CVTM: A Content-Venue-Aware Topic Model for Group Event Recommendation

Yulu Du, Xiangwu Meng, and Yujie Zhang

Abstract—Event recommendation is essential to help people find attractive events to attend, but it intrinsically faces cold-start problem. The previous studies exploit multiple contextual factors to overcome the cold-start problem in event recommendation. However, they do not consider the correlation among different contextual factors. Moreover, suggesting events for a group of users also has not been well studied. In this paper, we first discover the correlation between organizer and textual content, i.e., the events held by the same organizer tend to have more similar content. Based on this observation, we present a content-venue-aware topic model (CVTM) to capture group interests on an event from two perspectives: content and venue. The correlation between organizer and content is modeled in CVTM to alleviate the sparsity of textual content, and then we can further extract group interests on content of an event more accurately. Finally, a group event recommendation method using CVTM is proposed. We conduct comprehensive experiments to evaluate the recommendation performance of our model on two real-world datasets. The results demonstrate that the proposed model outperforms the state-of-the-art methods that suggest upcoming events for groups. Besides, CVTM can learn semantically coherent latent topics which are useful to explain recommendations.

Index Terms—Group recommendation, probabilistic generative model, event recommendation, event-based social networks

1 Introduction

The increasing popular event-based social networks (EBSNs) [1], such as Meetup[1] and Douban Event,[2] provide online platforms for users to create, discover and share offline social events, such as concerts, exhibitions and parties. Each event published in EBSNs is associated with some attributes including an organizer who creates the event, a location where the event will be held, a timestamp when the event will start and textual content describing the event. As a large volume of events are published incessantly in EBSNs, it is difficult to find attractive events for users. Personalized event recommender systems appear as an effective solution to alleviate such an information overload. One significant difference between event recommendation and other recommendation tasks, e.g., movie recommendation, is that event possesses very short life cycle and only events have been published in EBSNs but not yet started should be recommended to users. Because few user feedback exists for events that have not taken place, serious cold-start problem arises naturally in event recommendation scenario [2], [3].

1. http://www.meetup.com
2. http://beijing.douban.com

• *The authors are with the Beijing Key Lab of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China. E-mail: {yuludu, mengxw, zhangyj}@bupt.edu.cn.*

Manuscript received 31 Aug. 2017; revised 17 Feb. 2019; accepted 3 Mar. 2019. Date of publication 8 Mar. 2019; date of current version 3 June 2020. (Corresponding author: Xiangwu Meng). Recommended for acceptance by J. Ye. Digital Object Identifier no. 10.1109/TKDE.2019.2904066

Some previous studies [2], [3], [4], [5] exploited multiple contextual factors (e.g., spatial, temporal, content and social information) to alleviate cold-start problem in event recommendation. These studies mainly focus on suggesting upcoming events for individuals, but ignore generating recommendations for a group of users who want to attend events together, e.g., attending concerts with friends, and traveling with families. The main challenge lies in finding the events to satisfy all group members with divers preferences. For example, some members in a group may like attending concerts, while others may prefer to travel. The traditional methods aiming to recommend for individuals do not consider this preference conflict between group members and can not generate recommendations for a group of users directly. Group recommender systems [6] have been proposed to solve this problem. One of the most popular group recommendation approaches is based on aggregation, which models group preference by aggregating group members' preferences using some predefined strategies [6], [7], [8]. The main drawback of this method is that it overlooks the interactions among group members [9]. Recent research [9], [10] proposed some model-based approaches which consider interactions among members by modeling the generative process of a group and show better performance than aggregation-based methods, but the cold-start problem in event recommendation for a group of users has not been well studied.

Although various contextual factors have been incorporated in the previous work on event recommendation, they exploit each contextual factor individually and ignore the correlations among different contextual factors. Intuitively, the events held by same organizer may have more similar content. For instance, the events organized by National

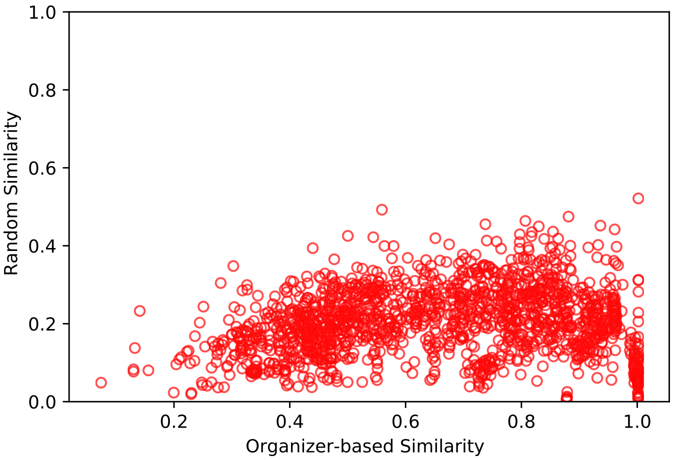


Fig. 1. The correlation between organizer-based similarity and random similarity on Douban Event dataset.

Museum of China have similar content, because they are mostly associated with the exhibition of artworks, such as historical relics, paintings, sculptures. We make a quantitative analysis to investigate the correlation between organizer and content of event. We first employ Latent Dirichlet Allocation (LDA) [11] to learn 50 latent topics from the textual content (i.e., title, tags and brief introduction) of events, and represent each event as a vector representing probability distribution over latent topics. After that, the similarity between two events is measured by the standard cosine similarity between the vectors of two events. Inspired by [12], we define *organizer-based similarity* to measure the similarity between two events that are held by same organizer, and *random similarity* to measure the similarity between two events that are randomly selected and held by different organizers. We randomly select 5000 events from Douban Event dataset, which is described in Section 5.1.1. For each event, we calculate the average organizer-based similarity and average random similarity following [12]. Fig. 1 plots the correlation between organizer-based similarity and random similarity. We notice that the plot in Fig. 1 exhibits strong biases towards the lower-right region, which indicates the high correlation between organizer and content of event, i.e., the events sharing same organizer have more similar content than the events held by distinct organizers. We argue that this finding is helpful to overcome the sparsity of textual content and further improve the performance of group event recommendation. Moreover, we also defined *venue-based similarity* to measure the similarity between two events held at same venue, but we did not observe the significant correlation between venue and content of event.

In light of this, we propose a Bayesian probability generative model, called Content-Venue-aware Topic Model (CVTM), to extract groups' venue preferences and content preferences. The correlation between organizer and textual content is modeled in CVTM to improve the quality of groups' content preferences and the effectiveness of group event recommendation. We build the model based on the following three factors that have not been exploited by previous work: 1) Each group is relevant to a content topic, e.g., a group of music funs is relevant to concert topic, and a group of tourists may be relevant to travel topic. To alleviate the sparsity of textual content, each content topic is discovered based on the co-occurrence patterns of textual content and organizer following our intuition and investigation results shown in Fig. 1. 2) A group is also associated with a venue topic, which indicates some latent features of venues, such as ticket price, facility and capacity. For example, a group of users who like drama are more likely to visit the venues where drama can be held (e.g, theaters). 3) Users consider content topic and venue topic of each group, when they make a decision to join a group. Based on the three new factors, we model the content topics, venue topics, and the process of group formation in a unified way. Based on CVTM, we propose a group event recommendation method.

The main differences and contributions of this paper can be summarized as follows:

- We investigate the correlation between organizer and content of event, i.e., the events held by same organizer have more similar content than those held by different organizers. We argue that modeling correlation between organizer and content is helpful for improving group event recommendation.
- We propose content-venue-aware topic model to learn group's content preferences and venue preferences. Moreover, the correlation between textual content and organizer is modeled to alleviate the sparsity of textual content, where some events are described with very few words.
- We conduct comprehensive experiments to evaluate our model on two real-world datasets. The experimental results show that our model outperforms the state-of-the-art methods and has a good interpretability.

The remainder of this paper is organized as follows. Section 2 introduces some related work on event recommendation and group recommendation. Section 4 details our model structure, generative process, parameter estimation and group event recommendation method based on our model. We report the experimental results and analysis in Section 5 and conclude the paper in Section 6.

2 Related Work

2.1 Event Recommendation in EBSNs

The definition of EBSNs is first provided by Liu et al. [1], who also propose two main challenges for EBSNs: community detection and event recommendation. The latter has captured more attentions in recent years [1], [2], [3], [13], [14]. Liu et al. [1] utilize online and offline social relationships between users to generate event recommendations. Qiao et al. [14] propose a latent factor model to capture heterogenous social network and geographical features of events in EBSNs. However, they ignore the cold-start problem in event recommendation, where the events to be recommended receive very few responses from users. With the growing popularity of EBSNs, more additional information can be used to alleviate cold-start problem in event recommendation. Macedo et al. [3] propose four methods utilizing social, content, temporal, spatial information, respectively, to predict scores of users to future events. After that, the predictive scores are combined using learning-to-rank techniques to generate final ranking lists. Zhang et al. [2] focus on local cold-start event recommendation task, where the recommended events have not taken place and held in the city where the target user lives. They employ Bayesian Poisson factorization as basic union to model each contextual factor and a collective matrix factorization model is created to combine multiple contextual factors. Their experimental results demonstrate that organizer plays an essential role in local cold-start event recommendation, and temporal factor is not very effective for this task. Pham et al. [13] propose a generic graph-based model for recommending events to users in EBSNs, where the rich information is modeled with a heterogeneous graph.

