COM: a Generative Model for Group Recommendation

Quan Yuan†*, Gao Cong†, Chin-Yew Lin‡
†School of Computer Engineering, Nanyang Technological University, Singapore 639798
{qyuan1@e., gaocong@}ntu.edu.sg
‡Microsoft Research, Beijing, China 100080
cyl@microsoft.com

ABSTRACT

With the rapid development of online social networks, a growing number of people are willing to share their group activities, *e.g.,* having dinners with colleagues, and watching movies with spouses. This motivates the studies on group recommendation, which aims to recommend items for a group of users. Group recommendation is a challenging problem because different group members have different preferences, and how to make a trade-off among their preferences for recommendation is still an open problem.

In this paper, we propose a probabilistic model named COM (COnsensus Model) to model the generative process of group activities, and make group recommendations. Intuitively, users in a group may have different influences, and those who are expert in topics relevant to the group are usually more influential. In addition, users in a group may behave differently as group members from as individuals. COM is designed based on these intuitions, and is able to incorporate both users' selection history and personal considerations of content factors. When making recommendations, COM estimates the preference of a group to an item by aggregating the preferences of the group members with different weights. We conduct extensive experiments on four datasets, and the results show that the proposed model is effective in making group recommendations, and outperforms baseline methods significantly.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information Filtering

Keywords

Group Recommendation; Collaborative Filtering; Topic Models

1. INTRODUCTION

Recommender systems (RS) aim to suggest items for users based on their preferences, and they have been widely deployed to assist users to select items in various fields, such as movies (Netflix), products (Amazon), restaurants (Yelp), etc. A number of rec-
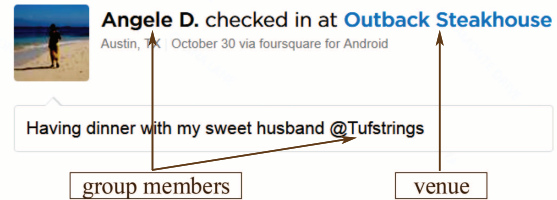
Figure 1: An Example of Group Check-in

ommendation techniques have been proposed, such as user/item-based collaborative filtering (CF) [17, 21], clustering CF [24], matrix factorization [11], etc., and most of them focus on producing recommendations for individual users. However, people often participate in activities together with others, *e.g.,* having dinners with colleagues, watching movies with spouses, and having picnics with friends. This calls for recommendation techniques for a group. Unfortunately, recommender systems designed for individuals are not effective in making recommendations for a group of people. Furthermore, an increasing number of group event records are becoming available on the web, since users often share their group activities on social networks, such as Facebook, Meetup, and Foursquare. For example, the Foursquare check-in in Figure 1 shows that user *Angele D.*, together with her husband, visited the Outback steakhouse. The availability of group event data promotes the research interest on how to make effective recommendations for a group of users [3, 4, 6, 8, 13, 26], which not only facilitates groups making decisions, but also helps web services improve user engagement.

However, making accurate recommendations for groups is not an easy task, because a group consists of multiple users who have different preferences. How to make a trade-off among their preferences to recommend items for a group is challenging. Previous solutions to group recommendation can be divided into two types: memory-based and model-based approaches. Memory-based approaches further fall in two categories based on the aggregation strategy: *preference aggregation* strategy first aggregates the profiles of group members into a new profile, and then employs recommendation techniques designed for individuals to make group recommendations [15, 28]; *score aggregation* strategy first calculates a recommendation list for each group member, and then aggregates these lists for group recommendation [4, 7, 14, 18, 20]. However, both strategies overlook the interactions between group members, and use trivial methods to aggregate members' preferences. Different from memory-based approaches, model-based methods exploit the interactions among members by modeling the generative process of a group [13, 26]. However, as to be detailed in Section 2, the assumptions of these models may not hold in real life.

To achieve better accuracy, we propose a latent Dirichlet allocation (LDA) [5] based generative model, named COnsensus Model (COM), to make group recommendations. COM is novel since it is built based on the following three considerations that have not been exploited by previous work:

1. Each group is relevant to several topics, *e.g.,* a picnic group is relevant to hiking and dining topics, and a movie watching group consisting of families may be relevant to the romance and comedy topics. The item selection of a group is influenced by both these relevant topics and group members' personal considerations of content factors, such as the geographical distance to venues for venue recommendation, and casts of movies for movie recommendation.

2. Users in a group may behave differently as group members from as individuals, *e.g.,* a user may prefer horror movies when he is alone, but will select romantic movies when watching with his wife.

3. Different users have different influences in making decisions in a group, and the influence degree of a user in a group is topic-dependent: a movie fan is probably influential in making decisions for a movie watching group, but is less influential in a dining group, because the dining group is less relevant to the movie topic.

Based on the three considerations, we model the generative process of a group as follows: each group has a multinomial distribution over latent topics, and these topics attract a set of users to join. The item selection of a user is influenced by both the group topic that attracted her, and her personal considerations of content factors (Consideration 1). Note that it is the topic of the group, instead of the user's, that account for her item selection (Consideration 2). The final decision of a group is made by aggregating the selections of all users in the group: if a user is an expert in the relevant topics of the group, her selections will have a larger weight (Consideration 3). Based on the generative model, we propose a recommendation method to suggest items for a target group.

In summary, the contributions of this paper are three-fold:

- We propose a generative model COM for modeling the process of item selection of a group, which considers members' topic-dependent influences and members' group behaviors.

- We develop a recommendation method to make group recommendations based on COM, which is able to exploit both users' selection history and users' personal considerations of content factors.

- We evaluate the effectiveness of the proposed method by extensive experiments on four datasets for event venue recommendation and movie recommendation. The experimental results show that our proposed method outperforms five baselines significantly by various evaluation metrics.

The rest of this paper is organized as follows. In Section 2, we review related work. Section 3 introduces the proposed COM model and the recommendation method. We present experimental results in Section 4. Finally, Section 5 concludes this paper.

2. RELATED WORK

We first briefly review recommendation systems in general, and then focus on techniques for group recommendation.

2.1 Recommender Systems

Recommender systems can be classified into three categories: content-based, collaborative filtering (CF), and hybrid recommendation approaches [1]. The content-based approaches make recommendations based on the content features of users (*e.g.,* age,

gender, etc.) and items (*e.g.,* price, category, etc.), but do not exploit users' rating/selection history. Thus, when content features are not sufficient, the content-based techniques will fail to produce accurate recommendations. In contrast, CF approaches rely on rating/selection history, but do not make use of content features. Consequently, they suffer from the sparsity problem, *i.e.,* when the number of users' ratings is not enough for finding similar users, the performance of CF techniques will be bad. The hybrid approaches combine the content-based and CF methods to avoid their limitations. Our proposed model exploits both users' selection history and content information, and thus it belongs to the hybrid approaches. In addition, when the content information is not available, our model will be reduced as a CF model that makes recommendations based on users' selection history only.

CF approaches have been extensively studied and can be further divided into two categories, namely, *memory-based CF* and *model-based CF* [23]. Memory-based CF approaches employ rating/selection history to find similar users of the target user, and then compute a recommendation score for a candidate item by a weighted combination of historical ratings on the item from these similar users. In contrast, model-based CF builds recommendation models using data mining techniques, such as clustering [19], matrix factorization [11], probabilistic topic model [2, 25], etc. Our model exploits users' selection history in the model-based manner.

