_2})$, denoted in Equation 11

5 PARAMETER INFERENCE

5.1 Inference Overview

The goal of the inference procedure is to infer the parameters $\Theta = \{\theta^{a}_{u_1}, \theta^{p}_{u_2}, \theta_e, \theta_l, \theta_g, \eta_e, \eta^b, \psi_g, \tau_l\}$, given the observations $O = \{D^{e}_{u_1,u_2}, D^{g}_{u_1,u_2}, D^{l}_{u_1,u_2}, N_{ge}, N_{le}\}$. We suppress the hyperparameters in the Gamma distributions, namely, λ_a and λ_b, for simplicity. The posterior distribution of latent factors can be expressed as:

$$p(\Theta|O) = \frac{p(O|\Theta)p(\Theta)}{p(O)} \tag{13}$$

The denominator is the marginal probability of all the observed data defined as:

$$\begin{aligned} P(O) = &\prod_{k=1}^{|E|}\prod^{|U|^2}\iiint p(D^{e}_{u_1,u_2}; PI^{e}_{u_1,u_2})p(\Theta)\prod_{m=1}^{|L|}\prod^{|U|^2}\iiint p(D^{g}_{u_1,u_2}; PI^{g}_{u_1,u_2})p(\Theta) \\ &\prod_{k=1}^{|G|}\prod^{|U|^2}\iiint p(D^{l}_{u_1,u_2}; PI^{l}_{u_1,u_2})p(\Theta)\prod_{n=1}^{|W|}\prod_{e\in E}^{|E|}\iint p(C_{wn}; (\eta_e+\eta^b)^T\phi)p(\Theta) \\ &\prod^{|G|}\prod^{|E|}\iint p(N_{ge}; \eta^T\tau_l)p(\Theta)\prod^{|L|}\prod^{|E|}\iint p(N_{le}; \eta^T\tau_l)p(\Theta) \end{aligned} \tag{14}$$

Unfortunately, the probability expressed in Equation 14 is intractable due to the coupling of multiple factors in the integration. Inspired by [43] and [9], we develop an inference method based on variational inference algorithm [12]. The general idea of variational inference is to find a distribution to approximate the intractable posterior distribution. Specifically, a variational distribution $q(\Theta)$ is learned such that the KL divergence to the posterior distribution expressed in Equation 20 is minimized.

5.2 Auxiliary Variables

We add auxiliary variables to facilitate the inference. Recall that K denotes the dimension of the latent influence factors and Z denotes the number of the latent event topics. The auxiliary variables are designed for the inner-product of any pair of factors. For example, for the pair of factors $\theta^{a}_{u_1,1:Z}$ and η_e for the user u_1 and the event e respectively, we design Z latent auxiliary variables $\xi^{a}_{u_1e,z} \sim Poisson(\theta^{a}_{u_1,z}\eta_{e,z})$. We list all the auxiliary variables and the corresponding factor pairs in Table 2.

A set of co-related auxiliary variables should satisfy a constraint. There is a constraint for each Poisson component in the original model. For example, the constraint for the Poisson component depicted in Equation 5 is designed as:

$$D^{e}_{u_1,u_2} = \sum_z(\xi^{a}_{u_1e,z} + \xi^{p}_{u_2e,z} + \chi^{1}_{u_1u_2,z}) + \sum_k(\alpha^{a}_{u_1e,k} + \alpha^{p}_{u_2e,k} + \chi^{2}_{u_1u_2,k}) \tag{15}$$

The constraint for the $N_{ge} = \sum_z \epsilon_{ge,z}$ is designed for the Poisson component in Equation 8.

Table 2. Auxiliary variables and corresponding factor pairs

auxiliary variables	α^a	α^p	ξ^a	ξ^p	β^a	β^p	σ^a	σ^p	χ^1
factor pairs	$\theta_u^a\theta_e$	$\theta_u^p\theta_e$	$\theta_u^a\eta_e$	$\theta_u^p\eta_e$	$\theta_u^a\theta_g$	$\theta_u^a\theta_g$	$\theta_u^a\psi_g$	$\theta_u^p\psi_g$	$\theta_{u,1:Z}^a\theta_{u,1:Z}^p$
auxiliary variables	γ^a	γ^p	π^a	π^p	ρ	ϱ	ζ	ϵ	χ^2
factor pairs	$\theta_u^a\theta_l$	$\theta_u^p\theta_l$	$\theta_u^a\tau_l$	$\theta_u^p\tau_l$	$\eta_e\phi_w$	$\eta_e^b\phi_w$	$\eta_e\tau_l$	$\eta_e\psi_g$	$\theta_{u,1:K}^a\theta_{u,1:K}^p$

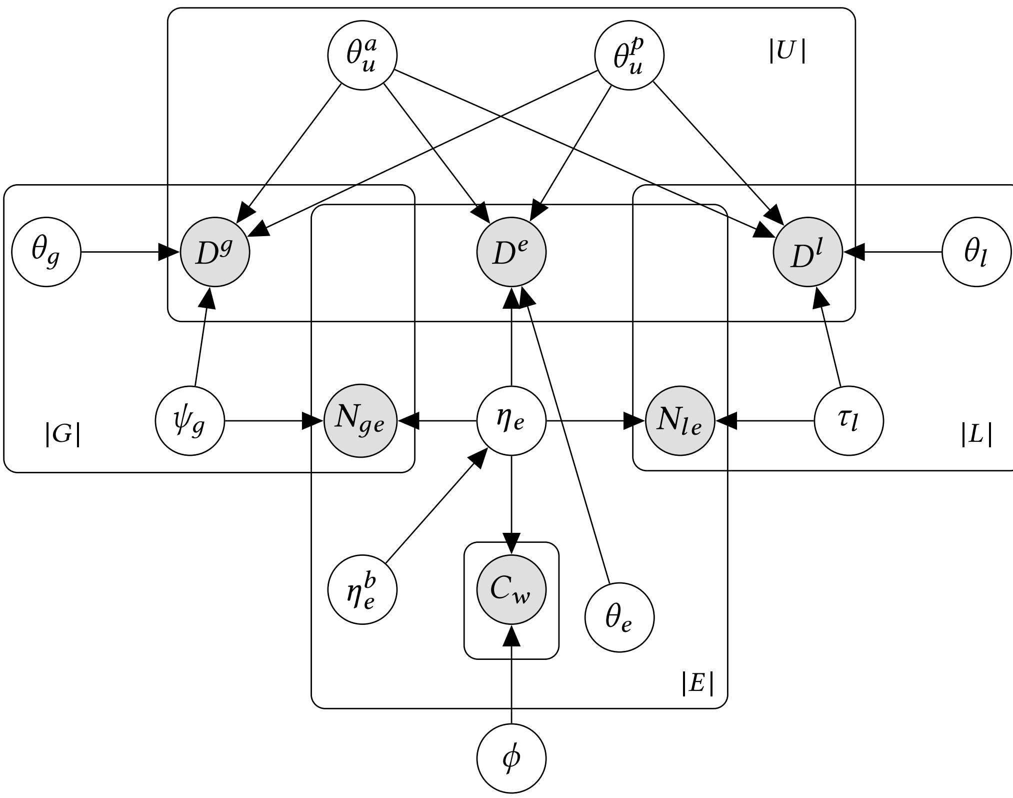

Fig. 2. Graphical model for our Framework

5.3 Conditional Distributions

We derive the complete conditional distribution for each latent variable, including both original variables and auxiliary variables. First, we derive the conditional distribution for original latent variables. For example, let us consider the location factor. The complete conditional distribution for the k-th dimension of the factor $\theta_{e,k}$, given the other latent variables is formulated as:

$$\begin{aligned}
&p(\theta_{e,k}|\Theta_{\neg\theta_{e,k}}, O) \\
&\propto \theta_{e,k}^{\lambda_{ea}-1}\exp(-\lambda_{eb}\theta_{e,k})\prod_{u_1,u_2}(\theta_{e,k}\theta^a_{u_1,k})^{\alpha^a_{u_2e,k}}\exp(-\theta_{e,k}\theta^a_{u_1,k})\prod_{u_1,u_2}(\theta_{e,k}\theta^p_{u_2,k})^{\alpha^p_{u_2e,k}}\prod\exp(-\theta_{e,k}\theta^p_{u_2,k}) \\
&\propto \theta_{e,k}^{\lambda_{ea}-1+\sum_{u_1}\sum_{u_2}(\alpha^a_{ae,k}+\alpha^p_{ue,k})}\exp(-(\lambda_{eb}+\sum_{u_1}\sum_{u_2}(\theta^a_{u_1,k}+\theta^p_{u_2,k}))\theta_{e_i}) \\
&= Gamma(\lambda_{ea}+\sum_{u_1}\sum_{u_2}(\alpha^a_{ae,k}+\alpha^p_{u_2e,k}), \lambda_{eb}+\sum_{u_1}\sum_{u_2}(\theta^a_{u_1,k}+\theta^p_{u_2,k}))
\end{aligned} \tag{16}$$

It can be observed from Equation 16 that a sum of independent Poisson random variables is itself a Poisson with the rate equal to the sum of the rates. Hence the new latent variables preserve the marginal distribution of the observed data O.