As described above, these previous studies do not aim to suggest events for a group of users, where the members participate events with face-to-face conversations. Besides, each contextual factor is modeled independently and the correlations among different contextual factors are not considered. Purushotham et al. [15] propose a probabilistic generative model based on collaborative topic regression to provide event recommendation for groups in EBSNs. This model jointly considers content of the events and interactions between group members. Their experimental results show that the method outperforms traditional aggregation-based group recommendation methods. However, this approach ignores some important contextual factors, such as organizer and venue, which influence user event attendance behaviors. Therefore, this approach may fail to alleviate cold-start problem effectively and achieves approving recommendation performance.

2.2 Group Recommendation

Group recommender systems, which aim to suggest items for a group of users, have attracted considerable attentions in recent years and have been applied in many domains, such as movies [8], [16], [17], TV [18], music [19], [20], travels [21], crowdfunding [22]. To incorporate members' preferences into group preference, many aggregation methods have been proposed and can be classified into three categories: 1) profile aggregation [18] aggregates group members' profiles (e.g., ratings) into group profile and use individual recommendation algorithms to generate recommendations; 2) score aggregation [6] first predicts relevant scores between users and items, and then group members' predicted scores are aggregated into group score; 3) rank aggregation [7] combines group members' recommendation lists, which are generated by traditional recommendation algorithms, to group recommendation list. Some heuristic strategies [23], e.g., average (AVG), least misery (LM), most pleasure (MP), relevant and disagreement (RD), are widely used for profile aggregation and score aggregation. Spearman footrule and Borda count are adopted for rank aggregation [7].

Recommending the upcoming events for groups in EBSNs has not been well studied. Liu et al. [24] explore social influence on group recommendation in EBSNs. Yuan et al. [9] propose a probabilistic generative model to recommend events for groups in EBSNs. These studies do not consider cold-start problem in event recommendation. Purushotham et al. [15] jointly model the textual content of events and generative processes of groups for recommending a list of events to a group in EBSNs, but the two most important factors, i.e., organizer and venue, in event recommendation [2], [3] are ignored in this work. Zhang et al. [25] propose a group recommendation method to integrate multiple contextual information including social relationships, time, location and semantic analysis. Salehi et al. [26] propose group recommendation method based on preference-oriented social network, which can make group recommendations when the preferences of some group members are unobserved. However, the correlations among contextual factors are not modeled in these studies.

TABLE 1
An Event and Its Attributes

Name: Star Wars Meetup
Organizer: Chicago Game Lovers Venue: Chicagoland Games Dice Dojo Category: Games Tags: Board Games, War Games, Card Games, Geek Culture

3 Problem Definition

In this section, we define some key concepts in our model and formulate group event recommendation problem. Throughout this paper, all vectors are denoted by bold lower case letters (e.g., $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$).

Definition 1 (Event). *An event is defined as a uniquely identified specific real-world activity, where participants can interact with each other face-to-face. An event has four attributes: identifier, organizer, venue and content. We use e to represent an event identifier, v_e to denote the venue where event e is held, and h_e to denote the organizer of event e. Besides, each event is associated with textual content, e.g., titles, categories and tags. We use the notation W_e to denote the collection of words describing event e. Table 1shows an example of an event.*

Definition 2 (Group Participation). *A group $g \in G$ consists of set of users $U_g \subset U$ who attend the same event. U denotes the set of users and $|U_g|$ denotes the size of group g, i.e., the number of users in the group g. G is the set of groups, and $|G|$ denotes the number of groups in G. For each group g, we create a group participation record $D_g = (U_g, e, v_e, h_e, W_e)$, which denotes that the members of group g attended the event e, which is created by organizer h_e, held at venue v_e and described by a set of words W_e. The dataset D consists of group participation records of groups, i.e., $D = \{D_g : g \in G\}$.*

It is worth to note that different with the online groups that organize events in some EBSNs (e.g., Meetup) , which have studied by previous work [3], [4], the group we defined is offline group where the members participate an event in the real-world. Moreover, the group we defined is ad-hoc, which means that we regard a group of users who participated two different events as two different groups, even they consist of same members.

Definition 3 (Content Topic). *Given a collection of words W, a content topic z is defined as a multinomial distribution over W, i.e., $\boldsymbol{\lambda}_z = \{\lambda_{z,w} : w \in W\}$ where each component $\lambda_{z,w}$ denotes the probability of content topic z associated with word w.*

TABLE 2
Notations of Model Parameters

Variables	Description
$\boldsymbol{\varphi}_g$	the venue-based preference of group g, represented by a multinomial distribution over a set of venue topics
$\boldsymbol{\theta}_g$	the content-based preference of group g, represented by a multinomial distribution over content topics of group g
$\boldsymbol{\phi}_r$	multinomial distribution over users specific to venue topic r
$\boldsymbol{\psi}_z$	multinomial distribution over users specific to content topic z
$\boldsymbol{\rho}_z$	multinomial distribution over organizers specific to content topic z
$\boldsymbol{\lambda}_z$	multinomial distribution over words specific to content topic z
$\boldsymbol{\eta}_r$	multinomial distribution over venues specific to venue topic r
$\alpha, \gamma, \sigma, \varepsilon$	Dirichlet priors to multinomial distributions $\boldsymbol{\theta}_g, \boldsymbol{\varphi}_g, \boldsymbol{\phi}_r$ and $\boldsymbol{\psi}_z$ respectively
π, τ, β	Dirichlet priors to multinomial distributions η_r, ρ_z and λ_z respectively

Definition 4 (Venue Topic). *Given a collection of venues V, a venue topic r is defined as a multinomial distribution over V, i.e., $\boldsymbol{\eta}_r = \{\eta_{r,v} : v \in V\}$ where each component $\eta_{r,v}$ denotes the probability of venue topic r associated with venue v.*

Problem 1 (Group Event Recommendation). *Given a group participation dataset D, a query group g_q and a set of upcoming events E_c, our goal is to recommend a list of k upcoming events that the group members in U_{g_q} are likely to participate. Because the upcoming events E_c have not started yet when recommendation is required, group event recommendation task faces a serious cold-start problem.*

4 Content-Venue-Aware Topic Model

In this section, we first describe the model structure, generative process and parameter estimation. After that, we propose a group event recommendation method based on our model.

4.1 Model Structure

For the ease of presentation, we first introduce the notations in our model and list them in Table 2. Fig. 2 shows the graphical representation of CVTM. It is a probabilistic generative model jointly over the group members, venue, organizer and words in textual content of event. Our input data, i.e., group participation $< U_g, e, v_e, h_e, W_e >$, is modeled as observed variables, shown as shaded circle in Fig. 2. The event e in group participation is not incorporated in our model explicitly, because it is not helpful to overcome cold-start problem in event recommendation. The content topic and venue topic are modeled as latent variables denoted by z and r, respectively, shown as unshaded circle in Fig. 2. Compared with the existing models [9], [24], CVTM discovers two kinds of latent topics: content topic and venue topic. Moreover, we model the group decision process with three new factors, which will be described in the following.

Organizer-Aware Content Topic Discovery. CVTM adopts latent content topics to model interests of all members in a group as a whole. Most of topic models represent a topic as the multinomial distribution over the words. However, we observe that the textual content of many events is very short and less meaningful. The classic topic models (e.g., LDA) may not infer topics with high quality from sparse corpus where each document is very short. Intuitively, the events held by same organizer may have similar content. For example, an organizer who has distributed a travel event is more likely to create another travel activity in the future. The experimental analysis on the Douban Event dataset reported in Fig. 1 demonstrates that the similarity among the events created by same organizer is higher than the similarity among events held by different organizers. In light of this, we discover content topics from both organizers and textual content of events in our model. Specifically, we assume that the set of words W_e and the organizer h_e of event e are drawn from two content topic specific multinomial distributions $\boldsymbol{\lambda}_z$ and $\boldsymbol{\rho}_z$, respectively. The correlation between organizer and content of an event can be indicated by the generative process of organizer and content in our model. The generation process of CVTM in Algorithm 4.2 shows that both the organizer and each word in the content of an event are drawn from multinomial distributions $\boldsymbol{\rho}_z$ and $\boldsymbol{\lambda}_z$, respectively. In other words, each content topic z in CVTM is not only associated with a word distribution $\boldsymbol{\lambda}_z$, but also with a distribution over organizers $\boldsymbol{\rho}_z$. The correlation between organizer and content of an event is also indicated in the graphical representation of the model shown in Fig. 2, where the organizer node h and each word node w depend on the content topic node z.

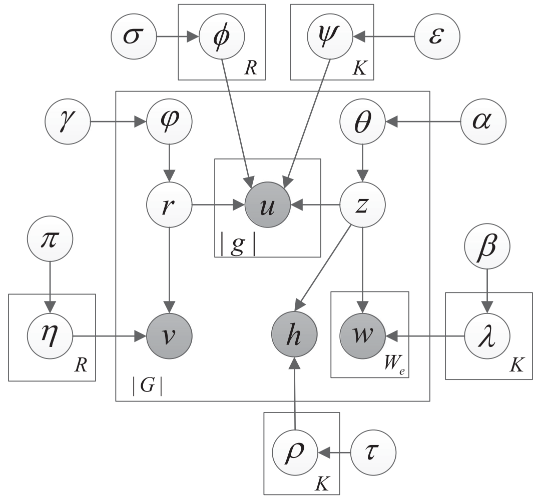


Fig. 2. The graphical representation of CVTM.