2.2 Group recommendation

Group recommendation techniques have been proposed for various domains, such as web/news pages [20], tourism [16], restaurants [14], music [7], TV programs [28], and movies [18]. Group recommendation methods in the earlier studies fall into two categories [3]: the *preference aggregation* approaches first aggregate the profiles of the group members into one profile, and makes recommendations based on the aggregated profile [15, 28]. The *score aggregation* approaches, in contrast, first produce recommendations for each group member respectively, and then aggregate their recommendation results for the group [4, 7, 14, 18, 20].

Compared with preference aggregation, score aggregation typically enjoys better flexibility [3, 10, 18], and thus receives more research attention. The score aggregation approaches usually employ either average (AVG) or least misery (LM) strategy to aggregate the recommendations of individuals. The AVG strategy averages the recommendation scores of all group members as the final score, aiming to maximize the overall satisfactions of a group [15, 28]; the LM strategy takes the smallest recommendation score of group members as the final score, and tries to make everyone happy [4]. For LM, the recommendation score of an item is largely influenced by the user who dislikes it most, even if all the others like it very much. For AVG, an item's relevance to different users might be diverse, and the recommendation results might be unfair to some users. Baltrunas *et al.* [4] compare different aggregation approaches, and find that there is no clear winner, and the effectiveness of an approach depends on the group size and inner group similarity. Amer-Yahia *et al.* [3] go one step further, and argue that an item is a good recommendation for a group if the group members have small disagreements on the item, where the disagreement is defined as the difference among relevance of the item to different group members.

Recently, several model-based approaches have been proposed. Seko *et al.* [22] develop a content-based group recommendation method based on the assumption that the choice made by a group is influenced by item genres. However, this approach can be only applied to pre-defined groups, *e.g.,* couples, while in real-life, groups are often ad-hoc. Carvalho *et al.* [6] introduce game theory into group recommendation by treating a group event as a noncooper-

ative game among members, and transform the recommendation problem into finding the Nash equilibrium. However, this method cannot suggest a specific item, since the equilibrium contains a set of items.

The model-based approaches proposed in [8, 13, 26] adopt the same setting as ours. Ye *et al.* [26] assume that when selecting items, a group member will follow her friends' opinions. They propose a probabilistic generative model to produce group recommendations by aggregating the preferences of pairwise friends in the group, where one influences the other. However, the strong assumption of pairwise influence in a group may not be true, especially when the group is large. In addition, a group does not always consist of friends. Liu *et al.* [13] propose a topic model approach based on the assumption that the influential users will become the representatives of all groups to make item selections, irrespective of the group topics and the influential users' expertise. However, users' influences should be topic-dependent: a user may be influential in a group because of her expertise on the group's topics, but may not be in another group. Gorla *et al.* [8] assume that the recommendation score for an item depends on its relevance to each group member and its relevance to the group as a whole. They propose an information-matching based framework to make group recommendations. However, this framework has a high time complexity, which is $O(|U|^2|I|^2)$ for each target group, where $|U|$ and $|I|$ are the sizes of user and item sets, respectively. We implemented and ran this method, but it could not finish on our datasets after 5 days and we stopped it. It can only finish on very small data we tried.

In our experiments, we will report comparison results with the approaches in the three proposals [3,13,26]. In the existing studies, these approaches have not been empirically compared with each other.

Finally note that the "group recommendation" defined in [31] concentrate on personalized recommendation of event-based groups to a user, which is a totally different task.

3. CONSENSUS MODEL

We first define the group recommendation problem in Section 3.1, and then introduce the proposed COnsensus Model (COM) in Section 3.2. After that, the inference algorithm and the recommendation method are presented in Section 3.3 and 3.4, respectively. Finally, we present how to incorporate content information into the model in Section 3.5. All the notations used in this paper are listed in Table 1.

3.1 Problem Statement

Let U, I, G be the user, item and group sets, respectively. A *group* $g \in G$ consists of a set of users (group members) $\boldsymbol{u}_g = \{u_{g,1}, u_{g,2}, ..., u_{g,|g|}\}$, where $u_{g,i} \in U$, and $|g|$ is the size of the group, *i.e.,* the number of users in g. In addition, a group g is associated with an item $i_g \in I$, if i_g is selected by group g. We define a *group event* by $<\boldsymbol{u}_g, i_g>$, *i.e.,* the item selection event of the group members as a whole. For example, the members of a picnic group selecting a venue for picnic is a group event, and a family selecting a movie to watch is also a group event.

Then, given a target group g_t, the problem of group recommendation is defined as recommending a list of items that users in g_t may be interested in.

3.2 COnsensus Model for Group Recommendation

We model the generative process of a group event based on the following intuitions:

Table 1: Symbols

Symbol	Description
U, I, G	user set, item set, group set
K	the number of latent topics
N	the size of recommendation list
g, $|g|$	the group g, the size of group g
$\boldsymbol{u}_g$	the members in group g
i_g	the item selected by group g
u	user $u \in U$
z_j, c_j	the topic and switch of the user-item pair j
$\boldsymbol{i}^{(\cdot)}$	the set of items that are generated when $c = (\cdot)$
$\boldsymbol{\theta}_g$	distribution of topics specific to group g
ϕ_z^{ZU}	distribution of users specific to topic z
ϕ_z^{ZI}	distribution of items specific to topic z
ϕ_u^{UI}	distribution of items specific to user u
λ_u	the parameter of Bernoulli distribution specific to user u for sampling the binary switch c
$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\eta}$	Dirichlet prior vector for $\boldsymbol{\theta}$, ϕ^{ZU}, ϕ^{UI} and ϕ^{ZI}
$\boldsymbol{\gamma}$	Beta prior for λ, where $\boldsymbol{\gamma} = \{\gamma, \gamma_t\}$
$n_{g,z,\neg j}^{GZ}$	number of times topic z is assigned to group g, excluding the jth user-item pair
$n_{z,u,\neg j}^{ZU}$	number of times user u is drawn from topic z, excluding the jth user-item pair
$n_{z,i,\neg j}^{ZI}$	number of times item i is drawn from topic z, excluding the jth user-item pair
$n_{u,i,\neg j}^{UI}$	number of times item i is drawn from user u, excluding the jth user-item pair
$n_{u,c,\neg j}^{UC}$	number of times switch c is drawn for user u, excluding the jth user-item pair

- **Intuition 1**: Each group is relevant to several topics with different degrees of match, *e.g.,* a picnic group is more relevant to the hiking and dining topics than to the body-building topic. The topics of a group attract users to join the group.

- **Intuition 2**: When selecting an item, users in a group have two considerations. The first is topics, *i.e.,* a user tends to select the items that are related to the group topic, which attracted her to join the group. The second is users' personal considerations of content factors, such as the geographical distance for venue recommendation, cast lists of movies for movie recommendation, etc. Most of these factors are user-specific, and cannot be captured by topics. In addition, different users make different trade-offs between group topics and personal considerations of content factors: some users tend to select the items that match the group topics best, while some may treat the personal considerations more important.

- **Intuition 3**: Users behave differently when selecting items as members in a specific group and when selecting items as individuals. In a group, a user tends to match her preference to the topics of the group.

- **Intuition 4**: The preference of a group to a candidate item is determined by the preferences of the group members [3, 8]. In addition to this, we exploit the following new intuition: the influence of each member on the item selection of the group is topic-dependent.

Specifically, we use a multinomial distribution $\boldsymbol{\theta}_g$ over latent topics to model the topic preferences of group g (**Intuition 1**). In addition, each latent topic z has a multinomial distribution ϕ_z^{ZU} over user set, which represents the relevance of users to the topic z, and a multinomial distribution ϕ_z^{ZI} over item set, which represents the relevance of items to the topic z. Here $\phi_{z,i}^{ZI}$ reflects given a topic z,

how likely the item i is selected; $\phi_{z,u}^{ZU}$ reveals the appealing degree of topic z to the user u, or the user u's expertise on topic z. To model **Intuition 1** that users join a group because of different topics, for each member in group g, a latent topic z is sampled from its topic distribution θ_g, and then a user u is drawn according to ϕ_z^{ZU}.