Then we derive the conditional distribution for auxiliary variables. For example, the variables involved in Equation 3 include topic factors $\{\theta^a_{u_1,1:Z}, \theta^p_{u_2,1:Z}, \eta_e\}$ and influence factors $\{\theta^a_{u_1,1:K}, \theta^p_{u_2,1:K}, \theta_e\}$. Let

$$y^e_{u_1,u_2} = (\alpha^a_{u_1e,1:Z}, \alpha^p_{u_2e,1:Z}, \xi^a_{u_1e,1:K}, \xi^p_{u_2e,1:K}, \chi_{u_1u_2,1:Z+K}) \tag{17}$$

where $y^e_{u_1,u_2}$ denotes a variable representing the concatenation of the auxiliary variables associated with the variable $PI^e_{u_1,u_2}$. The conditional distribution of $y^e_{u_1,u_2}$ can be formulated as in Equation 18.

$$p(y|O,\Theta) = \frac{p(y, Y = D^e_{u_1,u_2})}{p(Y = D^e_{u_1,u_2})} = \frac{p(y)p(D^e_{u_1,u_2}|y)}{p(D^e_{u_1,u_2})} \tag{18}$$

Recall that the sum of Poisson distributions preserves the marginal distributions, from which we can obtain $p(D^e_{u_1,u_2}|y) = 1$ given that $D^e_{u_1,u_2} = \sum_i y_i$. Furthermore, $p(y)$ and $p(D^e_{u_1,u_2})$ are computed by Equations 19 and 20 respectively.

$$\begin{aligned}
p(y) =& p(\alpha^a_{u_1e,1:Z}, \alpha^p_{u_2e,1:Z}, \xi^a_{u_1e,1:K}, \xi^p_{u_2e,1:K}, \chi_{u_1u_2,1:Z+K}) \\
&= \exp(-D^e_{u_1,u_2})\prod_z(\eta^a_{u_1e,z}\theta^a_{u_1e,z})^{\xi^a_{u_1e,z}}(\eta^p_{u_2e,z}\theta^p_{u_2e,z})^{\xi^p_{u_2e,z}}(\theta^a_{u_1u_2,z}, \theta^p_{u_1u_2,z})^{\chi_{u_1u_2,z}} \\
&\prod_k(\theta_{e,k}\theta^a_{u_1e,k})^{\alpha^a_{ue,k}}(\theta_{e,k}\theta^p_{u_2e,k})^{\alpha^p_{ue,k}}(\theta^a_{u_1u_2,k}, \theta^p_{u_1u_2,k})^{\chi_{u_1u_2,k}} \\
&/\prod_z(\xi^a_{u_1e,z}!)(\xi^p_{u_2e,z}!)(\chi_{u_1u_2,z}!)\prod_k(\alpha^a_{ue,k}!)(\alpha^p_{ue,k}!)(\chi_{u_1u_2,k}!)
\end{aligned} \tag{19}$$

$$\begin{aligned}
p(D^e_{u_1u_2}) =& exp(-D^e_{u_1u_2})[\sum_z(\eta^a_{u_1e,z}\theta^a_{u_1e,z} + \eta^p_{u_2e,z}\theta^p_{u_2e,z} + \theta^a_{u_1u_2,z}, \theta^p_{u_1u_2,z}) \\
&+ \sum_k(\theta_{e,k}\theta^a_{u_1e,k} + \theta_{e,k}\theta^p_{u_2e,k} + \theta^a_{u_1u_2,k}, \theta^p_{u_1u_2,k})]^{D^e_{u_1u_2}}/D^e_{u_1u_2}!
\end{aligned} \tag{20}$$

With the above equations, we can easily observe that the conditional probability of the auxiliary variables $p(y|O,\Theta) = p(y)/p(D^e_{u_1u_2})$ is a Multinomial distribution. The parameter of the Multinomial distribution for each auxiliary variable is the product of two associated original variables. The Multinomia probability can be interpreted as:

$$\begin{aligned} & p(\alpha^a_{u_1 e,1:K}, \alpha^p_{u_2 e,1:K}, \xi^a_{u_1,1:Z}, \xi^p_{u_2,1:Z}, \chi_{u_1 u_2,1:Z+K} | O, \Theta) \\ =& Mult(\alpha^a_{u_1 e,1:K}, \alpha^p_{u_2 e,1:K}, \xi^a_{u_1,1:Z}, \xi^p_{u_2,1:Z}, \chi_{u_1 u_2,1:Z+K}; \theta^a_{u_1,1:K}\theta_e, \theta^p_{u_2,1:K}\theta_e, \theta^a_{u_1,1:Z}, \theta^p_{u_2,1:Z}\eta_e, \theta^a_{u_1}\theta^p_{u_2}) \end{aligned} \tag{21}$$

We show all the complete conditionals in Table 3. The parameters for the Multinomial distribution are depicted as parameters in the variational family form.

5.4 Coordinate Updates

We define a mean-field variational family where all the variables, including auxiliary variables, are independent variables with distributions shown in Table 3. Note that we optimize the variational parameters. Each variational parameter is updated while the other variables are fixed. For example, the update for the variational parameters for the event participant influence factor $\theta_{e,k}$ is formulated as:

$$\hat{\theta}^{shp}_{e,k} = \mathbb{E}(\lambda_{ea} + \sum_{u_1}\sum_{u_2}(\alpha^a_{u_1 e,k} + \alpha^p_{u_2 e,k})) \tag{22}$$

$$\hat{\theta}^{rte}_{e,k} = \mathbb{E}(\lambda_{eb} + \sum_{u_1}\sum_{u_2}(\theta^a_{u_1,k} + \theta^p_{u_2,k})) \tag{23}$$

where $\mathbb{E}(\cdot)$ denotes the expectation. Recall that the value of variables $\alpha^a_{u_1 e,k}$ and $\alpha^p_{u_2 e,k}$ are governed by Multinomial distribution as denoted in Equation 21. Consequently, we derive $\mathbb{E}(\alpha^a_{u_1 e,k}) = D^e_{u_1,u_2}\kappa^{D^e}_{u_1 e,k_1}$, where k_1 is the index for the corresponding probability. Regarding the expectation of $\theta^a_{u_1,k}$ and $\theta^p_{u_2,k}$, we employ the knowledge that the expectation of Gamma variable is the ratio of the shape parameter and the rate parameter, e.g. $\mathbb{E}(\theta^a_{u_1,k}) = \theta^{ashp}_{u_1,k}/\theta^{arte}_{u_1,k}$. By calculating the expectation value for each element, we derive the update form as:

$$\hat{\theta}^{shp}_{e,k} = \lambda_{ea} + \sum_{u_1} D^e_{u_1,u_2}\kappa^{D^e}_{k_1} + \sum_{u_2} D^e_{u_1,u_2}\kappa^{D^e}_{k_2} \tag{24}$$

$$\hat{\theta}^{rte}_{e,k} = \lambda_{eb} + \sum_{u_1}\frac{\theta^{ashp}_{u_1,k}}{\theta^{arte}_{u_1,k}} + \sum_{u_2}\frac{\theta^{pshp}_{u_2,k}}{\theta^{prte}_{u_2,k}} \tag{25}$$

Another type of parameters that we need to update is the variational parameters of Multinomial variables. Similarly, they are updated via the expectation values. Note that the variational parameters of Multinomial variables are represented by Gamma variables. The expectation of the logarithm of a Gamma variable is, for example, $E(\log\theta^a_{u,k}) = \Psi(\theta^{ashp}_{u,k}) - \log\theta^{arte}_{u,k}$, where $\Psi(\cdot)$ is the digamma function. Hence we update one entry in κ^{D^e}, one of the Multinomial parameters, as:

$$\kappa^{D^e}(\alpha^a_{u1e,k}) \propto \exp(\Psi(\theta^{ashp}_{u_1,k}) + \Psi(\theta^{shp}_{e,k}) - \log\theta^{ashp}_{u_1,k}\theta^{shp}_{e,k}) \tag{26}$$

The entry corresponds to the variable $\alpha^a_{u_1 e,k}$. The algorithm for the parameter inference is described in Table 4.