Venue Topic Modeling. We argue that the venue where an event is held also plays an important role in group decision making for attending an event. We illustrate this point from threefold perspectives. First, a venue is associated with ticket price, which impacts a user's decision. Some users are more likely to attend free event, while others like pay for the events they attended. Moreover, even the events with same content are held at different venues may have different prices. For instance, the ticket price of a concert held at a gymnasium are different with the ticket price of a concert held at a bar for same singer. Second, different venues have different limitations of the number of participants, which may impacts the group decision. Intuitively, some users are more likely to attend the large events which attracted by many people, while others may prefer to participate small events. Third, different venues have different functions, which decide what kinds of events can be held. Some venues hold various kinds of events, but other venues can only hold one kind of events. As an example, concerts, football games and exhibitions all can be held at National Stadium in Beijing, while a bar can only hold concerts. Based on above intuitions, a venue topic r in CVTM is represented by a multinomial distribution η_r over venues to capture latent features of venues and used to model group members interests on venues.

Joint Modeling For Group Participation. Visually, a group of users decide to attend an event taking into account the venue where the event is held and the content of the event that attracts users to participate. For instance, a user who likes attending Rock concerts may consider the genre of a concert and the venue for it. Only if the genre of the concert is rock and the venue is bar or club, the user would attend the concert. Therefore, we assume that each group member join in the group influenced by both content topic and venue topic of this group. Based on this assumption, we infer group content interests as a multinomial distribution over a set of content topics, denoted as $\boldsymbol{\theta}_g$ and model group venue interests as a multinomial distribution over a set of venue topics, denoted as $\boldsymbol{\varphi}_g$. To model the formation of a group based on the two kinds of topics, each venue topic r in CVTM is not only associated with a venue distribution $\boldsymbol{\eta}_r$, but also with a multinomial distribution over users $\boldsymbol{\phi}_r$. Similarly, each content topic z is also represented as a multinomial distribution over users $\boldsymbol{\psi}_z$. Thus, each member u of a group in our model is generated by content topic z and venue topic r simultaneously with the probability of $\psi_{z,u} \times \phi_{r,u}$.

Some previous studies [2], [3] show that temporal influence does not play an important role in event recommendation task. The possible reason is that the most events start during certain time. We investigate the temporal patterns of events in Douban Event and find that the start time of 80 percent events falls in seven to eleven in the evening and 75 percent events are held at weekends. Besides, some events (e.g., exhibition and travel) may be held for several days. In this case, it is difficult to explore temporal influence on event recommendation according to the start time of an event. Consequently, we do not consider temporal influence in our model.

4.2 Generative Process

We formally describe the probability generative process of the proposed model in Algorithm 4.2. To avoid overfitting, we assign a Dirichlet prior [11] over each multinomial distribution (i.e., $\boldsymbol{\theta}_g, \boldsymbol{\varphi}_g, \boldsymbol{\rho}_z, \boldsymbol{\lambda}_z, \boldsymbol{\eta}_r, \boldsymbol{\psi}_z$ and $\boldsymbol{\phi}_r$). Suppose that the members U_g in group g make decision on event attendance. The group first draws a content topic z from the multinomial distribution over content topics θ_g and draws a venue topic r from the multinomial distribution over venue topics $\boldsymbol{\varphi}_g$. With the chosen content topic z, organizer h is generated from the multinomial distribution over organizers $\boldsymbol{\rho}_z$, and each word in event description W_e is generated from multinomial distribution over words $\boldsymbol{\lambda}_z$. With the chosen venue topic r, venue v is generated by the venue topic's venue distribution $\boldsymbol{\eta}_r$. Each member $u \in U_g$ is generated in terms of probability $\psi_{z,u} \times \phi_{r,u}$. Note that, the event e attended by the group is not generated explicitly in our model. The reason is twofold: 1) modeling the generative process of an event is not helpful to alleviate the cold-start problem in group event recommendation, where the recommended events start in the future and have not been attended by users; 2) In most cases, an event can be identified by its organizer, content and venue.

Algorithm 1. Probability Generative Process in CVTM

```
for each content topic z do
    Draw ψ_z ~ Dirichlet(ε);
    Draw ρ_z ~ Dirichlet(τ);
    Draw λ_z ~ Dirichlet(β);
end
for each venue topic r do
    Draw ϕ_r ~ Dirichlet(σ);
    Draw η_r ~ Dirichlet(π);
end
for each D_g in D do
    Draw θ_g ~ Dirichlet(α);
    Draw φ_g ~ Dirichlet(γ);
    Draw r ~ Multinomial(φ_g);
    Draw z ~ Multinomial(θ_g);
    Draw v ~ Multinomial(η_r);
    Draw h ~ Multinomial(ρ_z);
    for each member u in group g do
        Draw u ~ Multinomial(ϕ_r) × Multinomial(ψ_z);
    end
    for each word w ∈ W_e do
        Draw w ~ Multinomial(λ_z);
    end
end
```

4.3 Parameter Estimation

To estimate the model parameters $\boldsymbol{\Phi} = \{\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\psi}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\rho}}\}$ in our proposed CVTM, our goal is to maximize the marginal log-likelihood of the group participation data (U_g, e, v_e, h_e, W_e). However, the log likelihood is difficult to maximized directly. Therefore, we follow the two-step Gibbs sampling algorithm [9] to maximize the complete data likelihood in Equation (1). We first sample content topic z_i for ith group participation, and then sample venue topic r_i for ith group participation. For each latent variable (e.g., z_i), a collapsed Gibbs sampling method, which is a Markov Chain Monte Carlo (MCMC) algorithm, computes the full conditional probability for the assignment of the variable conditioned on all the other assignments (e.g., z_j).

$$\begin{aligned} & p(\boldsymbol{U_g}, \boldsymbol{v}, \boldsymbol{h}, \boldsymbol{W}, \boldsymbol{z}, \boldsymbol{r} | \alpha, \beta, \gamma, \tau, \varepsilon, \pi, \sigma) \\ = & \int p(z|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\alpha) d\boldsymbol{\theta} \int p(\boldsymbol{r}|\boldsymbol{\varphi}) p(\boldsymbol{\varphi}|\gamma) d\boldsymbol{\varphi} \int p(\boldsymbol{v}|\boldsymbol{r}, \boldsymbol{\eta}) p(\boldsymbol{\eta}|\pi) d\boldsymbol{\eta} \\ & \int p(\boldsymbol{h}|z, \boldsymbol{\rho}) p(\boldsymbol{\rho}|\tau) d\boldsymbol{\rho} \int p(\boldsymbol{W}|z, \boldsymbol{\lambda}) p(\boldsymbol{\lambda}|\beta) d\boldsymbol{\lambda} \\ & \int \int p(\boldsymbol{U_g}|\boldsymbol{r}, z, \boldsymbol{\phi}, \boldsymbol{\psi}) p(\boldsymbol{\phi}|\sigma) p(\boldsymbol{\psi}|\varepsilon) d\boldsymbol{\phi} d\boldsymbol{\psi}. \end{aligned} \tag{1}$$

More specifically, in the first step, we iteratively sample venue topics for all group participation data (U_g, e, v_e, h_e, W_e).

For ith group participation, a venue topic index is drawn from the following distribution:

$$p(r_i = r|\boldsymbol{r}_{\neg i}, \boldsymbol{z}, \boldsymbol{U_g}, \boldsymbol{v}, \boldsymbol{h}, \boldsymbol{W}) \\ \propto (n_{g,r}^{\neg i} + \gamma) \frac{n_{r,v}^{\neg i} + \pi}{\sum_{v'}(n_{r,v'}^{\neg i} + \pi)} \prod_{u \in U_g} \frac{n_{r,u}^{\neg i} + \sigma}{\sum_{u'}(n_{r,u'}^{\neg i} + \sigma)}, \quad (2)$$

where $r_{\neg i}$ represents venue topic assignments for all group participation records except the current one; $n_{g,r}$ denotes the number of times that venue topic r is assigned to group g. $n_{r,u}$ is the number of times that venue topic r is assigned to user u. $n_{r,v}$ is the number of times that venue topic r is assigned to venue v. $n_{\cdot}^{\neg i}$ denotes the count excluding the current record.

In the second step, after venue topic $\boldsymbol{r}$ is sampled, the content topic assignment z_i for ith group participation record is drawn from the following distribution:

$$p(z_i = z|\boldsymbol{r}, \boldsymbol{z}_{\neg i}, \boldsymbol{U_g}, \boldsymbol{v}, \boldsymbol{h}, \boldsymbol{W}) \propto (n_{g,z}^{\neg i} + \alpha) \frac{n_{z,h}^{\neg i} + \tau}{\sum_{h'}(n_{z,h'}^{\neg i} + \tau)} \\ \times \prod_{u \in U_g} \frac{n_{z,u}^{\neg i} + \varepsilon}{\sum_{u'}(n_{z,u'}^{\neg i} + \varepsilon)} \prod_{w \in W_e} \frac{n_{z,w}^{\neg i} + \beta}{\sum_{w'}(n_{z,w'}^{\neg i} + \beta)}, \quad (3)$$

where $n_{g,z}$ counts the number of times that content topic z is assigned to group g. $n_{z,h}$ denotes the number of times that content topic z is assigned to organizer h. $n_{z,u}$ represents the number of times that user u is sampled from topic z, and $n_{z,w}$ is the number of times that token w is sampled from topic z. Note that both venue topic and content topic can be assigned to a group at most once. Therefore, the values of $n_{g,z}$ and $n_{g,r}$ are binary. For instance, $n_{g,z} = 1$ if the content topic z is assigned to group g. Otherwise, $n_{g,z} = 0$.