A user in a group selects items either based on the group topics that attracted her to join the group, or her personal considerations of content factors (**Intuition 2**). We use a switch c to decide which one accounts for the item selection of a user, *i.e.,* if $c = 1$, the item is sampled based on the topic-specific multinomial distribution over items ϕ^{ZI}; if $c = 0$, the item is drawn from the user-specific multinomial distribution of items ϕ^{UI}. Since different users will make different trade-offs between group topics and personal considerations of content factors (**Intuition 2**), in our model, the switch c is drawn from a user-specific Bernoulli distribution with parameter λ_u. In other words, user u is influenced by group topics with probability λ_u, and is influenced by her personal considerations with probability $1 - \lambda_u$. Note that the Bernoulli distribution for λ_u has a Beta prior $\gamma = \{\gamma, \gamma_t\}$.

Next we illustrate the model using an example. Suppose a picnic group is more relevant to both hiking and dining topics than body-building topic. The three topics are sampled from the topic distribution of the picnic group, which attract users u_1, u_2 and u_3, respectively. Then, these three users determine which venue to visit based on the group topics and their personal considerations of content factors such as traveling distance. Suppose u_1 does not mind traveling, and the topic "hiking" has a more significant influence to his selection. Then she may select a distant venue that matches the hiking topic best. Thus, the switch c for u_1 is more likely to be 1. u_2 and u_3 will also make trade-offs between group topics and their personal considerations to select the venue for picnic.

Different from COM, previous topic model based approaches [13, 26] assume that when selecting items, a group member only considers her own topic preference. The assumption may not hold, because users in a group may behave differently as group members from that when they make choices as individuals (**Intuition 3**). For example, suppose u_1 is interested in both hiking and movie topics. In previous approaches, u_1 may select a theater for the picnic group because of her interest in movie topic. In contrast, in our model, u_1 join the picnic group because of the hiking topic, and thus her selection will be related to hiking rather than movie.

In summary, the generative process of a collection of group events is as follows:

- For each topic z_k, $k = 1, ..., K$
    - Draw $\phi_k^{ZU} \sim Dirichlet(\beta)$;
    - Draw $\phi_k^{ZI} \sim Dirichlet(\eta)$;
- For each user u_v, $v = 1, ..., |U|$
    - Draw $\phi_v^{UI} \sim Dirichlet(\rho)$;
    - Draw $\lambda_v \sim Beta(\gamma)$;
- For each group g
    - Draw $\theta_g \sim Dirichlet(\alpha)$;
    - For each group member
        * Draw $z \sim Multinomial(\theta_g)$;
        * Draw $u \sim Multinomial(\phi_z^{ZU})$;
        * Draw switch $c \sim Bernoulli(\lambda_u)$;
        * If $c = 0$
            · Draw $i \sim Multinomial(\phi_u^{UI})$;
        * If $c = 1$
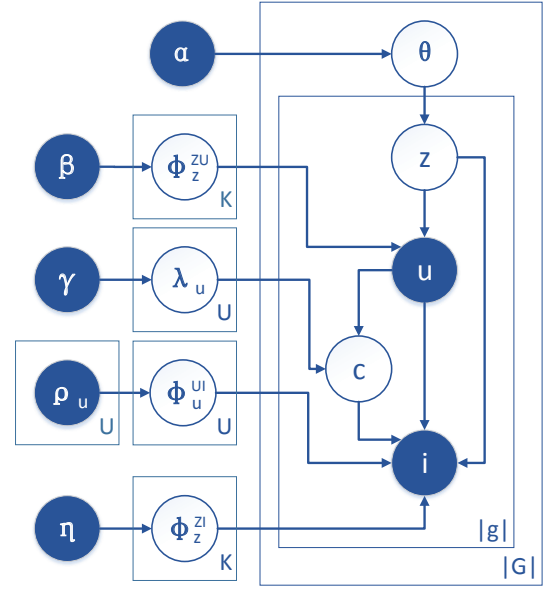   · Draw $i \sim Multinomial(\phi_z^{ZI})$;

Figure 2: The Graphical Model of COM

The graphical model is shown in Figure 2.

Note that different users in a group will sample different items in the model, which is in accordance with our experience: users may have different preferences over items, and thus are likely to make different choices. In fact, the item selection of a group is often made by two steps: group members express their own opinions on item selections first, and then these selections are weighted and a consensus is reached. As to be detailed in Section 3.4, we propose a recommendation method that can aggregate the selections of group members based on their topic-dependent influences, and produce a single recommendation for the target group.

We remark that the aforementioned generative process is also applicable to the groups with pre-defined members, since these groups also have topic distributions. Consider some students plan to form a club group and the topics of the group are dining, hiking, etc. The group is formed because its topics attract the members. If someone is not interested in any of the group topics, she will not join the group. Thus, the generative process of the club can also be explained by the proposed model, where the group members are sampled from the students.

3.3 Parameter Estimation

The total likelihood of the group event corpus is:

$$
\begin{aligned}
& P(\mathbf{z}, \mathbf{u}, \mathbf{c}, \mathbf{i} | \alpha, \beta, \rho, \eta, \gamma) \\
= {} & \int P(c|\lambda) P(\lambda|\gamma, \gamma_t) d\lambda \cdot \int P(z|\theta) P(\theta|\alpha) d\theta \cdot \\
& \int P(u|z, \phi^{ZU}) P(\phi^{ZU}|\beta) d\phi^{ZU} \cdot \\
& \int \int P(i|u, z, c, \phi^{UI}, \phi^{ZI}) P(\phi^{UI}|\rho) P(\phi^{ZI}|\eta) d\phi^{UI} d\phi^{ZI} \quad (1)
\end{aligned}
$$

We employ collapsed Gibbs sampling to obtain samples of the hidden variable assignment, and to estimate the unknown parameters $\{\lambda, \phi^{ZU}, \phi^{UI}, \phi^{ZI}\}$. For ease of presentation, we define a user u together with the item i selected by u as a user-item pair $j = (u, i)$, where the user of j is $u \in U$, and the item of j is $i \in I$.

166

Since there are two latent variables in the model, namely z and c, we employ two-step Gibbs sampling method. We first sample topics z_j for all user-item pairs j, and then sample switches c_j for all j. For each latent variable (*e.g.*, z_j), a Gibbs sampling method computes the full conditional probability for the assignment of the variable conditioned on all the other assignments (*e.g.*, $z_{\neg j}$). However, it is challenging to get the full conditional probability because of the complex interdependencies between user u, topic z, switch c and item i: u is sampled based on z, which influences the sampling of c, while i is sampled based on either z or u depending on c.

To solve this problem, we separate the items generated based on topics, and the items generated based on users' personal considerations of content factors. Then, the last part of Equation 1 becomes:

$$\int\int P(i|u,z,c,\phi^{UI},\phi^{ZI})P(\phi^{UI}|\rho)P(\phi^{ZI}|\eta)d\phi^{UI}d\phi^{ZI}$$
$$= \int P(i^{(0)}|u,c,\phi^{UI})P(\phi^{UI}|\rho)d\phi^{UI} \cdot$$
$$\int P(i^{(1)}|z,c,\phi^{ZI})P(\phi^{ZI}|\eta)d\phi^{ZI} \qquad (2)$$

where $i^{(0)}$ is the set of items that are sampled based on users' personal considerations of content factors (*i.e.*, $c = 0$), and $i^{(1)}$ is the set of items that are sampled based on topics (*i.e.*, $c = 1$).