6 RECOMMENDATION PROCEDURE

6.1 Participant Influence Calculation for Recommendation

Given a user u and a set of candidate events that are associated with some existing participants, we intend to rank the candidate events and recommend some top events to the user. Using the inference

Table 3. latent variables and complete conditionals

Variable	Type	Complete conditional	Parameters (V)
$\theta^a_{u,k}$	Gamma	$\lambda^a_{ua} + \sum_{u\neq v} \chi^2_{uv,k} + \sum_e \alpha^a_{ue,k} + \sum_g \beta^a_{ug,k} + \sum_l \gamma^a_{ul,k}$ $\lambda^p_{ub} + \sum_{u\neq v} \theta^p_{v,k} + \sum_e \theta_{e,k} + \sum_g \theta_{g,k} + \sum_l \theta_{l,k}$	$\theta^{ashp}_{u,k}, \theta^{arte}_{u,k}$
$\theta^p_{u,k}$	Gamma	$\lambda^p_{ua} + \sum_{u\neq v} \chi^2_{uv,k} + \sum_e \alpha^p_{ue,k} + \sum_g \beta^p_{ug,k} + \sum_l \gamma^p_{ul,k}$ $\lambda^a_{ub} + \sum_{u\neq v} \theta^a_{v,k} + \sum_e \theta_{e,k} + \sum_g \theta_{g,k} + \sum_l \theta_{l,k}$	$\theta^{pshp}_{u,k}, \theta^{prte}_{u,k}$
$\theta^a_{u,z}$	Gamma	$\lambda^a_{ua} + \sum_{u\neq v} \chi^1_{uv,z} + \sum_e \xi^a_{ue,z} + \sum_g \sigma^a_{ug,z} + \sum_l \pi^a_{ul,z}$ $\lambda^p_{ub} + \sum_{u\neq v} \theta^p_{v,z} + \sum_e \eta_{e,z} + \sum_g \psi_{g,z} + \sum_l \tau_{l,z}$	$\theta^{ashp}_{u,z}, \theta^{arte}_{u,z}$
$\theta^p_{u,z}$	Gamma	$\lambda^p_{ua} + \sum_{u\neq v} \chi^2_{uv,z} + \sum_e \xi^p_{ue,z} + \sum_g \sigma^p_{ug,z} + \sum_l \pi^p_{ul,z}$ $\lambda^a_{ub} + \sum_{u\neq v} \theta^a_{v,z} + \sum_e \eta_{e,z} + \sum_g \psi_{g,z} + \sum_l \tau_{l,z}$	$\theta^{pshp}_{u,z}, \theta^{prte}_{u,z}$
$\theta_{e,k}$	Gamma	$\lambda_{la} + \sum_{u_1} \sum_{u_2} (\alpha^a_{u_1 e,k} + \alpha^p_{u_2 e,k})$, $\lambda_{eb} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\theta^{shp}_{e,k}, \theta^{rte}_{e,k}$
$\theta_{g,k}$	Gamma	$\lambda_{ga} + \sum_{u_1} \sum_{u_2} (\beta^a_{u_1 g,k} + \beta^p_{u_2 g,k})$, $\lambda_{gb} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\theta^{shp}_{g,k}, \theta^{rte}_{g,k}$
$\theta_{l,k}$	Gamma	$\lambda_{la} + \sum_{u_1} \sum_{u_2} (\gamma^a_{u_1 l,k} + \gamma^p_{u_2 l,k})$, $\lambda_{lb} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\theta^{shp}_{l,k}, \theta^{rte}_{l,k}$
$\eta_{e,z}$	Gamma	$\lambda^t_{ea} + \sum_w \rho_{we,z} + \sum_g \epsilon_{ge,z} + \sum_l \zeta_{le,z} + \sum_{u_1} \sum_{u_2} (\xi^a_{u_1 e,z} + \xi^p_{u_2 e,z})$, $\lambda^t_{eb} + \sum_w \phi_{w,z} + \sum_g \psi_{g,z} + \sum_l \tau_{l,z} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\eta^{shp}_{e,z}, \eta^{rte}_{e,z}$
$\psi_{g,z}$	Gamma	$\lambda^t_{ga} + \sum_e \epsilon_{ge,z} + \sum_{u_1} \sum_{u_2} (\rho^a_{u_1 e,z} + \rho^p_{u_2 e,z})$, $\lambda^t_{eb} + \sum_e \eta_{e,z} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\psi^{shp}_{g,z}, \psi^{rte}_{g,z}$
$\tau_{l,z}$	Gamma	$\lambda^t_{la} + \sum_e \zeta_{le,z} + \sum_{u_1} \sum_{u_2} (\pi^a_{u_1 e,z} + \pi^p_{u_2 e,z})$, $\lambda^t_{lb} + \sum_e \eta_{e,z} + \sum_{u_1} \sum_{u_2} (\theta^a_{u_1,k} + \theta^p_{u_2,k})$	$\tau^{shp}_{l,z}, \tau^{rte}_{l,z}$
α^a α^p χ^1 ξ^a ξ^p χ^2	Mult	$\log \theta^a_{u,1:K} + \log \theta_e$ $\log \theta^p_{u,1:K} + \log \theta_e$ $\log \theta^a_{u,1:K} + \log \theta^p_{u,1:K}$ $\log \theta^a_{u_1,1:Z} + \log \eta_e$ $\log \theta^p_u + \log \eta_e$ $\log \theta^a_{u,1:Z} + \log \theta^p_{u,1:Z}$	κ_{D^e}
β^a β^p σ^a σ^p	Mult	$\log \theta^a_{u,1:K} + \log \theta_g$ $\log \theta^p_{u,1:K} + \log \theta_g$ $\log \theta^a_{u,1:Z} + \log \psi_g$ $\log \theta^p_{u,1:Z} + \log \psi_g$	κ_{D^g}
γ^a γ^p π^a π^p	Mult	$\log \theta^a_{u,1:K} + \log \theta_l$ $\log \theta^p_{u,1:K} + \log \theta_l$ $\log \theta^a_{u,1:Z} + \log \tau_l$ $\log \theta^p_{u,1:Z} + \log \tau_l$	κ_{D^l}
ρ ϱ	Mult	$\log \eta_e + \log \phi_w$ $\log \eta^b + \log \phi_w$	κ_w
ϵ	Mult	$\log \eta_e + \log \psi_g$	κ_e
ς	Mult	$\log \eta_e + \log \tau_l$	κ_l

Table 4. Inference algorithm

Initialization: 1: Build user-event matrix, where the entries denote the occurrences that a user participate in the event. Employ Matrix Factorization (MF) method to obtain factors θ_e and user factors $\theta_u^{a\|p}$. 2.Set group factors θ_g and location factor θ_l as the average of factors of all the related events. 3: Employ LDA to obtain the initial settings of latent event topics, obtaining the factors η_e and ϕ. 4: Set the auxiliary variables as using the expectation of the corresponding Multinomial Distribution. For example, $\alpha^a_{u_1 e,k} = \theta^a_{u_1,k}\theta_e/a$, where a is the normalization value for the distribution.
Repeat: 5: For k = {1,2,...,K} a. Update the variational parameters for the original variables related to influence factors following Equations 24 and 25 6: For z = {1,2,...,Z} a. Update the variational parameters for the original variables related to latent topics following Equation 24 and 25 7: For k = {1,2,...,K} a. Update the variational parameters for the auxiliary variables related to influence factors following Equations 26 5: For z = {1,2,...,Z} a. Update the variational parameters for the auxiliary variables related to latent topics following Equation 26 **Until convergence**

algorithm in Section 5, we are able to learn from the observed historical data the latent factors for users, events, groups, and locations, as well as the latent topics of events, groups and locations. Let e denote one of the candidate events, which is associated with the group g_e and the location l. For each existing participant v in the event e, we compute the strength of participant influence to the user u. Specifically, we first compute $PI^e_{v,u}$, $PI^{g_e}_{v,u}$, and $PI^{l_e}_{v_u}$ using Equation 3. Then an overall participant influence $PI_{u,v,e}$ is calculated by summing up the above three types of participant influence as depicted in Equation 27.

$$PI_{u_1 u_2,e} = \lambda_1 PI^e_{u_1,u_2} + \lambda_2 PI^{g_e}_{u_1,u_2} + \lambda_3 PI^{l_e}_{u_1,u_2} \tag{27}$$

where λ_1, λ_2 and λ_3 are the weights that need to be learned. With the calculated participant influence strength, we compute the preference score $S(u,e)$ by taking the average of the participant influence for all the existing participants as shown in Equation 28.

$$S(u,e) = \frac{1}{|V_e|}\sum_{v}^{V_e} PI_{v,u,e} \tag{28}$$

where V_e is the set of existing participants in the event e.