After a sufficient number of sampling iterations, the approximated posteriors can be used to estimate model parameters as follows:

$$\hat{\psi}_{z,u} = \frac{n_{z,u} + \varepsilon}{\sum_{u'}(n_{z,u'} + \varepsilon)} \quad \hat{\phi}_{r,u} = \frac{n_{r,u} + \sigma}{\sum_{u'}(n_{r,u'} + \sigma)} \\ \hat{\rho}_{z,h} = \frac{n_{z,h} + \tau}{\sum_{h'}(n_{z,h'} + \tau)} \quad \hat{\lambda}_{z,w} = \frac{n_{z,w} + \beta}{\sum_{w'}(n_{z,w'} + \beta)} \\ \hat{\eta}_{r,v} = \frac{n_{r,v} + \pi}{\sum_{v'}(n_{r,v'} + \pi)}. \quad (4)$$

Time Complexity of Parameter Estimation. Suppose our parameter estimation algorithm needs I iterations to reach convergence. In each iteration, it requires go through all group participation dataset D. For each record in D, it requires $O(K(|\overline{U_g}| + |\overline{W_e}|))$ and $O(R|\overline{U_g}|)$ operations to compute the posterior distributions for sampling content topic z and venue topic r, respectively, where $|\overline{U_g}|$ denotes the average size of group, i.e., the average number of members in a group and $|\overline{W_e}|$ represents the average number of words introducing an event. In real group event recommendation scenario, suggesting upcoming events for very large group is not very common [27] and the number of tags and categories of event can be regarded as very short text. Moreover, both $|\overline{U_g}|$ and $|\overline{W_e}|$ do not increase with the growth of the size of dataset. Thus, we can regard them as constant and ignore them in time complexity analysis. Finally, the whole time complexity of our parameter estimation algorithm is $O(I(R + K)|D|)$, where $|D|$ denotes the number of records in dataset D. Thus, the time complexity is linear to $|D|$.

4.4 Group Event Recommendation Using CVTM

Once we have estimated the model parameter set $\hat{\boldsymbol{\Phi}} = \{\hat{\boldsymbol{\psi}}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\eta}}\}$ are estimated, we can calculate the probability of upcoming event $e \in E_c$ attended by a target group g_t. Because the groups are ad-hoc, which means $g_i \neq g_j$, even if $U_{g_i} = U_{g_j}$, given a target group g_t, we have to discover its content topic distribution and venue topic distribution based on group members U_{g_t} by performing Gibbs sampling as follows:

$$p(r_i = r, z_i = k|\boldsymbol{z}_{\neg i}, \boldsymbol{r}_{\neg i}, U_{g_t}) \\ \propto (n_{g_t,r}^{\neg i} + \gamma)(n_{g_t,z}^{\neg i} + \alpha) \prod_{u \in U_{g_t}} \hat{\psi}_{k,u} \hat{\phi}_{r,u}. \quad (5)$$

After Gibbs sampling achieves convergence, we estimate the content topic distribution θ_{g_t} and venue topic distribution φ_{g_t} for target group g_t as follows:

$$\hat{\theta}_{g_t,z} = \frac{n_{g_t,z} + \alpha}{\sum_{z'}(n_{g_t,z'} + \alpha)} \qquad \hat{\varphi}_{g_t,r} = \frac{n_{g_t,r} + \gamma}{\sum_{r'}(n_{g_t,r'} + \gamma)}. \quad (6)$$

Based on the generative process of CVTM, described in Algorithm 4.2, we calculate the recommendation score of a candidate event e given a target group g_t as follows:

$$p(e, v_e, h_e, W_e|U_{g_t}, \hat{\theta}_{g_t}, \hat{\varphi}_{g_t}, \boldsymbol{\Phi}) \propto \sum_z \hat{\theta}_{g_t,z} \hat{\rho}_{z,h_e} (\prod_{w \in W_e} \hat{\lambda}_{z,w})^{\frac{1}{|W_e|}} \\ \times \sum_r \hat{\varphi}_{g_t,r} \hat{\eta}_{r,v_e} (\prod_{u \in U_{g_t}} \hat{\psi}_{z,u} \hat{\phi}_{r,u})^{\frac{1}{|U_{g_t}|}}. \quad (7)$$

Time Complexity of Recommendation. The group event recommendation using CVTM consists of two parts. The first part aims to learn θ_{g_t} and φ_{g_t} by performing Gibbs sampling algorithm using Equation (5). Suppose the algorithm needs I iterations to reach convergence. In each iteration, it requires $O((R + K) \times |U_{g_t}|)$ operations to compute multinomial distributions for the target group over content topic z and venue topic r, respectively. The second part calculating the recommendation score requires $O((K \times R \times |U_{g_t}| + K \times |W|) \times |E_c|)$ operations, where E_c denotes the set of candidate events and $|E_c|$ is the number of events in E_c, to calculate the recommendation scores of all candidate events for target group g_t.

5 Experiments

In this section, we first present our settings of experiments and then report the experimental results.

5.1 Experimental Settings

5.1.1 *Datasets*

We conduct experiments on two real-world datasets crawled from Douban Event and Meetup, two most popular EBSNs, respectively. EBSN provides an online platform for users to create, organize, distribute and share offline events, such as concerts, exhibitions, travels and salons. Once an event is distributed on EBSNs, any registered user can RSVP[3] to it with

3. RSVP stands for a French phrase "répondez s'il vous plaît", meaning "please reply".

TABLE 3
The Statistics of DoubanEvent and Meetup Datasets

Dataset	Douban Event		Meetup	
City	Beijing	Shanghai	Chicago	Phoenix
#Users	1146	3204	1976	1048
#Events	1050	2286	23882	13702
#Organizers	390	758	722	369
#Venues	539	1055	1813	922
#RSVPs	10697	28123	54798	30089
#Groups	10000	10000	10000	10000
Avg. #Events per group	5.5	10.02	2.96	5.15
Sparsity	99.11%	99.62%	99.88%	99.79%

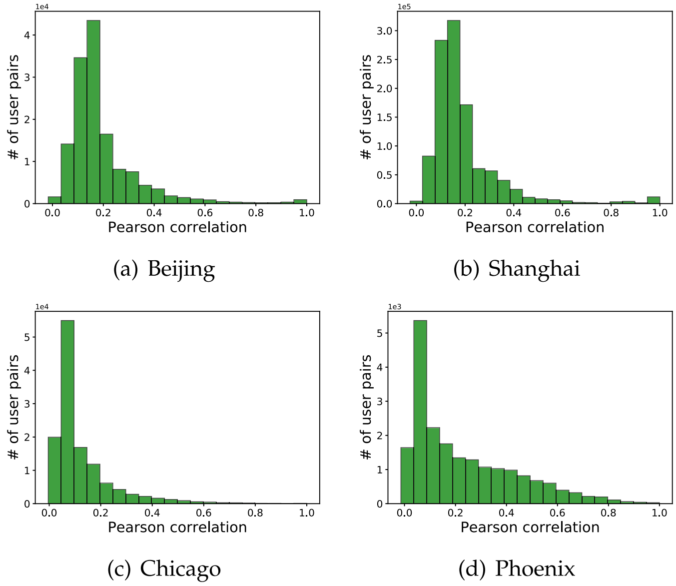
"Yes" or "No" representing whether the user will attend the event or not. We use Douban API to collect the events in Beijing and Shanghai, the two largest cities in China, from September, 2016 to June, 2017. The Meetup dataset is collected by [3] from January, 2010 to April, 2014 using the Meetup REST API.[4] We select the events held in Chicago and Phoenix from the Meetup dataset for our experiments. For each event in both datasets, we obtain its organizer, participants, textual content, venue and start time. The textual content of each event consists of title, categories, tags and introduction. To make our model simple, only categories and tags are considered as textual content in experiments. Because the datasets do not include any groups required for group event recommendation task, we have to generate some synthetic groups, described in Section 5.1.2. Before that, we conduct data preprocessing to reduce the sparsity of datasets. For Douban Event dataset, we only use the events held between September, 2016 and December, 2016. For Meetup dataset, the events from January, 2013 to December, 2014 are used for experiments. Moreover, for each city, we remove the users who attend less than 10 events to filter noisy data. Table 3 summarizes basic statistics of the datasets. We also report the sparsity of the datasets, as the percentage of missing RSVPs in the user-event matrix.

5.1.2 Evaluation Methods

As we discussed in the previous sections, event recommendation inherently suffers from cold-start problem, because most of the events that will be recommended to the users have not started yet and few responses received from users. For this reason, we evaluate our model for group event recommendation in the cold-start scenario, where all the events to be recommended have not received any response, thus the effectiveness of our model alleviating cold-start problem can be evaluated.

In order to simulate the realistic scenario of group event recommendation, where the events required to be recommended are cold-start, we propose a modified time-dependent cross-validation method [28]. For each city, we first sort the events chronologically according to their start time. Then, we partition events into four parts according to their start time and perform 3-fold time-dependent cross-validation. As an example, the events in Beijing and Shanghai start in four months, from September, 2016 to December 2016, and the events in each month are token into one partition.

4. http://www.meetup.com/meetup_api/



(a) Beijing (b) Shanghai

(c) Chicago (d) Phoenix

Fig. 3. User-to-user similarity.