Based on the new equation of total likelihood, we can derive the full conditional distribution of topic z_j and switch c_j assignments for each user-item pair j. If the item of j is drawn based on topics, *i.e.*, $c_j = 1$, we sample z_j according to the following probability:

$$P(z_j = k|z_{\neg j}, u, i^{(1)})$$
$$= \frac{\int P(z|\theta)P(\theta|\alpha)d\theta}{\int P(z_{\neg j}|\theta)P(\theta|\alpha)d\theta} \cdot \frac{\int P(u|z,\phi^{ZU})P(\phi^{ZU}|\beta)d\phi^{ZU}}{\int P(u|z_{\neg j},\phi^{ZU})P(\phi^{ZU}|\beta)d\phi^{ZU}} \cdot$$
$$\frac{\int P(i^{(1)}|z,c,\phi^{ZI})P(\phi^{ZI}|\eta)d\phi^{ZI}}{\int P(i^{(1)}|z_{\neg j},c,\phi^{ZI})P(\phi^{ZI}|\eta)d\phi^{ZI}}$$
$$\propto \frac{n_{g_j,k,\neg j}^{GZ} + \alpha_k}{\sum_{k'\in Z}(n_{g_j,k',\neg j}^{GZ} + \alpha_{k'})} \cdot \frac{n_{k,u,\neg j}^{ZU} + \beta_u}{\sum_{u'\in U}(n_{k,u',\neg j}^{ZU} + \beta_{u'})} \cdot \frac{n_{k,i,\neg j}^{ZI} + \eta_i}{\sum_{i'\in I}(n_{k,i',\neg j}^{ZI} + \eta_{i'})} \quad (3)$$

where g_j is the group of j. If the item of j is drawn based on user's personal considerations of content factors, *i.e.*, $c_j = 0$, we have:

$$P(z_j = k|z_{\neg j}, u, i^{(0)})$$
$$\propto \frac{n_{g_j,k,\neg j}^{GZ} + \alpha_k}{\sum_{k'\in Z}(n_{g_j,k',\neg j}^{GZ} + \alpha_{k'})} \cdot \frac{n_{k,u,\neg j}^{ZU} + \beta_u}{\sum_{u'\in U}(n_{k,u',\neg j}^{ZU} + \beta_{u'})} \qquad (4)$$

After sampling topics for all user-item pairs, we draw a switch c_j for each j according to the following posterior probability. When $c_j = 1$, we have:

$$P(c_j = 1|c_{\neg j}, z, u, i)$$
$$= \frac{\int P(c|\lambda)P(\lambda|\gamma,\gamma_t)d\lambda}{\int P(c_{\neg j}|\lambda)P(\lambda|\gamma,\gamma_t)d\lambda} \cdot \frac{\int P(i^{(0)}|u,c,\phi^{UI})P(\phi^{UI}|\rho)d\phi^{UI}}{\int P(i^{(0)}|u,c_{\neg j},\phi^{UI})P(\phi^{UI}|\rho)d\phi^{UI}} \cdot$$
$$\frac{\int P(i^{(1)}|z,c,\phi^{ZI})P(\phi^{ZI}|\eta)d\phi^{ZI}}{\int P(i^{(1)}|z,c_{\neg j},\phi^{ZI})P(\phi^{ZI}|\eta)d\phi^{ZI}} \qquad (5)$$

Note that since $c_j = 1$, the second term in the right hand side of Equation 5 is 1. Thus, we cancel this part, and get:

$$P(c_j = 1|c_{\neg j}, z, u, i)$$
$$\propto \frac{n_{u,(1),\neg j}^{UC} + \gamma}{n_{u,(0),\neg j}^{UC} + n_{u,(1),\neg j}^{UC} + \gamma + \gamma_t} \cdot \frac{n_{z_j,i,\neg j}^{ZI} + \eta_i}{\sum_{i'\in I}(n_{z_j,i',\neg j}^{ZI} + \eta_{i'})} \qquad (6)$$

Similarly, we calculate the sampling probability for $c_j = 0$:

$$P(c_j = 0|c_{\neg j}, z, u, i)$$
$$\propto \frac{n_{u,(0),\neg j}^{UC} + \gamma_t}{n_{u,(0),\neg j}^{UC} + n_{u,(1),\neg j}^{UC} + \gamma + \gamma_t} \cdot \frac{n_{u,i,\neg j}^{UI} + \rho_i}{\sum_{i'\in I}(n_{u,i',\neg j}^{UI} + \rho_{i'})} \qquad (7)$$

After sampling a sufficient number of iterations, we calculate the parameters ϕ^{ZU}, ϕ^{UI}, ϕ^{ZI} and λ as follows:

$$\widehat{\phi}_{z,u}^{ZU} = \widehat{P}(u|z) = \frac{n_{z,u}^{ZU} + \beta_u}{\sum_{u'\in U}(n_{z,u'}^{ZU} + \beta_{u'})} \qquad (8)$$

$$\widehat{\phi}_{u,i}^{UI} = \widehat{P}(i|u) = \frac{n_{u,i}^{UI} + \rho_i}{\sum_{i'\in I}(n_{u,i'}^{UI} + \rho_{i'})} \qquad (9)$$

$$\widehat{\phi}_{z,i}^{ZI} = \widehat{P}(i|z) = \frac{n_{z,i}^{ZI} + \eta_i}{\sum_{i'\in I}(n_{z,i'}^{ZI} + \eta_{i'})} \qquad (10)$$

$$\widehat{\lambda}_u = \widehat{P}(c = 1|u) = \frac{n_{u,(1)}^{UC} + \gamma}{n_{u,(0)}^{UC} + n_{u,(1)}^{UC} + \gamma + \gamma_t} \qquad (11)$$

3.4 Recommendation

When making recommendations for a target group g_t, we first discover its topic distribution based on the group members u_{g_t}. The distribution, denoted by θ_{g_t}, can be learnt by performing Gibbs sampling on u_{g_t} according to the following equation:

$$P(z_j = k|z_{\neg j}, u_j = v, u_{\neg j}) \propto \widehat{\phi}_{k,v}^{ZU}(n_{g_t,k,\neg j}^{GZ} + \alpha_k) \qquad (12)$$

Since the recommendations should match the topic distribution θ_{g_t}, based on the generative model, we define the recommendation score for candidate item i as follows:

$$P(i|u_{g_t}, \theta_{g_t}) \triangleq \prod_{u\in u_{g_t}}\sum_{z\in Z}\theta_{g_t,z} \cdot \widehat{\phi}_{z,u}^{ZU}(\widehat{\lambda}_u \cdot \widehat{\phi}_{z,i}^{ZI} + (1 - \widehat{\lambda}_u) \cdot \widehat{\phi}_{u,i}^{UI}) \quad (13)$$

Equation 13 embeds **Intuition 4** (when selecting items, different users in a group have different influence scores, and the influence scores are topic dependent) as follows: if the topic z is more relevant to group g_t, and a user u is an expert in z, then u will be more influential in item selection. Recall that the expertise of a user u on topic z is modeled by $\phi_{z,u}^{ZU}$. In Equation 13, $\theta_{g_t,z} \cdot \phi_{z,u}^{ZU}$ is the influence score of user u in group g_t for a given topic z, and $\lambda_u \cdot \phi_{z,i}^{ZI} + (1 - \lambda_u) \cdot \phi_{u,i}^{UI}$ is u's preference to a candidate item i given topic z. We margin out the topics, and get the overall preference of u to i. Then the preferences of all members to i are multiplied as the group preference to i. The rationale is three-fold: 1) the preference of a group to an item depends on the preferences of all individuals; 2) ranking an item based on the product of preferences is equal to the geometric mean of these individuals' preferences. Compared with the traditional strategies that calculate the arithmetic mean of preferences (averaging) or concentrate on the smallest preference (least-misery), the aggregated preference score by geometric mean is less sensitive to extreme values; 3) this definition matches the proposed model well.