We rank all the candidate events by the preference score $S(u,e)$, and top-k events are selected as recommendations to the user u.

6.2 Weight Learning

We design a method analogously to the BPR optimization criterion [32] for learning the weights λs, which correspond to the participant influence regarding different aspects. To achieve this, we employ a regression model. Given a user u, we compute the preference for all events by Equation 28. Then we select the preference for any event pair, one of which is participated by the user, denoted as $e_{u,y}$ and the other is not, denoted as $e_{u,n}$. We design the regression ranking model as:

$$R(u, e_{u,y}, e_{u,n}) = \frac{1}{1 + \exp(-(S(u, e_{u,y}) - S(u, e_{u,n})))} \tag{29}$$

Then we find the optimal weights λ by maximizing the following log likelihood function:

$$\mathcal{L} = \log \prod_{u \in U} \prod_{e_{u,y}} \prod_{e_{u,n}} R(u, e_{u,y}, e_{u,n}) \tag{30}$$

The optimization for Equation 30 can be achieved by the standard gradient-based method. For event recommendation given a user, we compute the score for all the candidate events with Equation 28, incorporating the learned weights. Then the candidate events are ranked by the computed scores. Events with higher scores are recommended to the user.

6.3 Handling Cold-start Situations

Event recommendation becomes a cold-start problem when the candidate event is a newly created event that is not observed in the historical data. Indeed this case is common in EBSN sites since most events are newly created events. Our proposed framework is able to handle this kind of cold-start situations since it considers not only the topics of the event, but also the group and location of the event. Precisely, our proposed model is able to infer the latent topics from the collection of text descriptions of the new events. For modeling new events in the candidate events, we estimate the influence factor by averaging all the influence factors of all the events in the same group. To obtain the latent topics of this new event, we apply a topic modeling method.Furthermore, groups and locations are not new in most newly created events, which can be represented by the parameters learned by our model. Consequently, our proposed model is able to compute the participant influence given a new event.

7 EXPERIMENTS

7.1 Datasets and Evaluation Metrics

7.1.1 Data Collection. We prepare our raw data from Meetup.com, a popular EBSN website. Besides Meetup, there are several EBSN websites such as DoubanEvent and FacebookEvent. Our problem settings apply to most EBSN websites. For example, in DoubanEvent and FacebookEvent websites, each event is associated with a location and users are able to know who has decided to participate in an event. Regarding the group settings, the hosts of events on DoubanEvent and FacebookEvent are certain hub pages that have plenty of followers, which we can treat as a group. However, among these common EBSNs, only Meetup.com provides the Application Programming Interface (API) to access the exact time that a user participated in an event. Since the exact time data is an essential information, we select Meetup.com as a representative of various EBSN websites to conduct our experiments. All the data are acquired via the official API of Meetup.com. We have collected the data for four cities, namely, Los Angeles (LA), London (LD), Singapore (SG), and Hong Kong (HK), ranked by the volume of data. People in these four cities also have different lifestyles. Specifically, for each city, we have obtained the events that have been created between 1st, August 2013 and 1st, August 2015, as well as the location and time that each event was held at. A piece

Table 5. Statistics of raw data

City	LA	LD	SG	HK
Number of Event	228942	109383	26539	18593
Number of User	476782	554044	118875	66136
Number of RSVP	1358494	1066670	253126	143344
Number of Location	26613	19207	5007	3587
Number of Group	7138	7510	2006	1353
Avg. #participant per event	5.93	9.75	9.53	7.71
Avg. #event per location	8.60	5.69	5.30	5.18
Avg. #event per group	32.07	14.56	13.23	13.74

of “yes" RSVP data represents that a user registers, or participates in an event. Three elements can be extracted from each RSVP, namely, the user ID, the event ID, and the timestamp the user registered the event, from which we can identify the order the users registered the event. The API of Meetup.com provides the “create time" and “last modified time" for each RSVP. We choose “last modified time" as the timestamp of the RSVP since a user may cancel or change his RSVP. Note that the organizer of an event is regarded as the first participant of the event. We collected all the “yes" RSVPs associated with each event. Some identical events were held several times, each time with different event ids. We aggregate such events and treat them as a single event. For example, events held by the same group at the same location with the same title and descriptions are treated as one event. We remove stopwords and stem the words in the event text descriptions. We remove events that have more than fifteen participants. The reason is that for events with too many participants, the possibility of interactions between two users during the events will become relatively smaller. We notice that more than 90% of the events have less than 15 participants in our dataset. The statistics of the raw data is depicted in Table 5. The volume of the dataset for each city shows that Meetup.com is far more popular in LA and LD than in SG and HK. Moreover, the average number of events held in each group and at each location are twice larger in LA than in other cities, indicating that events are denser in LA. There are also relatively fewer participants, which may also be attributed to its denser events. To obtain some insights about user preference on event locations, we apply K-means clustering on event location to detect regions. Then we observe that most users are likely to participate in events held in certain one or two regions.

7.1.2 Data Preprocessing. We follow the method in [26] for deriving several datasets for our experiments from the above raw data. Since each event has its lifetime, i.e. creation time and activity time, the candidate events to be recommended should only include those which have already been created, but not yet occurred. Such setup can facilitate a more realistic simulation of the practical situation. For each city, we derive four datasets named as Part I, II, III, IV. The datasets are created as follows: We define four time points, simulating the situation that the recommender system is trying to recommend events at this time point. We generate a dataset for each time point. The events that are created within 6 months before the time point, as well as the corresponding RSVPs with timestamps before this particular time point, are treated as the training data of the corresponding dataset. We further select events which have been created but not yet occurred at this particular time point, and select the corresponding RSVPs of the events after the time point, to form the data for testing. Note that the candidate events in our experiments are naturally a mixture of existing events and new events, which allows our experiments to include the cold-start situations. We present the statistics of one preprocessed datasets for each city in Table 6. For each city, all the four

Table 6. Statistics of a particular dataset for each city after preprocessing

Dataset	LA	LD	SG	HK
Number of Event	35236	17822	5478	4050
Number of User	36615	27571	7140	5104
Number of RSVP	121917	73312	22277	15696
Number of Location	7623	5692	1499	1130
Number of Tuple $< u^a, u^p, e >$	329875	235354	88763	61738
Number of Group	4878	3160	852	627
IP>0.4[4]	26.6%	25.2%	25.4%	20.5%

datasets share similar statistical characteristics. We also conduct initial investigation on participant influence using the method described in Section 3 . It shows that all the influence probability are larger than 30% for all the datasets, which implies the existence of participant influence.

To evaluate the performance of event recommendation, we employ three evaluation metrics. The first metric is normalized discounted cumulative gain (NDCG) evaluation metric truncated to the top 10 recommendations, namely, NDCG@10. NDCG is a measure of ranking quality in information retrieval tasks, which is often used to measure effectiveness of web search engine algorithms. In event recommendation task, it can measure the recommendation quality of an event for a given user based on its position in the recommendation list. The formulae of NDCG@10 is described in Equation 31.

$$NDCG@10 = \frac{DCG@10}{IDCG@10} \tag{31}$$

$$DCG@10 = \sum_{i=1}^{10} \frac{2^{rel_i} - 1}{\log_2(i+1)} \tag{32}$$

where $rel_i \in \{0, 1\}$ and IDCG@10 is the DCG@10 value of an ideal ranking.

The other evaluation metrics are Precision and Recall, which are not related to the rankings of recommended events in contrast with the NDCG metric. Pre@N and Rec@N measures the ability of the recommendation system to recommend events that are finally selected by the users, where N is the number of recommended events. G is the number of selected events. S_N is the set of overlap between recommended events and selected events by the user. The formula of Pre@N and Rec@N are described in Equations 33 and 34 respectively.

$$Pre@N = \frac{|S_N|}{N} \tag{33}$$

$$Rec@N = \frac{|S_N|}{G} \tag{34}$$

7.2 Quantitative Evaluation

7.2.1 Experimental Setup. We fix each Gamma shape and rate hyper-parameter to 0.3 except the shape hyper-parameter for the background topic η^b. We set the shape hyper-parameter of η^b to 0.2. The topic factors of events, η and β, are initialized with LDA [2] on the event descriptions for each dataset separately. Note that LDA generates distributions over words for the topics, whereas β, the distribution of words over topics, is computed as the normalized counterpart of the results generated from LDA. The topics of locations and groups are initialized by minimizing the squared errors between the observed values and the cosine similarity of topics. We split the dataset into

60%, 20%, 20% as training, validating, and testing set respectively. The training, validating and testing dataset are the same for our model and all the comparative methods. We conduct grid search method on the validation set within the range [10,30,50,70,90] to determine a suitable dimension.