For first fold validation, we use the events in the first month for training and remaining events are used to test. For second fold validation, we exploit the events in first two months for training and remaining events are used to test. For third fold validation, the events in first three months are used for training. Because each experiment have more events for training than the previous validation, the data sparsity of each fold is different. For example, the training data of first fold is more sparse than that of second fold. Thus, the effectiveness of recommendation methods on different degrees of data sparsity can be evaluated. The similar cross-validation settings are performed on Meetup dataset.

Because both Douban Event and Meetup datasets do not contain any group information, it is necessary to create some synthetic groups to evaluate our model for group event recommendation task. Note that simulating group is a common setting in group recommendation evaluations [7], [9], [16], [29], because there are very few datasets that contain group information. Due to the datasets are very sparse, we generate groups based on similarities between users following [7], [8], where Pearson correlation coefficient (PCC) is used to measure the user-to-user similarity. Figs. 3a, 3b, 3c and 3d show the user-to-user similarity distribution for Beijing, Shanghai, Chicago and Phoenix, respectively. The groups are defined as those containing users with user-to-user similarity higher than a certain threshold. We set the threshold as average of user-to-user similarity following [8]. Thus, the thresholds for group generation of Beijing, Shanghai, Chicago and Phoenix are set to 0.197, 0.192, 0.128 and 0.233, respectively. Finally, for each city, we generate 2000 groups with different sizes from 2 to 6. The average number of the events attended by each group is shown in Table 3. We do not consider the group consists of users more than 6 for two reasons. First, it is not common scenario in the real world that large groups attend events. Second, it is difficult to find enough groups, the size of which is larger than 6.

To measure the effectiveness of our model, we employ Recall and Normalized Discounted Cumulative Gain (NDCG), both of which are standard metrics for evaluating the effectiveness of group recommender systems [7], [16], [29]. Specifically, we define Recall@k and NDCG@k in Equations (8) and (10) to measure the accuracy of a group event recommender that produces a list of k events for each group.

$$Recall@k = \frac{|R_k \cap T|}{|T|}, \quad (8)$$

where R_k denotes the set of recommended k events. T denotes the set of events in test set.

$$DCG@k = rel_1 + \sum_{i=2}^{k} \frac{rel_i}{log(i+1)}, \quad (9)$$

$$NDCG@k = \frac{DCG@k}{IDCG@k}, \quad (10)$$

where rel_i is the graded relevance of the event at position i. If the event at position i in the recommendation list appears in the test set, $rel_i = 1$, if otherwise, $rel_i = 0$. IDCG@k is the maximum DCG@k value of all possible recommendation lists with length k. The higher Recall@k and NDCG@k indicate the better performance of recommendation. We compute average of Recall@k and NDCG@k for overall groups as final evaluation metrics. The precision is not considered to evaluate our model for two reasons, because the precision is defined as the proportion of relevant events in recommendation lists and misses unknown positives [30], which is also discussed in our previous work [31].

5.1.3 Comparison Approaches

To the best of our knowledge, there are few literatures focusing on recommending events for a group of users. Hence, we compare our models with some competitive group recommendation methods that are not designed to suggest events specially. On the other hand, some personalized event recommendation methods are compared with our approach. Because these methods are designed to recommending evens for individuals, we adopt some aggregation strategies introduced in Section 2.2 to combine recommendations for individual users to recommendations for the whole group. In our experiments, we compare our method with following approaches.

COM [9] is a probabilistic generative model considering topic-dependent influences of group members. Moreover, the personal considerations of content factors can be incorporated into COM via prior distribution. There are two kinds of content factors, i.e., geographical distance and movie cast, considered in [9] for venue recommendation and movie recommendation, respectively. In our experiments, we only incorporate geographical distance into COM for our group event recommendation task.

PCGR [15] extents collaborative topic regression [32], which combines latent Dirichlet allocation and matrix factorization, to model the group generative process and content influence for group event recommendation. To the best of our knowledge, PCGR is the most related work to our task, i.e., suggesting events for a group of users in EBSNs.

HBGG [10] jointly models group geographical topic, group mobility regions and social relations among groups, users and items. HBGG is designed to recommend venues for groups, but we use it to recommend events in our experiments.

MCLRE [3] combines multiple contextual factors including social, temporal, content and geographic influences into event recommendation for individual users. Each contextual factor is first modeled independently. In the end, all contextual factors are combined through learning to rank technique. Because MCLRE is designed to recommend for individual user, we employ two aggregation strategies, i.e., AVG and RD [6], to incorporate predicted scores of group members into group predicted score, which is used to sort the candidate events to generate recommendation list. Moreover, we also adopt rank aggregation method, i.e., using Borda count (BC) [7] to aggregate the ranked list of each group member to get the final group ranking. Thus, three group event recommendation methods based on MCLRE, i.e., MCLRE-AVG, MCLRE-RD and MCLRE-BC, are evaluated and compared with our model in experiments.

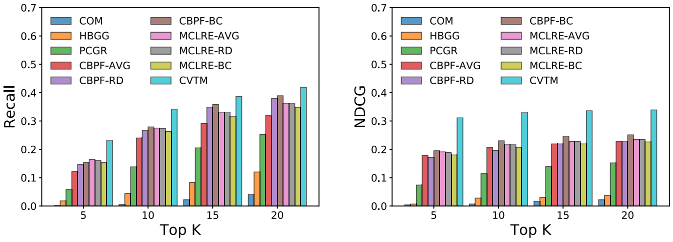
CBPF [2] aims to solve local cold-start event recommendation problem and utilizes Bayesian Poisson factorization as basic unit to model social relation, venue, organizer and textual content of event respectively. Then units are connected through collective matrix factorization. Because CBPF is designed to making recommendations for individual users, the same aggregation strategies used for MCLRE are employed to aggregation the individual recommendations produced by CBPF into group recommendation. Finally, three comparison methods based on CBPF, i.e., CBPF-AVG, CBPF-RD and CBPF-BC are evaluated in our experiments.

In order to evaluate the effect of three new factors that we proposed in Section 4.1, we design six variant versions. CVTM-V1 is the first version which does not consider venue topic modeling component. CVTM-V2 ignores the organizer-aware content topic modeling component. CVTM-V3 generates each group member only based on content topic z. CVTM-V4 generates each group member only based on venue topic r. CVTM-V5 does not generate each word in the content of an event and CVTM-V6 does not generate organizer of an event.

5.2 Recommendation Effectiveness

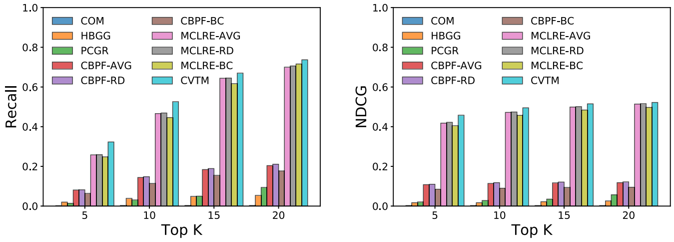
In this subsection, we report the comparison results between our models and other baselines with well-turn parameters for group event recommendation in cold-start scenario on Meetup and Douban Event datasets, respectively. Figs. 4, 5, 6 and 7 report the Recall and NDCG scores of recommendation methods for Beijing, Shanghai, Chicago and Phoenix, respectively. We only show the Recall@k and NDCG@k scores where k is set to 5,10,15,20, as greater value of k is usually ignored for a top-k recommendation task [9], [33].

Recommendation on Douban Event. Figs. 4a and 4b present the Recall and NDCG values of group event recommendation methods on Beijing dataset, respectively. Recall@10 and NDCG@10 of CVTM are about 0.342 and 0.331, respectively. Obviously, our proposed CVTM outperforms other comparison methods, and the improvements, in terms of Recall@10, are 6740, 677.27, 147.83, 42.5, 28.09, 22.58, 24.36, 25.27, and 30.04 percent compared with COM, HBGG,



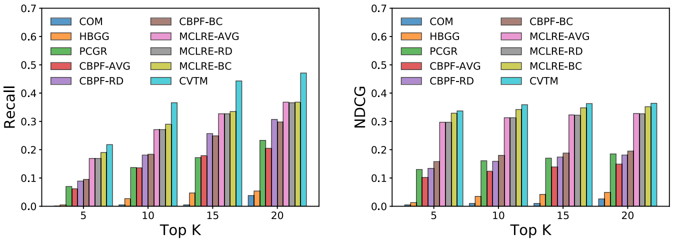
(a) Recall (b) NDCG

Fig. 4. Top-k performance on Beijing dataset.



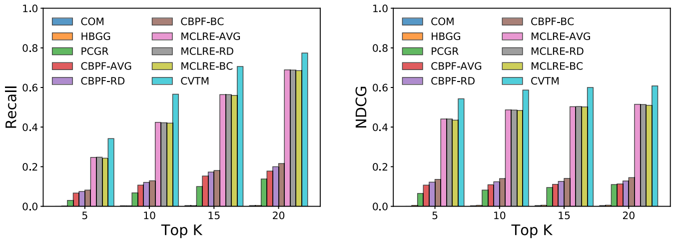
(a) Recall (b) NDCG

Fig. 6. Top-k performance on Chicago dataset.



(a) Recall (b) NDCG

Fig. 5. Top-k performance on Shanghai dataset.

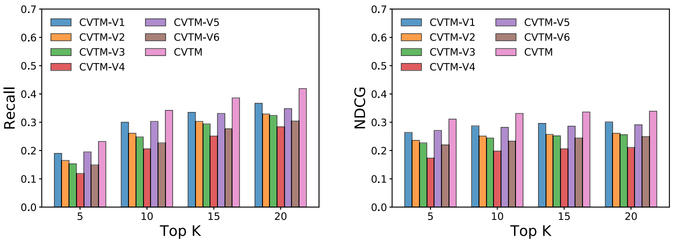


(a) Recall (b) NDCG

Fig. 7. Top-k performance on Phoenix dataset.