3.5 Incorporation of Content Information

The Dirichlet prior ρ to ϕ_u^{UI} allows us to incorporate different content information into the model. We illustrate the incorporation using two recommendation tasks, namely, venue recommendation and movie recommendation.

Venue recommendation for groups: People often visit venues together with others for shopping, dining, etc. Venue recommendation for a group aims to recommend the venues that the group members are interested in. For venue recommendation, geographical distance is an important factor to consider [27, 30]. Previous studies reported that users tend to visit nearby venues, and the willingness of visiting a venue decreases with the increase of distance from their current locations. Here, we adopt a power-law function of distance to model the willingness of a user moving from one venue to another as [29] does. More specifically, the willingness of a user to visit a d-km far away venue is defined by Equation 14.

$$\text{wi}(d) = \omega \cdot d^\kappa \qquad (14)$$

where ω and κ are parameters of the power law function, which can be learned by maximum likelihood estimation.

Then, given a user u, the set of venues that she has visited I_u, we calculate $P(i|I_u)$ for each candidate venue i according to the geographical distance, and use this value for $\rho_{u,i}$. Based on the Bayes rule, $P(i|I_u)$ is calculated as follows:

$$\rho_{u,i} = P(i|I_u) \propto \quad P(i)P(I_u|i)$$
$$= \quad P(i) \prod_{i' \in I_u} P(i'|i) \qquad (15)$$

where $P(i'|i)$ is proportional to the willingness value in Equation 14, in which d is the distance between venues i' and i. □

Movie recommendation for groups: When selecting a movie to watch, a user may consider several factors, such as genre, cast, etc. We take the cast as an example to illustrate how to exploit content information. Intuitively, users tend to watch the movies stared by their favorite actors or actresses. We incorporate user u's cast-based considerations to a movie i by modifying the prior $\rho_{u,i}$ as follows:

$$\rho_{u,i} \propto \sum_{s \in S_i} P(s|u) \qquad (16)$$

where s is a movie star, and S_i is the cast list of movie i. $P(s|u)$ is estimated based on the occurrences of s in u's watching history. □

Note that the $|U| \times |I|$ dimensional matrix ϕ^{UI} requires a large amount of space, in which each value of $\phi_{u,i}^{UI}$ is determined by both the count $n_{u,i}^{UI}$ and the prior ρ_i (Equation 9). Since for each user u, the values of most items in ϕ^{UI} are 0 ($n_{u,i}^{UI} = 0$ and $\rho_{u,i} = 0$), we can use sparse matrix to store ϕ^{UI} to reduce the space complexity.

4. EXPERIMENTS

We first introduce the setup of the experiments in Section 4.1, and then present the experimental results in Section 4.2, in which we compare the recommendation accuracy of our model with five baselines on four datasets. After that, we analyze which factor influences group members' choices more significantly in Section 4.3. In the end, we show some sample topics discovered by the proposed model to examine their semantics in Section 4.4.

4.1 Experimental Setup

4.1.1 Datasets

Four real-world datasets are used in our experiments. The first dataset is used in previous work [12], which is collected from Plan-

Table 2: Statistics of the Datasets

Dataset	Plancast	Jiepang	MovieLens -Simi	MovieLens -Rand
#Users	41,705	28,888	891	3,689
#Group Events	13,885	23,621	3,000	3,000
#Items	8,016	9,746	441	1,518
Avg. Group Size	20.30	4.68	5	5
Avg. #Items for a Group	1.00	1.01	14.97	3.73
Avg. #Records for a User	6.76	3.83	16.83	4.06
Avg. #Friends for a User	43.29	13.00	N/A	N/A
Avg. #Records for an Item	1.73	2.42	100.06	7.37

cast[1], an event-based social network (EBSN). In Plancast, a user can follow others' calendars, and join different events. An event involves a group of members, and is held at a venue. A venue is associated with a geographical coordinate. We treat an event as a group, where the users involved in the event are the group members, and the venue of the event is the item selected by them.

Second, we collect 45 million check-ins from Jiapang[2], a location-based social network (LBSN). As shown in Figure 1, LBSNs allow users to share their geographical information by check-ins, where a check-in has a user, time and venue, indicating the user visited the venue at that time. Each a venue in Jiepang is associated with its geographical coordinate. However, Jiepang does not contain explicit group information, and we extract implicit group check-ins as follows: we assume if a set of friends visit the same venue at the same time, they are the members of a group. Specifically, the set of individual check-ins made by friends within 0.5 hour is regarded as a group check-in. For both Jiepang and Plancast datasets, we aim to recommend venues for given groups.

The last two datasets are extracted from 1M MovieLens dataset[3] by following the approach in [4]. MovieLens allows users to rate the movies they have watched by stars ranging from 0 to 5. Two kinds of groups are considered in the experiments: similar and random, denoted by MovieLens-Simi and MovieLens-Rand, respectively. Groups in MovieLens-Simi have larger inner similarities between members, while groups in MovieLens-Rand are randomly formed. The two datasets simulate two kinds of groups in real-life: the groups formed by people who have similar preferences, and the groups that happen to be formed by a set of people. For each dataset, we randomly select 3000 groups with 5-members. We also evaluated groups of size 3 and 8, and obtained similar results. The details for generating the datasets can be found in [4]. Given a group, if every member gives 4 stars or above to a movie, we assume that the movie is adopted by the group. We also collect the cast list of each movie from IMDB[4] as content information.

The information of the four datasets is shown in Table 2. For each dataset, we randomly mark off 20% of group events as the test set to evaluate the recommendation accuracy of different methods.

4.1.2 Evaluation Metrics

Following previous work [3,4,8,13], we evaluate the accuracy of different methods with three metrics, namely average *precision@N* (*Pre@N*), average *recall@N* (*Rec@N*) and normalized discounted

---

[1] http://plancast.com/
[2] http://jiepang.com/
[3] http://grouplens.org/datasets/movielens/
[4] http://www.imdb.com/interfaces/

cumulative gain (*nDCG*), where N is the number of recommendations. We consider three values of N (*i.e.,* 5, 10, 20), where 5 is the default value.

Precision@N is the fraction of the top N recommendations that are adopted by a group, while *recall@N* is the fraction of items adopted by a group (true items) that are contained in the top N recommendations. Formally, given a group, the *precision@N* and *recall@N* are calculated as:

$$precision@N = \frac{|\{\text{top } N \text{ recommendations}\} \cap \{\text{true items}\}|}{|\{\text{top } N \text{ recommendations}\}|} \quad (17)$$

$$recall@N = \frac{|\{\text{top } N \text{ recommendations}\} \cap \{\text{true items}\}|}{|\{\text{true items}\}|} \quad (18)$$

We average the *precision@N* and *recall@N* of all testing groups as the *Pre@N* and *Rec@N*, respectively. Note that the average number of true items of a group in Plancast and Jiepang is close to 1 (Table 2). In this case, *Pre@N* is proportional to *Rec@N*, since $|\{\text{top } N \text{ recommendations}\}|$ is N times greater than $|\{\text{true items}\}|$. Thus, to save space, we only report *Rec@5* for the two datasets.

nDCG measures how well a method can rank the true item higher in the recommendation list. It is calculated as follows:

$$DCG = rel_1 + \sum_{i=2}^{N} \frac{rel_i}{log_2(i)} \quad (19)$$

$$nDCG = \frac{DCG}{IDCG} \quad (20)$$

where $rel_i = 1$ if the ith item in the recommendation list is adopted by the group, and $rel_i = 0$ otherwise. *IDCG* is the maximum possible discounted cumulative gain (*DCG*) with optimal top N recommendations. We average the *nDCG* values of all groups as the final result. In the experiments, N is fixed at 10.