7.2.2 Comparative Methods. We compare the effectiveness of our proposed event recommendation model (PIERT) with three state-of-the-art models as well as our previous model PIER. All the four models have been reported to achieve better performance than lots of baselines in the event recommendation task. These comparative models are described below.

Participant Influence for Event Recommendation (PIER) [20] This is our previous model considering participant influence for event recommendation. The method models the participant influence regarding group, location, time, and words. In contrast with our proposed model, the PIER model is a simpler model that does not consider latent topics. For PIER model, we use the same parameter settings and the same validating process as the PIERT model.

Multi-Contextual Learning to Rank (MCLRE) [26] The MCLRE method exploits some contextual information of the users and events. Several features are distilled from the contextual information, which are fed into a learning-to-rank algorithm. Specifically, four types of contextual information, namely, the group information of users and events, the textual description of events, the location of the events, and the time of the event, are utilized to derive the features. Let x denote the feature set for a user-event pair and let y denote whether the user participates the event (1 for yes and 0 for no). The goal is to learn the function $h(x)$ such that the implication: $h(x_i) > h(x_j) \Leftrightarrow y_i > y_j$ holds for any user-event pairs. For MCLRE model, we follow the parameter settings as depicted in [26].

Collective Bayesian Poisson Factorization (CBPF) [43] The CBPF method is a collective matrix factorization model which takes Bayesian Poisson factorization as its basic unit to model user responses to events, social relation, and content text separately. An event is represented as a weighted combination of the organizer, the location, and the textual information. Given a user and an event, a rating is computed to evaluate the events, based on the multiple types of factors during training. The factor representation of two users tends to become similar if they are friends. Since our dataset does not provide the information of friendship, we alternatively treat two users as friends if they belong to at least one common group. We follow the parameter settings described in [43]. Besides, we tune the dimension of latent factors in the range [30,50,70,90,120] on the validation set to determine a suitable dimension.

BMF combining social group influence and individual preference (SogBmf) [7] The SogBmf method designs two types of social relations between two users. The first type of relation models the online relation. Specifically, the Jaccard similarity regarding the group membership is computed to represent the weight between two users. The second relation models the offline relation, i.e. a weight capturing the offline event co-participation. Besides the social relations, mixture ratings indicating individual preference are exploited. Gaussian regularization terms incorporating these relations are added to the Bayesian Personalized Ranking (BPR) model [32]. We follow the parameter settings described in [7]. The dimension of latent factors is tuned in the range [30, 50, 70, 90, 120] on the validation set.

7.2.3 Recommendation Results. We evaluate the effectiveness of our model PIERT, PIER, MCLRE, CBPF and SogBmf in event recommendation with NDCG@10, Pre@10, and Rec@10. The performance of each dataset as well as the average of all datasets for each city are depicted in Tables 7, 8, 9 and 10. It can be observed that our proposed models PIERT and the simplified variant PIER generally outperform MCLRE, CBPF, and SogBmf in all the datasets under all the metrics. Both PIERT and PIER considers participant influence. The results indicate that users are indeed affected

Table 7. Recommendation performance in Los Angeles data. The symbol $*$ and † denotes that PIERT and PIER is better respectively than the comparative models (i.e. SogBmf, CBPF, MCLRE) with statistical significance. Note that the statistical significance tests are done on each dataset, but not for the "average".

Metrics	Dataset	Models				
		SogBmf	CBPF	MCLRE	PIER	PIERT
NDCG@10	I	0.2445	0.2430	0.2488	0.2706†	0.2923*
	II	0.2421	0.2427	0.2492	0.2691†	0.2858*
	III	0.2411	0.2413	0.2501	0.2713†	0.2823*
	IV	0.2393	0.2460	0.2525	0.2683†	0.2886*
	average	0.2418	0.2433	0.2494	0.2698	0.2873
Pre@10	I	0.0892	0.0834	0.1093	0.1133	0.1267*
	II	0.0773	0.0945	0.1025	0.1104†	0.1186*
	III	0.0858	0.0962	0.1045	0.1144†	0.1212*
	IV	0.0901	0.0992	0.1033	0.1143†	0.1162*
	average	0.0856	0.0933	0.1049	0.1131	0.1206
Rec@10	I	0.2633	0.2704	0.2805	0.2892†	0.3044*
	II	0.2807	0.2766	0.2797	0.2816	0.3036*
	III	0.2782	0.2715	0.2927	0.2901	0.2971
	IV	0.2725	0.2758	0.2876	0.2912	0.2936*
	average	0.2737	0.2736	0.2856	0.2880	0.2996

Table 8. Recommendation performance in London data. The symbol $*$ and † denotes that PIERT and PIER is better respectively than the comparative models (i.e. SogBmf, CBPF, MCLRE) with statistical significance. Note that the statistical significance tests are done on each dataset, but not for the "average".

Metrics	Dataset	Models				
		SogBmf	CBPF	MCLRE	PIER	PIERT
NDCG@10	I	0.2021	0.2063	0.2276	0.2393†	0.2511*
	II	0.1982	0.2076	0.2267	0.2281	0.2428*
	III	0.2027	0.2092	0.2144	0.2275†	0.2432*
	IV	0.1956	0.2059	0.2138	0.2280†	0.2412*
	average	0.1997	0.2073	0.2206	0.2307	0.2446
Pre@10	I	0.0823	0.0767	0.0802	0.0796	0.0907*
	II	0.0743	0.0752	0.0768	0.0828†	0.0912*
	III	0.0835	0.0779	0.0735	0.0903†	0.0941*
	IV	0.0829	0.0784	0.0736	0.0835	0.0908*
	average	0.0807	0.0766	0.0760	0.0846	0.0917
Rec@10	I	0.2823	0.2834	0.2844	0.2933†	0.3030*
	II	0.2762	0.2788	0.2815	0.2901†	0.3024*
	III	0.2653	0.2813	0.2748	0.2845	0.2880*
	IV	0.2882	0.2794	0.2769	0.2808	0.2869
	average	0.2780	0.2810	0.2794	0.2871	0.2950

by the existing participants occasionally when they decide whether to participate in an event. The consideration of existing participants can improve the performance of event recommendation.

Table 9. Recommendation performance in Singapore data. The symbol $*$ and † denotes that PIERT and PIER is better respectively than the comparative models (i.e. SogBmf, CBPF, MCLRE) with statistical significance. Note that the statistical significance tests are done on each dataset, but not for the "average".

Metrics	Dataset	Models				
		SogBmf	CBPF	MCLRE	PIER	PIERT
NDCG@10	I	0.2036	0.2172	0.2215	0.2426†	0.2619*
	II	0.2289	0.2158	0.2212	0.2421†	0.2582*
	III	0.2196	0.2381	0.2485	0.2530	0.2550*
	IV	0.2297	0.2385	0.2402	0.2556†	0.2601*
	average	0.2204	0.2274	0.2329	0.2483	0.2588
Pre@10	I	0.0740	0.0801	0.0847	0.0944†	0.1012*
	II	0.0829	0.0831	0.0851	0.0958†	0.0994*
	III	0.0816	0.0847	0.0912	0.1001†	0.1028*
	IV	0.0851	0.0868	0.0907	0.1017†	0.1031*
	average	0.0809	0.0837	0.0879	0.0980	0.1016
Rec@10	I	0.2581	0.2732	0.2789	0.2827	0.2922*
	II	0.2747	0.2725	0.2764	0.2835†	0.2946*
	III	0.2842	0.2838	0.3004	0.3010	0.3002
	IV	0.2775	0.2921	0.2993	0.3023	0.3089*
	average	0.2736	0.2804	0.2888	0.2948	0.2990

Table 10. Recommendation performance in Hong Kong data. The symbol $*$ and † denotes that PIERT and PIER is better respectively than the comparative models (i.e. SogBmf, CBPF, MCLRE) with statistical significance. Note that the statistical significance test are done on each dataset, but not for the "average".