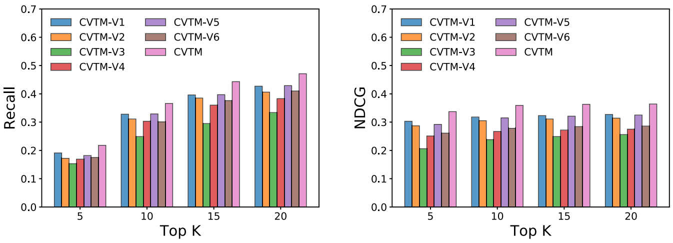
PCGR, CBPF-AVG, CBPF-RD, CBPF-BC, MCLRE-AVG, MCLRE-RD, and MCLRE-BC, respectively. There are several observations made from the results: 1) CVTM achieves higher recommendation performance than HBGG and PCGR, showing the benefit of joint modeling group interests on content and venue, because HBGG only models geographical effect and PCGR only considers content influence. 2) CVTM outperforms MCLRE-based methods and CBPF-based methods, illustrating the benefit of joint modeling for group participation and organizer-aware content topic discovery. MCLRE and CBPF are designed to recommend events for individuals. Although aggregation methods are used to incorporate their recommendations into group recommendations, the interactions between group members are not modeled. Moreover, both methods model organizer and content influence individually, but they ignore the correlation between organizer and content of an event. Thus, they fail to alleviate the sparsity of textual content. 3) Among comparison methods, MCLRE-based methods and CBPF-based methods outperform others. This is because that MCLRE and CBPF consider multiple contextual factors (i.e., social, content, geographical and temporal information), while COM and HBGG only consider geographical influence, and PCGR only incorporates the content information. 4) PCGR achieves higher Recall and NDCG values than HBGG, indicating that the contribution of textual content for improving performance is larger than the contribution of geographical factor. This conclusion is consistent with the experimental results in [2], [3]. 5) There is no clear winner among aggregation methods, i.e., AVG, RD and BC. For example, Fig. 4 shows that MCLRE-AVG and MCLRE-RD outperform MCLRE-BC, while CBPF-BC achieves higher Recall and NDCG values than CBPF-AVG and CBPF-RD. The previous study [23] shows that the performance of aggregation method is largely dependent on the algorithm used to generate individual recommendations. Fig. 5a and 5b show the Recall and NDCG values of group event recommendation methods on Shanghai dataset, respectively. We observe similar results that our model still outperforms other comparison methods on Shanghai dataset. Besides, compared with Fig. 4b, the improvements of our model over MCLRE-based methods in terms of NDCG are smaller in Fig. 5b.

Recommendation on Meetup. Figs. 6 and 7 report the performance of the recommendation models for Chicago and Phoenix in Meetup, respectively. From the Figures, we can see that the trend of comparison result is similar to that presented in Figs. 4 and 5, and the main difference is that MCLRE-based methods and our model perform better on Meetup dataset, while COM, HBGG, PCGR and CBPF-based methods achieve lower Recall and NDCG values. The possible reason is that the RSVPs data in Meetup is more sparse than that in Douban Event (e.g., 99.88 percent for Chicago versus 99.11 percent for Beijing). COM, HBGG and PCGR consider fewer contextual factors. Therefore, they perform worse on more sparse data. Moreover, the tags and categories of an event in Meetup dataset are less than those in Douban Event dataset. This leads to lower performance of PCGR, because it assumes that multiple topics can be extracted from the textual content of an event, which does not perform well over short texts. CVTM assigns one content topic to all words in textual content of an event. This assumption is more reasonable for modeling short texts. CBPF-based methods seem to be very sensitive to datasets, but the specific cause is not clear. On the other hand, the organizer of an event in Meetup is an online community (or groups), while any user can be an organizer and create events in Douban Event. The users in a community are more likely to attend the events held by the community than others. In other words, the organizer in Meetup may has larger influence on user's event attendance behavior than it in Douban Event. Thus, our model and MCLRE-based methods, which consider organizer influence, perform better on Meetup dataset.



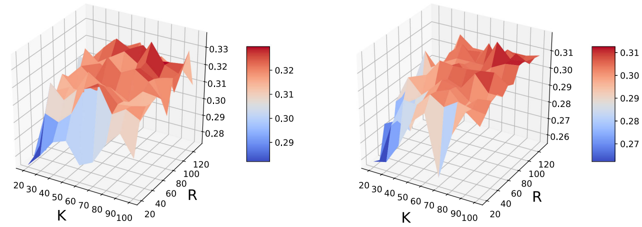
(a) Recall (b) NDCG

Fig. 8. Impact of different factors on Beijing dataset.



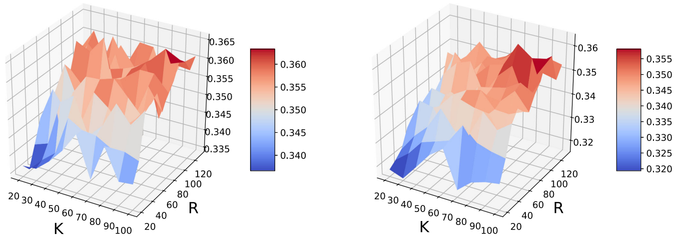
(a) Recall (b) NDCG

Fig. 9. Impact of different factors on Shanghai dataset.



(a) Recall@10 (b) NDCG@10

Fig. 10. Impact of K and R on Beijing dataset.



(a) Recall@10 (b) NDCG@10

Fig. 11. Impact of K and R on Shanghai dataset.

5.3 Impact of Different Factors

In this subsection, we study the impact of different factors on recommendation performance of CVTM. Six factors are investigated in this experiment by evaluating six variant versions introduced in Section 5.1.3. We compare CVTM with its six variant versions and report Recall and NDCG results on Shanghai dataset and Beijing dataset in Figs. 8 and 9, respectively.

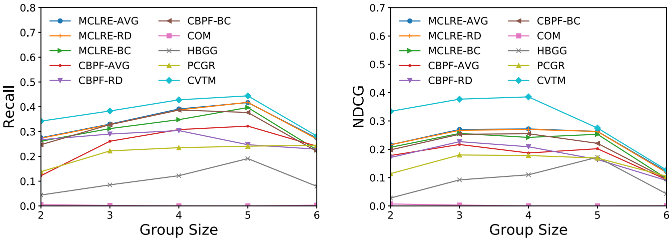
We observe that CVTM consistently outperforms the six variant versions in terms of Recall and NDCG on Beijing and Shanghai datasets, which indicates the benefit brought by each factor respectively. For example, the performance gap between of CVTM over CVTM-V1 shows the benefit of venue topic modeling. CVTM-V1 achieves higher accuracy than other variant versions, demonstrating that organizer-aware content topic modeling provides most contributions on recommendation performance. Besides, we observe that each factor has different contributions to improve recommendation performance on different datasets. Specifically, according to the performance of six variant versions on Beijing dataset, we can rank them as follows: CVTM-V1 $\geq$ CVTM-V5 $>$ CVTM-V2 $>$ CVTM-V3 $\geq$ CVTM-V6 $>$ CVTM-V4, while we can rank them as follows on Shanghai dataset: CVTM-V1 $\geq$ CVTM-V5 $>$ CVTM-V2 $>$ CVTM-V6 $\geq$ CVTM-V4 $>$ CVTM-V3. The possible reason is that the user preferences are influenced by content factor more than venue factor in Beijing, and the generation of group members based on content topic plays more important role than it based on venue topic, i.e., CVTM-V3 $>$ CVTM-V4, on Beijing dataset. The users in Shanghai consider venue factor more than content factor, which leads to CVTM-V4 $>$ CVTM-V3 on Shanghai dataset.

5.4 Influence of Parameters

In this subsection, we aim to validate the impact of varying parameters, i.e., the number of venue topics (R) and the number of content topics (K) in CVTM on the two datasets. For other hyperparameters in CVTM, we empirically set them with fixed values ($\alpha = 50/K, \gamma = 50/R, \beta = \tau = \pi = \delta = \varepsilon = 0.01$), following the studies [9], [33]. We conduct experiments with different hyperparameters and find that the performance of CVTM is not sensitive to these hyperparameters, but its performance is sensitive to the values of K and R.

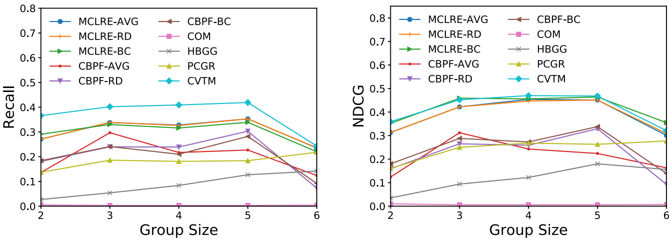
Figs. 10a and 10b show the Recall@10 and NDCG@10 on Beijing dataset, respectively. The number of venue topics R increases from 20 to 130 and the number of content topics K increases from 20 to 100. From the results, we observe that CVTM achieves the best performance, i.e., Recall@10=0.337 and NDCG@10=0.319, when $K = 80$ and $R = 100$. The accuracy does not change significantly when K is larger than 80 and R is larger than 100. Fig. 11 shows the similar results that our model achieves the highest NDCG@10=0.364 when K = 80 and R = 100, and highest Recall@10=0.365 when K = 50 and R = 90. The performance of CVTM tends to increase with the improvement of K and R. When K and R are larger than a threshold, accuracy of the model does not change significantly and begins to decrease. The reason is that the training data associated with a specific topic becomes extremely sparse, when K and R are too large. This easily leads to the overfitting in model training and reduces the recommendation accuracy.