For all metrics, larger value indicates better recommendation performance.

4.1.3 Recommendation Methods

We evaluate 7 methods in our experiments, namely CF-AVG, CF-LM, CF-RD [3], SIG [26], PIT [13], and the proposed methods COMP and COM. To the best of our knowledge, these state-of-the-art group recommendation methods have not been compared with each other in previous work, and our evaluation is the first experimental studies on them.

User-based CF with averaging strategy (CF-AVG): Given a candidate item i, CF-AVG first estimates the recommendation score of each user in the target group by user-based CF, and then uses the average of these scores as the recommendation score for the group.
User-based CF with least-misery strategy (CF-LM): Given a candidate item i, CF-LM first estimates the recommendation score of each user in the target group by user-based CF, and then uses the smallest score as the recommendation score for the group.
User-based CF with relevance and disagreement (CF-RD) [3]: This model calculates the recommendation score for a candidate item i based on the relevance and disagreement of the group, where the relevance is calculated based on either CF-AVG or CF-LM, and the disagreement can be either the average difference of recommendation scores of pair-wise group members, or the variance of members' recommendation scores.
Social influence based group recommendation (SIG) [26]: SIG is a topic model based approach, which has been introduced in Section 2. Since the MovieLens-Simi and MovieLens-Rand datasets have no friendship information, we do not report the results of SIG for them.

Personal impact topic model (PIT) [13]: PIT model assumes that different users have different impact scores, and in a group, the user who has a larger impact score is more likely to be selected as the representative. Given a group of users u_{g_t}, PIT model first samples a representative user r from u_{g_t} based on users' impact scores, and then r selects a topic based on her topic preference, and finally the topic generates an item for the group.
COnsensus Model Plain (COMP): To make a fair comparison with these baselines which do not exploit users' considerations of content factors, we use a symmetric Dirichlet prior for ϕ^{UI} to disregard the effect of content information.
COnsensus Model (COM): The proposed model incorporated with users' considerations of content factors.

All baselines are evaluated under the optimal settings. For the hyperparameters in COMP and COM, we take fixed values ($\alpha = 50/K$, $\beta = \eta = 0.01$, $\gamma = \gamma_t = 0.5$ and $\rho = 0.01$ for COMP). The prior ρ in COM encodes the content-based knowledge, and needs to be set empirically. Previous work fixes its value as 0.01 [9], and thus the sum of the prior is $0.01 \times |I|$. In this paper, we normalize the value of ρ of each user to a fix value $0.01 \times p \times |I|$, where the parameter p is used to tune the confidence in the prior knowledge.

4.2 Experimental Results

Precision and Recall Under Different N

We first fix the number of topics K at 250, and vary the number of recommendations N. The *Pre@N* and *Rec@N* values on the four datasets are plotted in Figure 3. Please note that the MovieLens-Simi/Rand datasets do not contain social relations, and thus the baseline SIG cannot be applied to them.
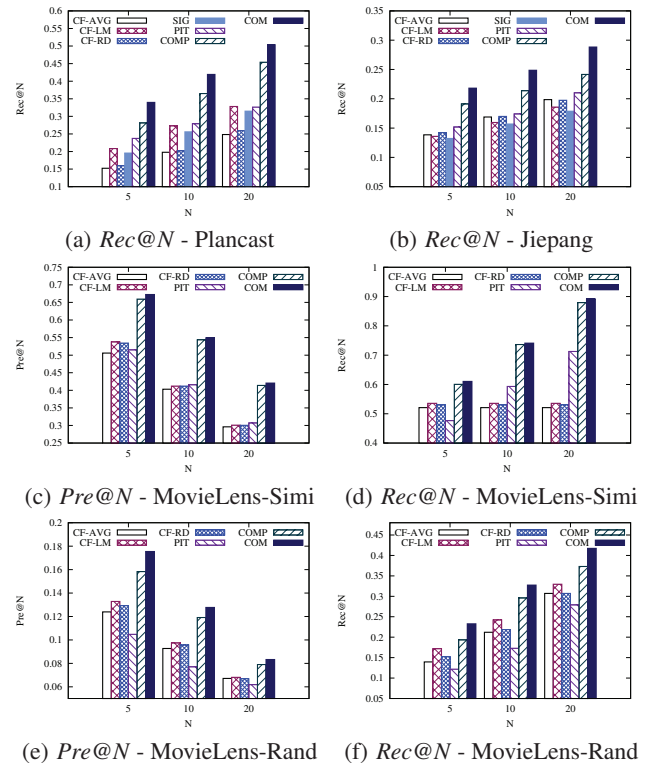

(a) *Rec@N* - Plancast   (b) *Rec@N* - Jiepang

(c) *Pre@N* - MovieLens-Simi   (d) *Rec@N* - MovieLens-Simi

(e) *Pre@N* - MovieLens-Rand   (f) *Rec@N* - MovieLens-Rand

Figure 3: *Pre&Rec* **under #recommendations made** (N)

From Figure 3(a) and 3(b), we can see that the CF-based approaches, namely, CF-AVG, CF-LM and CF-RD, do not perform well on the Plancast and Jiepang datasets. This is because the three methods exploit neither the difference of individuals, nor the inter-

actions among group members. They assume that users in a group make choices independently, and aggregate their choices for recommendations. The performance of SIG is not satisfactory, either. The reason is that SIG makes group recommendations based on the social relations between users in a group, and it requires tags of candidate items. However, in the groups of Plancast, only several or even none of the members are friends to each other, and neither of the two datasets has tag information. The lack of social relations and tags brings down its recommendation accuracy.

In contrast, PIT performs better than the CF based approaches on the Plancast and Jiepang datasets because PIT utilizes the interactions in a group by differentiating influences of users, and assumes that a user with a larger impact score will be influential in every group of the user. However, PIT ignores the fact that the influence of a user will be different across different topics. The performance of PIT on the two MovieLens datasets is not as good as on the Plancast and Jiepang datasets. This is because groups in the two datasets are loosely organized, and users select movie independently. Since no representative member exists to make item selections for a group, the basic assumption of PIT does not hold any more, which results in its bad recommendation accuracy.

Compared with the five baselines, our proposed method COMP always archives superior recommendation accuracy. For example, it outperforms CF-AVG, CF-LM, CF-RD, SIG and PIT by 84%, 34%, 76%, 43% and 19%, respectively, for $Rec@5$ on Plancast. The reasons are two-fold: on the one hand, COMP considers the behavior changes of users in a group; on the other hand, it estimates the topic-dependent influences of users in a group. Compared with COMP, COM further improves $Rec@5$ by more than 15% on Plancast, Jiepang and MovieLens-Rand, showing that COM is effective in incorporating the content information (geographical distance for Plancast and Jiepang, and cast list for MovieLens-Rand). The improvement on MovieLens-Simi is marginal, since its user-item selection matrix has a high density (about 4%). As a result, COMP, which only utilizes users' selection history, already achieves very good accuracy(*e.g., Rec@5* is 67.3%), and thus the value of relative improvement is small.