Metrics	Dataset	Models				
		SogBmf	CBPF	MCLRE	PIER	PIERT
NDCG@10	I	0.2477	0.2448	0.2528	0.2618†	0.2707*
	II	0.2334	0.2441	0.2563	0.2624	0.2751*
	III	0.2329	0.2335	0.2422	0.2547†	0.2775*
	IV	0.2343	0.2352	0.2415	0.2516†	0.2603*
	average	0.2371	0.2394	0.2482	0.2576	0.2709
Pre@10	I	0.0855	0.0893	0.0839	0.0925	0.1010*
	II	0.0754	0.0833	0.0850	0.0889	0.0926
	III	0.0733	0.0860	0.0811	0.0893	0.0927*
	IV	0.0831	0.0833	0.0818	0.0847	0.0914*
	average	0.0793	0.0855	0.0830	0.0889	0.0944
Rec@10	I	0.2803	0.3049	0.3255	0.3243	0.3379*
	II	0.2822	0.3157	0.3269	0.3340†	0.3450*
	III	0.3138	0.3052	0.3006	0.3171	0.3247*
	IV	0.3110	0.3088	0.3070	0.3175	0.3272*
	average	0.2968	0.3087	0.3150	0.3232	0.3337

We have also conducted statistical significance test to compare our models (i.e. PIERT and PIER) with the comparative models (i.e. SogBmf, CBPF, MCLRE) based on the paired t-test with $p < 0.05$. Each user in the test set has recommendation performance measured by our metrics. The statistical

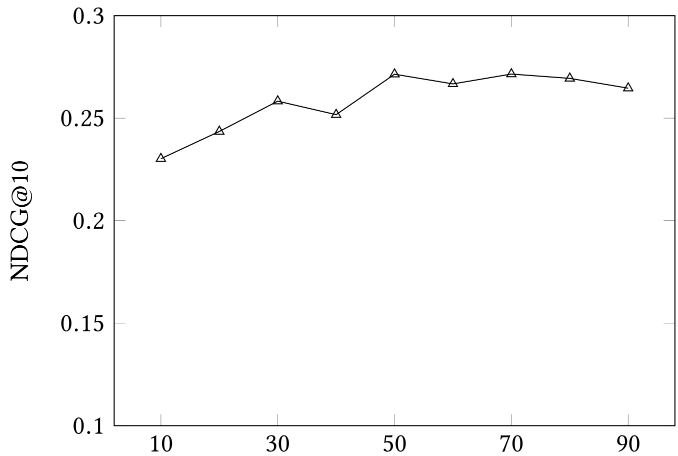

Fig. 3. Effect of varying the dimension of participant influence factors in our model.

significance test is conducted treating each user as a data point. Note that statistical significance tests are done on each dataset, but not for the “average". The “average" is calculated by averaging the results in four datasets for each city. The results show that our proposed generally outperforms all the comparative methods. In particular, the differences between our model (PIERT) and the comparative models under the evaluation metric NDCG@10 are statistically significant for almost all datasets. Under the metrics Pre@10 and Rec@10, our model also achieves superior performance than the comparative models, although the differences are not as significant as under NDCG@10. The MCLRE model generally performs better than CBPF and SogBmf. The reason is that the MCLRE model considers the most types of context information in an EBSN. CBPF is slightly better than SogBmf regarding the recommendation accuracy.

PIERT outperforms our previous simplified variant PIER. The reason is that PIERT considers the latent topics of events, and characterize groups and locations with the latent event topics. Furthermore, associations between topics and participant influence are considered. The results imply that users are attracted by the topics of the event when participating in the event.

7.2.4 Using Gaussian Instead of Poisson. In our proposed model, we employ Poisson distribution to generate the observed records. We investigate the effectiveness of Poisson distribution by replacing Poisson distribution with Gaussian distribution, leading to a Gaussian variant. Except the employed distribution, the framework of the Gaussian variant is the same with our proposed model such as the modeling of participant influence. To infer the parameters of the Gaussian variant, we follow the method introduced by a prototype of Gaussian factorization model [34] where a point estimate is found by maximizing the log-posterior. The dimensions of the latent influence facets and the topics factors are determined by grid-based search on the validation process. The performance of the Gaussian variant ($\mathcal{G}$) as well as our proposed model ($\mathcal{P}$) is presented in Table 11. We observe that our proposed model achieves better performance than the Gaussian on all the datasets. The empirical results demonstrate that Poisson distribution is superior than Gaussian distribution in capturing the characteristics of user behaviors in participating various events. The advantage of Poisson distribution can be partially attributed to its ability to model the budget of users. In our problem settings, the budget refers to the fact that a user has limited spare time to participate in events.

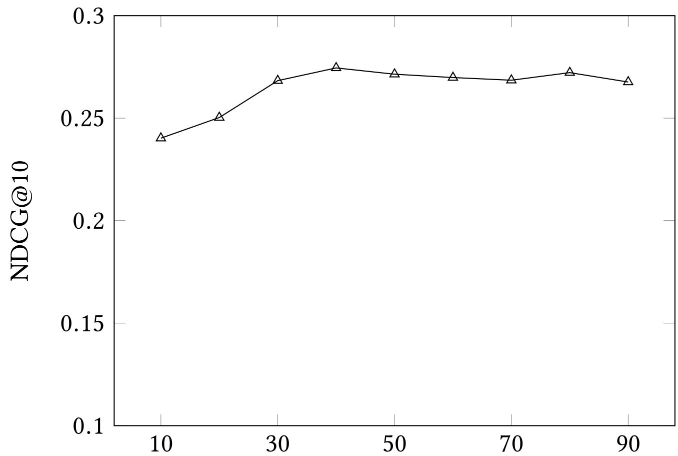

Fig. 4. Effect of varying the number of latent event topics in our model.

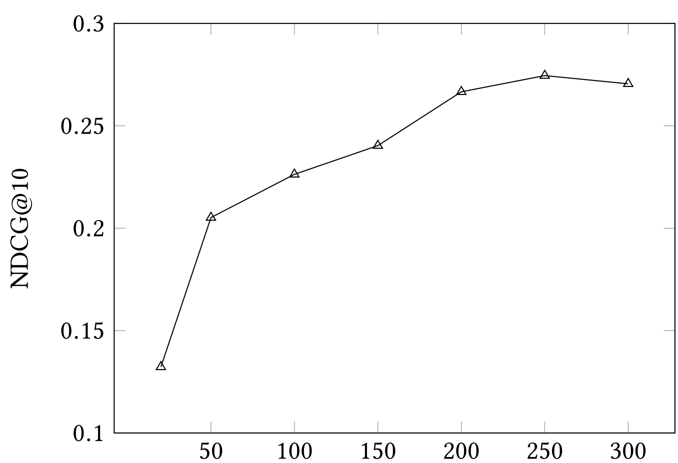

Fig. 5. Effect of varying the number of iterations

7.2.5 More Quantitative Analysis. We study the effect of the factor dimension on recommendation performance. The experiment is conducted on Hong Kong dataset. In our proposed model, there are two types of multi-dimensional vectors. The first type of vectors are participant influence vectors, including $\{\theta^a_{u,1:K}, \theta^p_{u,1:K}, \theta_e, \theta_g, \theta_l\}$. The dimensions of these vectors capture the number of participant influence facets. The second type of vectors are topic vectors, including $\{\theta^a_{u,1:Z}, \theta^p_{u,1:Z}, \eta_e, \psi_g, \tau_l\}$. The dimensions of these vectors capture the number of event latent topics. We vary the number of participant influence facets keeping the number of latent event topics as 30. The result is shown in Figure 3. It shows that the performance is relatively worse when the dimension is smaller than 50 and becomes stable in the range between 50 and 90 . Similarly, we vary the number of latent event topics keeping the number of participant influence facets as 30. The result is shown in Figure 4. The result shows that the performance is relatively worse when the dimension is smaller than 30 and becomes stable in the range between 50 and 90

Table 11. Evaluation of our Participant Influence framework with Gaussian and Poisson distribution. $\mathcal{G}$ denotes the variant with Gaussian distribution and $\mathcal{P}$ denotes our proposed model with Poisson distribution.