We observe that 100 venue topics required by our model seem too large for our datasets. Actually, a venue topic r can be not only represented by the multinomial distribution over venues $\hat{\eta}_r$, but also represented by the multinomial distribution over users $\hat{\phi}_r$, which is used to generate group members. In other words, in order to model the co-occurrence patterns of group members and venue, our model requires more venue topics to simultaneously generate group members and the venue than that required by general topic models, such as LDA, where no co-occurrence patterns are modeled. Moreover, if there are more co-occurrence patterns of group members and venue, more venue topics will be needed to achieve optimal performance. Similarly, a content topic z can



(a) Recall@10 (b) NDCG@10

Fig. 12. Recall@10 and NDCG@10 for different size of groups on Beijing dataset.



(a) Recall@10 (b) NDCG@10

Fig. 13. Recall@10 and NDCG@10 for different size of groups on Shanghai dataset.

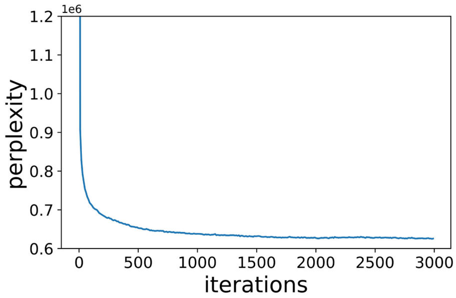


Fig. 14. Perplexity of the model.

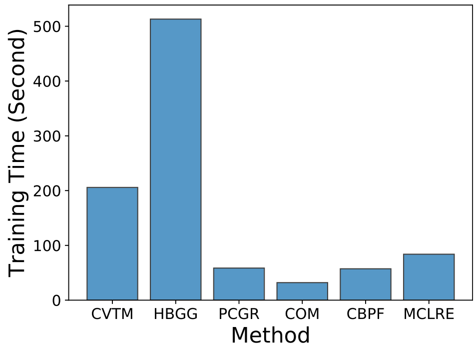


Fig. 15. Training time.

be represented by multinomial distributions over words, organizers and users, i.e., λ_z, ρ_z and ψ_z, respectively. Thus, our model requires more content topics to simultaneously describe the generative process of words, organizers and group members. Because there are fewer organizers than venues in the datasets, there are less co-occurrence patterns of group members and organizer than that of group members and venue. Therefore, compared with the number of content topics we needed, less venue topics are required, i.e., $K < R$.

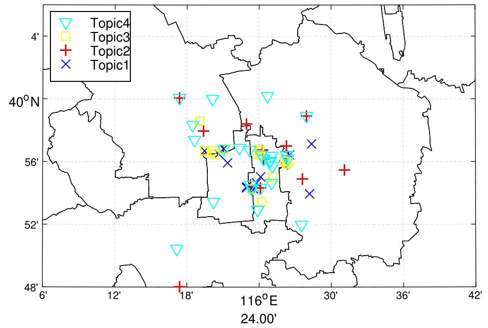
5.5 Performance for Different Size of Groups

In this experiment, we study the performance of each recommendation approach for groups of different sizes. The number of content topics is fixed at 80 and the number of venue topics is fixed at 100. We report the Recall@10 and NDCG@10 values for each approach on Beijing and Shanghai datasets in Figs. 12 and 13, respectively. We observe that CVTM outperforms other baselines for groups of different sizes on both datasets in terms of Recall@10 and NDCG@10. Among baseline approaches, MCLRE-based methods perform best for all groups. PCGR achieves higher performance than other two model-based group recommendation methods, i.e., COM and HBGG. This demonstrates that modeling the group interest on content of an event is more effective than modeling group mobility region and geographical influence. Moreover, we observe that the recommendation performance of each method does not changes significantly with different group sizes. The reason is that each group consists of users with high similarity instead of selecting randomly [7], which makes the performance does not decrease significantly for large groups.

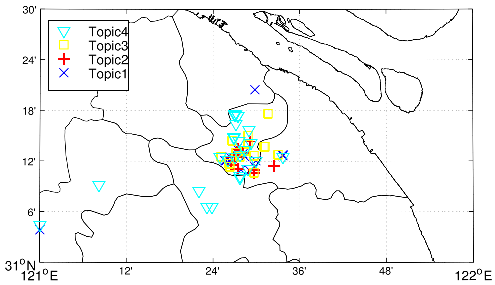
5.6 Recommendation Efficiency

This experiment is to evaluate the efficiency of CVTM on Beijing dataset. We compare our model with five baseline methods described in Section 5.1.3. All methods were implemented in Java (JDK 8) and run on a Windows 7 with 8G RAM. The time complexity analysis in Section 4.3 shows that the number of iterations has a great influence on efficiency of model training. Thus, we first investigate the number of iterations CVTM needed for achieving convergence by evaluating the perplexity of each iteration. Perplexity is widely used to evaluate the performance of topic models, and a lower perplexity score indicates better generalization performance [11]. Fig. 14 shows the perplexity of CVTM on Beijing dataset. We observe that the perplexity of CVTM decreases fast when the number of iterations is less than 500 and does not change significantly after 2500 iterations.

The running time of model training is reported in Fig. 15, which shows that CVTM is trained faster than HBGG. The reason is that the parameter estimation of two parts, i.e., group geographical model and social-based collaborative filtering, in HBGG requires more operations than our model. The training of PCGR and COM is faster than the training of CVTM, because PCGR only extracts topics from textual content and does not consider organizer and venue influence, and COM only models geographical information through prior distribution, which is not involved in model training. CBPF and MCLRE also cost fewer time for training than CVTM, because CBPF adopts variational inference, which is usually more efficiency than Gibbs sampling [34], and MCLRE is trained by stochastic gradient descent which can achieve convergence very fast. However, there are many methods to speed up the learning of our model. For example, parallel implementation of topic model [33] and fast Gibbs sampling [34], which shows 3× speedup at least compared with standard Gibbs sampling for LDA, can be easily adopted to reduce the training time of CVTM.



(a) Beijing



(b) Shanghai

Fig. 16. Geographical distribution of venue topic in Beijing and Shanghai.

5.7 Interpretability

In this subsection, we analysis two types of latent topics, i.e., content topic and venue topic, in our model. Table 5 shows five content topics discovered by CVTM on Beijing dataset. For each topic, we present the top ten words with the highest generation probabilities. We observe that the five topics are obviously different, and each of them is semantically coherent. For example, Topic 1 contains words related to tourism events, where users attend outdoor sports, such as hiking, skiing and climbing. Besides, users usually participate tourism activities with their friends and are likely to take some photograph on their journey. Topic 2 is associated with movie events, where users watch some movies, listen some stories about movies and even interact with actors. Topic 3 is related to music events, where the representative words are music genres, such as folk, country, rock, hiphop and Jazz. Topic 4 and Topic 5 are about art and party, respectively.

The venue topics are analyzed from geographical perspective and semantic perspective, respectively. To analyze the geographical distribution of venues in different venue topics, we first select four venue topics discovered by CVTM on Beijing dataset and Shanghai dataset, respectively. The geographical distributions of top ten venues in each venue topic on both datasets are shown in Fig. 16. We find that the venues in the same topic are not consistently closer to each other geographically than the venues in different topics, which is very different from the meaning of topics in [9], where the venues are clustered by topics. This indicates that the venue topics discovered by our model do not represent the geographical features of venues, but the semantic features shown in the following experimental analysis.

TABLE 4
Venue Topic Analysis on the Beijing Dataset

	Name of venue
Topic 1	1. National Centre for the Performing Arts 2. Beijing Tianqiao Performing Arts Center 3. The Ullens Center for Contemporary Art 4. National Theatre of China 5. The Capital Theatre
Topic 2	1. SDX Joint Publishing Company 2. Peking University 3. Beijing Sport University 4. China Renmin University Press 5. The Other Shore Bookstore
Topic 3	1. Modernsky Lab 2. Wukesong Arena 3. China International Exhibition Centre 4. Forbidden City Concert Hall 5. Haidian Theatre
Topic 4	1. Beijing Railway Station 2. XIZHIMEN Station 3. WUDAOKOU 4. Taikoo Li Sanlitun 5. HUIXINXIJIE NANKOU Station
Topic 5	1. Jianghu Bar 2. 13 Club 3. TrainSpotting Cafe&Bar 4. 1905 Film Club 5. CNEX Salon Cafe

We further analyze the semantic presentation of venue topics and report five venue topics discovered by CVTM on Beijing dataset in Table 4. For each venue topic, the names of top 5 venues with highest generation possibilities are presented in Table 4. From the results, we find that venue topics indicate the type of the venues. Specifically, Topic 1 clusters the large venues associated with arts, such as National Center for the Performing Arts, which is China's top performing arts center and The Capital Theater, which is the first theater built in Beijing after the founding of the PRC. Topic 2 gathers the venues related to education, including universities, presses, and bookstores. Topic 3 is similar to Topic1, but the top 5 venues in Topic 3 is multi-function and large-scale venues. For instance, both basketball games and concerts can be held at Wukesong Arena, a famous venue in Beijing. Topic 4 is about traffic stations,

TABLE 5
Content Topic Analysis on the Beijing Dataset

	tags
Topic 1	tourism, photography, travel, hiking, skiing, waterfall, friends, climbing, recreation, meeting
Topic 2	movie, art, film, story, actor, music, documentary, free, exchange, exhibition
Topic 3	music, folk, Jazz, country, master, tour, guitar, rock, hiphop, party
Topic 4	art, design, exhibition, exposition, solo show, fashion, museum, picture, international, contemporary
Topic 5	friends, single, love, game, food, party, photography, reading, salon, sport

which may be the meeting points of some outdoor activities. Topic 5 consists of the venues that are suitable for small-scale performance, such as clubs and bars.