Precision and Recall Under Different K

We fix the number of recommendations at 5, and vary the number of topics K from 50 to 400. The $Pre@5$ and $Rec@5$ values on the four datasets are plotted in Figure 4. Since CF-AVG, CF-LM and CF-RD do not involve topics, their $Pre@5$ and $Rec@5$ values do not vary with K.

For the topic model based approaches, namely, SIG, PIT, COMP and COM, their $Pre@5$ and $Rec@5$ values do not change much with varying the number of topics. In addition, we notice that SIG performs worse than CF-AVG, CF-LM and CF-RD on Plancast, but better than these CF-based approaches on Jiepang when $K \geq 250$. This is because the group events of Jiepang are extracted based on friendships, and thus they fit well with the assumption of the SIG. PIT's $Rec@5$ value is the best among the baselines on the Plancast and Jiepang datasets, but is worse than that of CF-AVG, CF-LM and CF-RD on MovieLens-Simi and MovieLens-Rand. Potential reason is the generative process of groups in the MovieLens datasets is different from that of PIT model. Our proposed method COMP outperforms the best baselines by about 20% on the four datasets. After incorporating users' personal considerations of content factors, COM further improve the $Rec@5$ values by more than 15% on the Plancast, Jiepang and MovieLens-Rand datasets, demonstrating the effectiveness of the proposed model.

***nDCG* Under Different K**

Next, we vary the number of topics K, and examine the $nDCG$ results of different approaches to see how well they can rank the
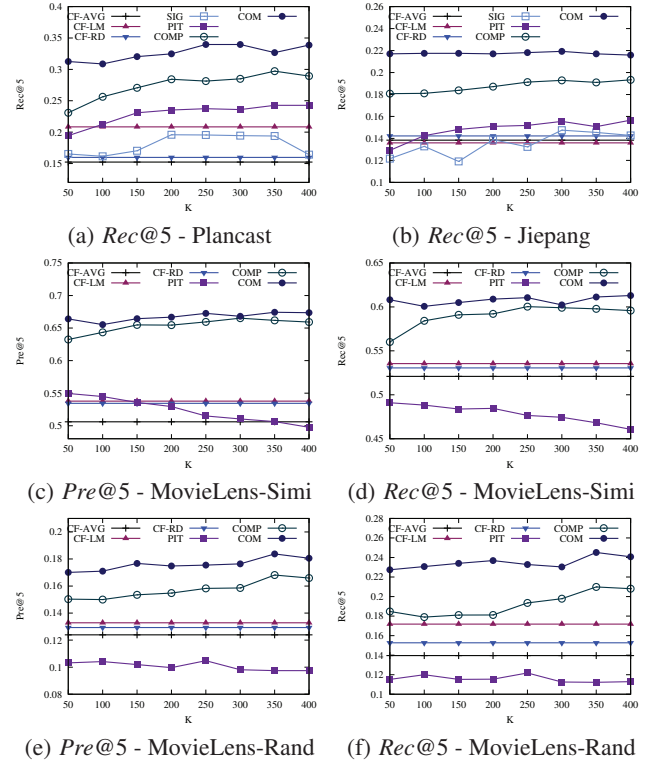

(a) $Rec@5$ - Plancast   (b) $Rec@5$ - Jiepang

(c) $Pre@5$ - MovieLens-Simi   (d) $Rec@5$ - MovieLens-Simi

(e) $Pre@5$ - MovieLens-Rand   (f) $Rec@5$ - MovieLens-Rand

Figure 4: $Pre@5$&$Rec@5$ **under #topics (K)**

true items higher. The results are plotted in Figure 5. We can see that the results display a similar trend with the previous experimental results based on $Pre@5$ and $Rec@5$: PIT performs the best among the baseline methods on Plancast and Jiepang, but the worst on MovieLens-Simi and MovieLens-Rand. However, our method COMP consistently outperforms the best baseline under different number of topics by more than 15% on the four datasets. COM achieves the best results, which are at least 16% greater than that of COMP on Plancast, Jiepang and MovieLens-Rand.

Effect of p

We next examine the effect of p on the recommendation accuracy of COM. Recall that after incorporating the users' personal considerations of content factors into the prior, we normalize the prior of each user to $p \cdot 0.01 \cdot |I|$. Parameter p is set to adjust the effect of the prior, *i.e.,* larger p implies that the distribution ϕ_u^{UI} is more influenced by the content information. We examine the recommendation accuracy of COM under different value of p ranging from 0.001 to 1000. The $Rec@5$ and $nDCG$ on four datasets are plotted in Figure 6. We can see that the recommendation performance remains relatively stable when varying the value of p. The $Pre@5$ follows a similar trend with $Rec@5$, and we omit it due to the space limitation.

Performance for Different Size of Groups

This set of experiments is to study the performance of each recommendation method for groups of different sizes. We group the Plancast groups into bins based on group size, and plot the $Rec@5$ and $nDCG$ curves of each method in Figure 7. The number of topics is fixed at 250. Due to the space limitation, the results on the other datasets are not given here. Figure 7 shows that the proposed methods COMP and COM outperform the baselines for groups of different sizes. Among the baselines, CF-AVG, CF-LM and CF-RD perform the worst, followed by SIG and PIT. Compared with that of other methods, the performance of CF-based approaches is
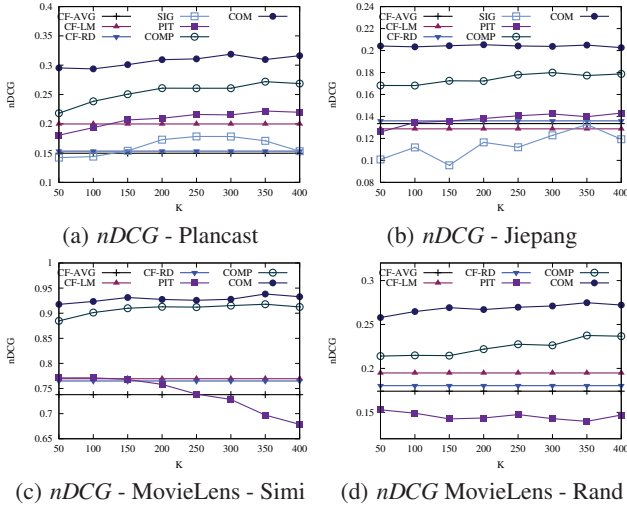
(a) nDCG - Plancast

(b) nDCG - Jiepang

(c) nDCG - MovieLens - Simi

(d) nDCG MovieLens - Rand

Figure 5: nDCG under #topics (K)

(a) Rec@5

(b) nDCG

Figure 7: Rec@5 **and** nDCG **for Different Size of Groups**

(a) λ_u vs K - Plancast

(b) λ_u vs K - Jiepang

(c) λ_u vs K - MovieLens-Simi

(d) λ_u vs K - MovieLens-Rand

Figure 8: Avg. λ_u **under #topics (K)**

better for groups of small size, because their group organizations are simple, and thus these simple aggregating strategies are good for making recommendations.

4.3 Weight of Topics in Item Selection

In this section, we study the weight of topics in users' item selections by investigating parameter λ_u, the probability that a user selects an item according to group topics.

We first study the effect of the number of topics K on the value of λ_u. Specifically, for each dataset, we plot the average λ_u of all users as a function of K. The curves of COMP and COM on the four datasets are shown in Figure 8. We can see that for COMP which does not exploit the content information, the average λ_u is almost not affected by the value of K, and its value is between 0.75 to 0.9. The results reveal that most of items are selected according to topics, but there is still a set of items that are selected based on users' personal considerations of content factors. Compared with the average λ_u of COMP, that of COM is much smaller, since additional content information is incorporated. The value of λ_u on Plancast (around 0.5) is larger than that on Jiepang (around 0.2) for COM, indicating that topics have larger weight in venue selections of Plancast users. In addition, the λ_u values of both COMP and COM on MovieLens-Simi are larger than those on MovieLens-Rand, because groups in MovieLens-Simi consist of people with high similarities, and thus topic is a very important consideration.