Metrics	Dataset	Los Angeles		London		Singapore		Hong Kong	
		$\mathcal{G}$	$\mathcal{P}$	$\mathcal{G}$	$\mathcal{P}$	$\mathcal{G}$	$\mathcal{P}$	$\mathcal{G}$	$\mathcal{P}$
NDCG@10	I	0.2577	0.2923	0.2323	0.2511	0.2385	0.2619	0.2624	0.2707
	II	0.2593	0.2858	0.2215	0.2428	0.2390	0.2582	0.2619	0.2751
	III	0.2613	0.2823	0.2185	0.2432	0.2511	0.2550	0.2658	0.2775
	IV	0.2585	0.2886	0.2235	0.2412	0.2494	0.2601	0.2489	0.2603
	average	0.2592	0.2873	0.2294	0.2446	0.2445	0.2588	0.2598	0.2724
Pre@10	I	0.1093	0.1267	0.0771	0.0907	0.0918	0.1012	0.0902	0.1010
	II	0.0935	0.1186	0.0781	0.0912	0.0994	0.0994	0.0931	0.0926
	III	0.1003	0.1212	0.0848	0.0941	0.0930	0.1028	0.0884	0.0927
	IV	0.1014	0.1162	0.0857	0.0908	0.0981	0.1031	0.0928	0.0914
	average	0.1011	0.1206	0.0814	0.0917	0.0956	0.1016	0.0911	0.0944
Rec@10	I	0.2824	0.3044	0.2801	0.3030	0.2878	0.2922	0.3262	0.3379
	II	0.2876	0.3036	0.2870	0.3024	0.2854	0.2946	0.3313	0.3450
	III	0.2854	0.2971	0.2819	0.2880	0.2819	0.3002	0.3192	0.3247
	IV	0.2833	0.2936	0.2793	0.2869	0.2966	0.3089	0.3116	0.3272
	average	0.2847	0.2996	0.2856	0.2950	0.2879	0.2990	0.3221	0.3337

We iteratively update the parameter in the inference procedure. The performance of our proposed model varying the number of iterations is shown in Figure 5. It shows that the performance becomes stable after 200 iterations.

7.2.6 Computation Efficiency. As mentioned previously, Poisson factorization only requires iterating over non-zero observations. Hence the computation complexity mainly depends on the size of observed records in the dataset. Specifically, the inference algorithm has a per-iteration computational complexity of $O(3(|D|^e + |D|^g + |D|^l)(Z + K) + (|N|_{ge} + |N|_{le} + |C|_{we})Z)$, where Z is the size of latent topics and K is the size of latent influence facets. $|D|^e$ is the size of observed records that a user participates in an event before another user, i.e. the size of identical tuples $< u^a, u^p, e >$. Similarly, $|D|^g$ and $|D|^l$ corresponds to the particular group or location that is associated with the event. $|N|_{ge}$, $|N|_{le}$, $|C|_{we}$ are the size of non-zero observations in the event-group matrix, event-location matrix, and event-word matrix respectively. Clearly, the complexity increases linearly according to the size of non-zero observations. Moreover, due to the data sparsity nature of recommendation scenarios, the non-zero observations occupy a small fraction of the matrices. For example, as shown in the LA dataset in Table 6, $|D|^e = 329875$, $|U| = 36615$ and $|E| = 35236$. $|D|^e$ is far smaller than size of the corresponding tensor, i.e. $|U| * |U| * |E|$, where $|U|$ and $|E|$ are the number of users and events respectively. In this dataset, the learning time cost in one iteration is roughly 113.7s.

7.3 Ablation Experiments

There are several components in our proposed model. To investigate the contribution of each component, we derive three models where one component is removed from our full model respectively. We compare the performance of our full model with these derived models using the LA datasets.

We derive the first model by removing participant influence modeling, which is a key component, from our full model. We denote the first model without participant influence as PIERT-PI. In this

Table 12. Ablation experiment using Los Angeles Data. The symbol $*$ denotes that PIERT is better than the derived models with statistical significance.

Metrics	Dataset	Models			
		PIERT-PI	PIERT-G	PIERT-L	PIERT
NDCG@10	I	0.2533	0.2420	0.2631	0.2923*
	II	0.2526	0.2417	0.2628	0.2858*
	III	0.2506	0.2405	0.2635	0.2823*
	IV	0.2492	0.2433	0.2617	0.2886*
	average	0.2514	0.2419	0.2628	0.2873
Pre@10	I	0.0901	0.0858	0.1106	0.1267*
	II	0.0923	0.0914	0.1073	0.1186*
	III	0.0915	0.0863	0.1075	0.1212*
	IV	0.0927	0.0918	0.1093	0.1162*
	average	0.0917	0.0888	0.1087	0.1206
Rec@10	I	0.2714	0.2644	0.2820	0.3044*
	II	0.2736	0.2681	0.2789	0.3036*
	III	0.2700	0.2672	0.2910	0.2971*
	IV	0.2688	0.2702	0.2857	0.2936*
	average	0.2709	0.2675	0.2844	0.2996

derived model, it is assumed that a user participates an event simply because the context of the event appeals to him, without the consideration of the existing participants. Precisely, the frequency that a user u participates an event e is generated from Poisson distribution with the preference parameter $Prefrence(u,e) = \sum_{d=1}^{Z+K}(\tilde{\theta}_{e,d}\theta_{u,d})$, where $\tilde{\theta}_e$ and θ_u are the factor associated with the event and the user respectively. The role of $Preference(u,e)$ replaces the participant influence defined in Equation 3. The other two models are derived by removing groups and locations from our full model. We denote these two models as PIERT-G and PIERT-L respectively. For example, we do not model D^g and N_{ge} for the PIERT-G model.

The performance of these three models, as well as our full model, is shown in Table 12. Note that the statistical significance tests are done on each dataset, but not for the "average". The results show that PIERT-PI performs worse than our full model. Hence it shows that participant influence can improve event recommendation performance. It is worthwhile investigating the influence of existing participants when designing an event recommendation system. Moreover, PIERT-G and PIERT-L also perform worse than PIERT, implying that both components contribute to our proposed full model. Specifically, PIERT-G performs worse than PIERT-L, indicating that the group information is more important than the location information in making event recommendation. The reason could be that users prefer to participate in events held by certain groups.

7.4 Qualitative Case Study

Besides making event recommendation, our proposed model is able to extract latent event topics from the event descriptions. The variable η_e captures the topics for the event z and each latent event topic is represented by a distribution over words, captured by the variable ϕ. Table 13 shows some examples of latent event topics and high-ranked words. Table 14 shows the titles of some example events that belong to each topic.

Moreover, the proposed model can learn the relations between latent event topics and the participant influence between users. As previously mentioned, each entry in the first Z dimensions

Table 13. Event topics interpretation

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
art	yoga	real	music	hike	spiritual	food	ride
artist	class	estate	perform	trail	life	wine	bike
workshop	practise	investment	band	mile	us	taste	group
draw	bring	investor	song	park	meditation	menu	bring
model	come	money	concert	organ	love	restaurant	pleasure
paint	mate	success	sing	road	experience	tea	station
light	body	deal	live	event	lives	dish	join
studio	relax	property	play	canyon	share	bring	mile
shoot	pleasure	market	jam	group	power	vegan	metro
creativity	breath	financial	drum	mountain	world	eat	lunch
work	water	webinar	jazz	anise	time	cook	meet
photograph	donation	make	night	take	practice	beer	hour
image	clothe	year	enjoy	water	inner	serve	bicycle
session	flow	start	rock	trip	god	drink	park
pose	wear	cash	musician	view	circle	potluck	adventure

of the variables $\theta^a_{u_1}$ and $\theta^p_{u_2}$ captures the degree that the corresponding topic will contribute to the participant influence from the user u_1 to the user u_2. In other words, when the context of the event, including the event descriptions, the group of the event, and the location of the event, is about a particular topic corresponding to larger entries of $\theta^a_{u_1}$ and $\theta^p_{u_2}$, the user u_2 is more likely to participate in the event if he/she knows that the user u_1 has indicate willingness for joining the event. For each user pair $< u_1, u_2 >$, our model produces a list of topics, ranked by the degree that the topic will contribute to the participant influence from u_1 to u_2. The contribution for the z-th topic, is evaluated by the product of $\theta^a_{u_1,z}$ and $\theta^p_{u_2,z}$. As a consequence, we are able to semantically explain what topics will increase the chance that the user u_1 influences the user u_2.

We present the participant influence between an example of the user pair, namely User 8837208 and User 104077492. In the dataset, we observe that User 8837208 influences User 104077492 in participating the following events:

(1) Bike Ride from North Hollywood Metro Station to Chinatown and a Chinese Lunch.
(2) Dinner Show - "The Company Men" vocal group perform at Rockwell Table & Stage.
(3) The Rascals: Once Upon a Dream at the Greek Theatre.
(4) See the comedy HOLLYWOOD & BROADWAY Gloria & Tallulah Talk (1/2 Price) & Dinner.
(5) Join us for BOWLING + Dinner at a Thai Restaurant afterward.