6 Conclusion

In this paper, we discovered the strong correlation between organizer and textual content, i.e., the events held by same organizer have more similar content than those held by different organizers. Based on our finding, we proposed a probabilistic generative model CVTM to jointly learn groups' content preferences and venue preferences. CVTM modeled group participation behaviors based on three factors: 1) Organizer-aware content topic discovery based on the correlation between organizer and content information to alleviate sparsity of textual content and further capture group interests on content of the event; 2) Venue topic modeling to capture group interests on the venue where event will be held; 3) Joint modeling for group participation to generate group members depended on both venue and content topics. We conducted extensive experiments on two real-world datasets crawled from Douban Event and Meetup, respectively, to evaluate the recommendation performance of our model. The results showed that our proposal outperforms other comparison methods in terms of both Recall and NDCG values. We validated the contributions of different factors to CVTM, and found that the organizer-aware content topic discovery plays a dominate role in improving recommendation performance. Besides, the proposed model can learn coherent latent topics which are helpful to explain the recommendations. However, the model training is not very efficient due to the Gibbs sampling for two kinds of latent topics. The group event recommendation based on the proposed model has a relative high time complexity and cannot support real-time recommendation on large dataset. In our future work, we will study fast offline training algorithm and efficient online recommendation method for our proposed model.

Acknowledgments

This work is supported by the Mutual Project of Beijing Municipal Education Commission of China.

References

[1] 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 *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 1032–1040.

[2] W. Zhang and J. Wang, "A collective bayesian poisson factorization model for cold-start local event recommendation," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 1455–1464.

[3] A. Q. Macedo, L. B. Marinho, and R. L. Santos, "Context-aware event recommendation in event-based social networks," in *Proc. 9th ACM Conf. Recommender Syst.*, 2015, pp. 123–130.

[4] Y. Jhamb and Y. Fang, "A dual-perspective latent factor model for group-aware social event recommendation," *Inf. Process. Manage.*, vol. 53, no. 3, pp. 559–576, 2017.

[5] L. Gao, J. Wu, Z. Qiao, C. Zhou, H. Yang, and Y. Hu, "Collaborative social group influence for event recommendation," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 1941–1944.

[6] S. Amer-Yahia, S. B. Roy, A. Chawlat, G. Das, and C. Yu, "Group recommendation: Semantics and efficiency," *Proc. VLDB Endowment*, vol. 2, no. 1, pp. 754–765, 2009.

[7] L. Baltrunas, T. Makcinskas, and F. Ricci, "Group recommendations with rank aggregation and collaborative filtering," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 119–126.

[8] O. Kaššák, M. Kompan, and M. Bieliková, "Personalized hybrid recommendation for group of users: top-n multimedia recommender," *Inf. Process. Manage.*, vol. 52, no. 3, pp. 459–477, 2016.

[9] Q. Yuan, G. Cong, and C.-Y. Lin, "Com: A generative model for group recommendation," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2014, pp. 163–172.

[10] Z. Lu, H. Li, N. Mamoulis, and D. W. Cheung, "Hbgg: A hierarchical bayesian geographical model for group recommendation," in *Proc. SIAM Int. Conf. Data Mining*, 2017, pp. 372–380.

[11] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, no. Jan, pp. 993–1022, 2003.

[12] H. Ma, "On measuring social friend interest similarities in recommender systems," in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2014, pp. 465–474.

[13] T.-A. N. Pham, X. Li, G. Cong, and Z. Zhang, "A general graph-based model for recommendation in event-based social networks," in *Proc. IEEE 31st Int. Conf. Data Eng.*, 2015, pp. 567–578.

[14] Z. Qiao, P. Zhang, Y. Cao, C. Zhou, L. Guo, and B. Fang, "Combining heterogenous social and geographical information for event recommendation," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 145–151.

[15] S. Purushotham and C.-C. J. Kuo, "Personalized group recommender systems for location-and event-based social networks," *ACM Trans. Spatial Algorithms Syst.*, vol. 2, no. 4, 2016, Art. no. 16.

[16] M. S. Pera and Y.-K. Ng, "A group recommender for movies based on content similarity and popularity," *Inf. Process. Manage.*, vol. 49, no. 3, pp. 673–687, 2013.

[17] L. Quijano-Sánchez, B. Díaz-Agudo, and J. A. Recio-García, "Development of a group recommender application in a social network," *Knowl.-Based Syst.*, vol. 71, pp. 72–85, 2014.

[18] Z. Yu, X. Zhou, Y. Hao, and J. Gu, "Tv program recommendation for multiple viewers based on user profile merging," *User Model. User-Adapted Interaction*, vol. 16, no. 1, pp. 63–82, 2006.

[19] J. F. McCarthy and T. D. Anagnost, "Musicfx: An arbiter of group preferences for computer supported collaborative workouts," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 1998, pp. 363–372.

[20] A. Crossen, J. Budzik, and K. J. Hammond, "Flytrap: Intelligent group music recommendation," in *Proc. 7th Int. Conf. Intell. User Interfaces*, 2002, pp. 184–185.

[21] K. McCarthy, M. Salamó, L. Coyle, L. McGinty, B. Smyth, and P. Nixon, "Cats: A synchronous approach to collaborative group recommendation," in *Proc. FLAIRS Conf.*, 2006, pp. 86–91.

[22] V. Rakesh, W.-C. Lee, and C. K. Reddy, "Probabilistic group recommendation model for crowdfunding domains," in *Proc. 9th ACM Int. Conf. Web Search Data Mining*, 2016, pp. 257–266.

[23] T. De Pessemier, S. Dooms, and L. Martens, "Comparison of group recommendation algorithms," *Multimedia Tools Appl.*, vol. 72, no. 3, pp. 2497–2541, 2014.

[24] X. Liu, Y. Tian, M. Ye, and W.-C. Lee, "Exploring personal impact for group recommendation," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage.*, 2012, pp. 674–683.

[25] Y. Zhang, "Grorec: A group-centric intelligent recommender system integrating social, mobile and big data technologies," *IEEE Trans. Serv. Comput.*, vol. 9, no. 5, pp. 786–795, Sep./Oct. 2016.

[26] A. Salehi-Abari and C. Boutilier, "Preference-oriented social networks: Group recommendation and inference," in *Proc. 9th ACM Conf. Recommender Syst.*, 2015, pp. 35–42.

[27] M. O'connor, D. Cosley, J. A. Konstan, and J. Riedl, "Polylens: A recommender system for groups of users," in *Proc. 7th Conf. Eur. Conf. Comput. Supported Cooperative Work*, 2001, pp. 199–218.

[28] P. G. Campos, F. Díez, and I. Cantador, "Time-aware recommender systems: A comprehensive survey and analysis of existing evaluation protocols," *User Model. User-Adapted Interaction*, vol. 24, no. 1/2, pp. 67–119, 2014.

[29] V. R. Kagita, A. K. Pujari, and V. Padmanabhan, "Virtual user approach for group recommender systems using precedence relations," *Inf. Sci.*, vol. 294, pp. 15–30, 2015.

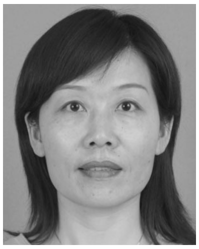
[30] P. Cremonesi, F. Garzotto, and R. Turrin, "Investigating the persuasion potential of recommender systems from a quality perspective:an empirical study," *ACM Trans. Interactive Intell. Syst.*, vol. 2, no. 2, pp. 1–41, 2012.

[31] P. Lv, X. Meng, and Y. Zhang, "Fere: Exploiting influence of multidimensional features resided in news domain for recommendation," *Inf. Process. Manage.*, vol. 53, no. 5, pp. 1215–1241, 2017.

[32] C. Wang and D. M. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 448–456.
[33] H. Yin, X. Zhou, B. Cui, H. Wang, K. Zheng, and Q. V. H. Nguyen, "Adapting to user interest drift for poi recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 10, pp. 2566–2581, Oct. 2016.
[34] I. Porteous, D. Newman, A. Ihler, A. Asuncion, P. Smyth, and M. Welling, "Fast collapsed gibbs sampling for latent dirichlet allocation," in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2008, pp. 569–577.

Yulu Du received the master's degree in 2013. He is currently working toward the PhD degree in the School of Computer Science, Beijing University of Posts and Telecommunications, Beijing. His research interests include recommender systems and intelligent information processing.

Xiangwu Meng received the PhD degree from the Institute of Software, Chinese Academy of Sciences, in 1997. He is a full professor with the Beijing University of Posts and Telecommunications. His research interests include network services, communication software, and artificial intelligence.



Yujie Zhang is an associate professor with the Beijing University of Posts and Telecommunications. Her research interests include recommender systems, network services, and intelligent information processing.