By analyzing the effect of topics on the participant influence on these two users, it shows that Topic 4 and Topic 8 are among the top ranked topics. It shows that User 8837208 is likely to influence User 104077492 when the event is about Topic 4 or Topic 8.

8 CONCLUSIONS

We have proposed a Poisson factorization model considering the participant influence for event recommendation. The proposed model also learns the latent event topics from event text descriptions and characterizes events, groups, and locations via a distribution of event topics. The associations between latent topics and participant influence are exploited to improve recommendation. Experimental results have demonstrated that our proposed model achieves better event recommendation

Table 14. Example events for topics

Topics	Example Events
1	Photographing Jewelry and other shiny and small objects
	Life Drawing Workshop - Model: Tracey
	Saturday Figure Drawing and Painting at Kline Academy
2	say YEP! - Good Morning Sunday Stretch
	Yogatation! (Yoga + Meditation) on the beach
	Tuesday Night Weekly Yoga
3	2014 Real Estate Update - The Changes You Need To Know
	Road to Financial Freedom Workshop
	Property Management and Due Diligence
4	Cantor Yonah Kliger and Temple Emanuel's live
	Free JAZZPOP LA Concerts and Mixer
	OC Singles for Christ Ministry
5	Forrestal Ecological Reserve
	Forest Magyar Christmas Party
	Griffith Park Northside Loop - 7 miles - 3 peaks and MORE from Travel Town
6	Goddess Wisdom Circle
	Sunday Morning Sufi Teachings in the Valley
	Community HU Chant - Let Go and Let God
7	Mix and Mingle at Ruth's Chris Steakhouse
	Chardonnay and Pinot Noir wine dinner
	Zin and BBQ Party
8	CicLAvia~ Heart of L.A. Bicycle Ride~ Mariachi Plaza Food/Music
	Bike Ride from North Hollywood Metro Station to Little Tokyo and a Japanese Lunch
	PV Ride - 50 miles with shorter route options!

performance than state-of-the-art models. As a result, event recommender systems can benefit from the consideration of participant influence.

REFERENCES

[1] M. Balabanović and Y. Shoham. Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, 1997.

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[3] J. Canny. Gap: A factor model for discrete data. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 122–129, 2004.

[4] C. Cheng, H. Yang, M. R. Lyu, and I. King. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, pages 2605–2611, 2013.

[5] A. Q. De Macedo and L. B. Marinho. Event recommendation in event-based social networks. In *HT (Doctoral Consortium/Late-breaking Results/Workshops)*, 2014.

[6] H. Gao, J. Tang, X. Hu, and H. Liu. Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM Conference on Recommender Systems*, pages 93–100, 2013.

[7] L. Gao, J. Wu, Z. Qiao, C. Zhou, H. Yang, and Y. Hu. Collaborative social group influence for event recommendation. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 1941–1944, 2016.

[8] P. Gopalan, J. M. Hofman, and D. M. Blei. Scalable recommendation with poisson factorization. *arXiv preprint arXiv:1311.1704*, 2013.

[9] P. K. Gopalan, L. Charlin, and D. Blei. Content-based recommendations with poisson factorization. In *Advances in Neural Information Processing Systems*, pages 3176–3184, 2014.
[10] J. He, X. Li, L. Liao, D. Song, and W. Cheung. Inferring a personalized next point-of-interest recommendation model with latent behavior patterns. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 137–143, 2016.
[11] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang. Social contextual recommendation. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, pages 45–54, 2012.
[12] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37(2):183–233, 1999.
[13] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.
[14] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura. Geo topic model: Joint modeling of user's activity area and interests for location recommendation. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining*, pages 375–384, 2013.
[15] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562, 2001.
[16] K. W.-T. Leung, D. L. Lee, and W.-C. Lee. Clr: a collaborative location recommendation framework based on co-clustering. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 305–314, 2011.
[17] H. Li, R. Hong, S. Zhu, and Y. Ge. Point-of-interest recommender systems: A separate-space perspective. In *Proceedins of the IEEE 15th International Conference on Data Mining (ICDM)*, pages 231–240, 2015.
[18] M. Li, A. Ahmed, and A. J. Smola. Inferring movement trajectories from GPS snippets. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, pages 325–334, 2015.
[19] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 831–840, 2014.
[20] Y. Liao, X. Lin, and W. Lam. Whether this participant will attract you to this event? exploiting participant influence for event recommendation. In *Proceedings of the IEEE 16th International Conference on Data Mining (ICDM)*, pages 1035–1040, 2016.
[21] B. Liu, Y. Fu, Z. Yao, and H. Xiong. Learning geographical preferences for point-of-interest recommendation. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1043–1051, 2013.
[22] S. Liu, B. Wang, and M. Xu. Event recommendation based on graph random walking and history preference reranking. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 861–864, 2017.
[23] 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 *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1032–1040, 2012.
[24] Y. Liu, C. Liu, B. Liu, M. Qu, and H. Xiong. Unified point-of-interest recommendation with temporal interval assessment. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1015–1024, 2016.
[25] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King. Recommender systems with social regularization. In *Proceedings of the fourth ACM International Conference on Web Search and Data Mining*, pages 287–296, 2011.
[26] A. Q. Macedo, L. B. Marinho, and R. L. Santos. Context-aware event recommendation in event-based social networks. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 123–130, 2015.
[27] E. Minkov, B. Charrow, J. Ledlie, S. Teller, and T. Jaakkola. Collaborative future event recommendation. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, pages 819–828, 2010.
[28] M. J. Pazzani and D. Billsus. Content-based recommendation systems. In *The Adaptive Web*, pages 325–341. Springer, 2007.
[29] S. Purushotham and C.-C. J. Kuo. Personalized group recommender systems for location- and event-based social networks. *ACM Trans. Spatial Algorithms Syst.*, 2:1–29, 2016.
[30] Z. Qiao, P. Zhang, Y. Cao, C. Zhou, L. Guo, and B. Fang. Combining heterogenous social and geographical information for event recommendation. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pages 145–151, 2014.
[31] Z. Qiao, P. Zhang, C. Zhou, Y. Cao, L. Guo, and Y. Zhang. Event recommendation in event-based social networks. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pages 3130–3131, 2014.
[32] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*, pages 452–461, 2009.

[33] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, pages 811–820, 2010.

[34] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, pages 1257–1264, 2008.

[35] J. Sang, D. Lu, and C. Xu. A probabilistic framework for temporal user modeling on microblogs. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 961–970, 2015.

[36] Y. Tan, M. Zhang, Y. Liu, and S. Ma. Rating-boosted latent topics: Understanding users and items with ratings and reviews. In *Proceedings of the 25th International Joint Conferences on Artificial Intelligence*, pages 2640–2646, 2016.

[37] J. Tang, C. Aggarwal, and H. Liu. Recommendations in signed social networks. In *Proceedings of the 25th International Conference on World Wide Web*, pages 31–40, 2016.

[38] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng. Learning hierarchical representation model for nextbasket recommendation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 403–412, 2015.

[39] X. Wang and Y. Wang. Improving content-based and hybrid music recommendation using deep learning. In *Proceedings of the ACM International Conference on Multimedia*, pages 627–636, 2014.

[40] T. Xu, H. Zhong, H. Zhu, H. Xiong, E. Chen, and G. Liu. Exploring the impact of dynamic mutual influence on social event participation. In *Proceedings of the 2015 SIAM International Conference on Data Mining*, pages 262–270, 2015.

[41] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann. Time-aware point-of-interest recommendation. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 363–372, 2013.

[42] C. Zeng, Q. Wang, S. Mokhtari, and T. Li. Online context-aware recommendation with time varying multi-arm bandit. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2025–2034, 2016.

[43] W. Zhang and J. Wang. A collective bayesian poisson factorization model for cold-start local event recommendation. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1455–1464, 2015.

[44] W. Zhang, J. Wang, and W. Feng. Combining latent factor model with location features for event-based group recommendation. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 910–918, 2013.

[45] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 83–92, 2014.

[46] Y. Zhang, H. Wu, V. Sorathia, and V. K. Prasanna. Event recommendation in social networks with linked data enablement. In *15th International Conference on Enterprise Information Systems*, pages 371–379, 2013.

[47] S. Zhao, T. Zhao, H. Yang, M. Lyu, and I. King. Stellar: Spatial-temporal latent ranking for successive point-of-interest recommendation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 315–321, 2016.

[48] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with GPS history data. In *Proceedings of the 19th International Conference on World Wide Web*, pages 1029–1038, 2010.