Then, we fix the number of topics at 250, and plot the distribution of λ_u of COM on the four datasets in Figure 9. We notice that on all datasets, the λ_u value of the majority of users is smaller than 0.4, showing that the personal considerations of content information is important for most of people. In addition, we see that the curve
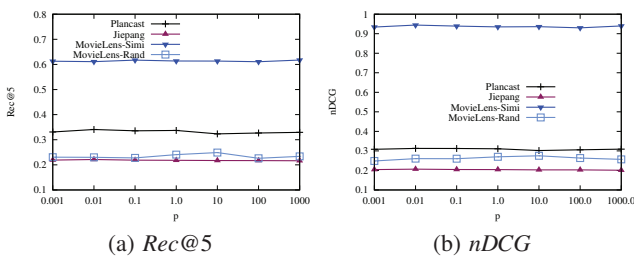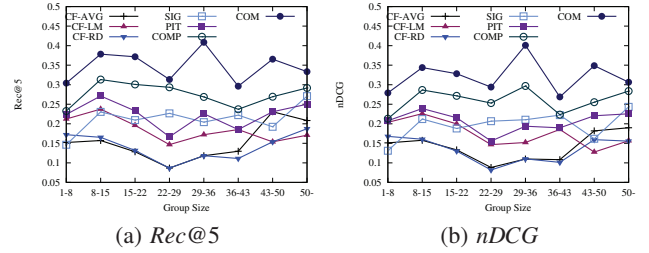

(a) Rec@5

(b) nDCG

Figure 6: Rec@5 **and** nDCG **under different** p
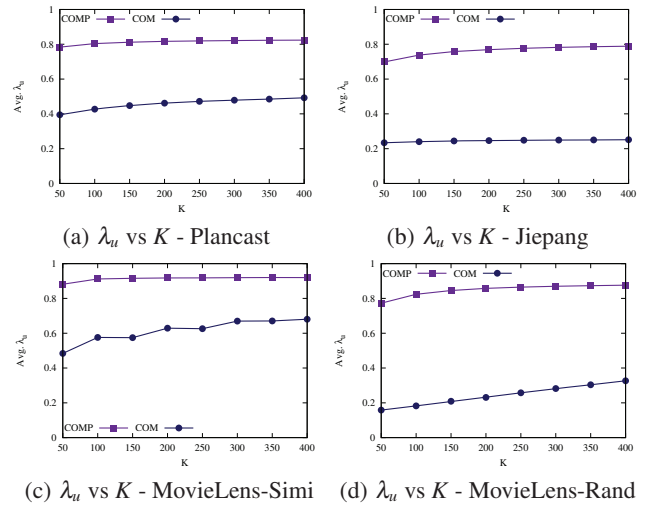
of Plancast has a long tail, indicating that a considerable portion of users treat topics important. In addition, λ_u on MovieLens-Simi reaches another peak at $\lambda_u = 1$, showing that in a group with high inner-similarity, a considerable portion of people select items based on topics.

4.4 Topic Analysis

We first investigate the venue distribution of each topic generated by COM on Plancast and Jiepang datasets, where the number of topics is set to 250. For each topic z, we rank the venues i based on $\phi_{z,i}^{ZI}$. The top 5 venues of 5 randomly selected topics on the two datasets are plotted in Figure 10. We observe that for each topic, the top ranked venues are close to each other. This is because topics are estimated based on users' group participation history. Since users tend to join groups held at their nearby venues due to the spatial constraint, the venues visited by each user fall in a small geographical region, and thus the top ranked venues of each topic are close to each other.

Then, we examine the movie distributions of topics of COM on the MovieLens-Rand dataset. Specifically, we set the number of topics at 50 for COM, and randomly select 5 topics. For each topic z, we rank the movies i based on the learnt $\phi_{z,i}^{ZI}$. The top 3 movies of the 5 topics are listed in Table 3. The name of each topic is generated from the top 10 movies' genres in IMDB by majority vote. We can find that the discovered topics are semantically meaningful.

5. CONCLUSION

Recommender systems have been studied for decades, but most of them are designed for individuals. How to make accurate recommendations for groups is still an open problem. In this paper, we propose a probabilistic model COM to simulate the generative
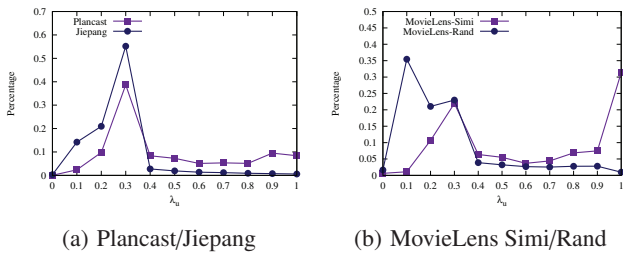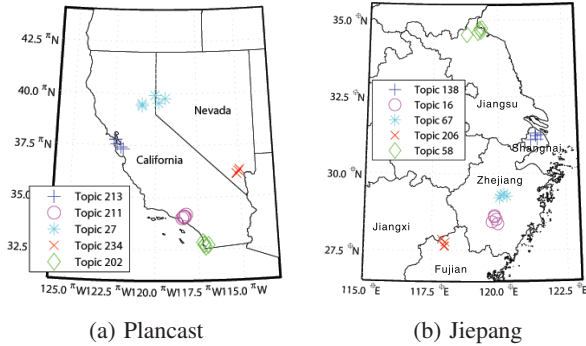
(a) Plancast/Jiepang (b) MovieLens Simi/Rand

Figure 9: Distribution of λ_u

(a) Plancast (b) Jiepang

Figure 10: Venue Distribution of Topics

Table 3: Representative Movies for COM Topics

Topic	Movies
Comedy	Big (1988), Romancing the Stone (1984), Four Weddings and a Funeral (1994)
Action	Raiders of the Lost Ark (1981) Star Wars: Episode V - The Empire Strikes Back (1980), Full Metal Jacket (1987)
Drama	Fight Club (1999), Trainspotting (1996), Magnolia (1999)
Thriller	Reservoir Dogs (1992), Gattaca (1997), Heavenly Creatures (1994)
Animation	Toy Story 2 (1999), Mulan (1998), Peter Pan (1953)

process of group events and make recommendations for a group of users. Since users' item selections are not only influenced by topics, but also by users' personal considerations of content factors, we incorporate the content information into the model. In addition, the proposed model considers the change of users' behaviors in a group from as individuals, and differentiates the influences of users in a group according to topics. Experimental results on four real-world datasets show that the proposed method outperforms five baselines significantly.

For the future work, it would be interesting to exploit social relations to make group recommendations, since the friendships in a group may influence the group's choices. In addition, several content factors are difficult to be incorporated into the our model, *e.g.,* time. We will investigate how to model such content factors.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng.*, 17(6):734–749, 2005.

[2] D. Agarwal and B.-C. Chen. flda: matrix factorization through latent dirichlet allocation. In *WSDM*, pages 91–100, 2010.

[3] S. Amer-Yahia, S. B. Roy, A. Chawla, G. Das, and C. Yu. Group recommendation: Semantics and efficiency. *PVLDB*, 2(1):754–765, 2009.

[4] L. Baltrunas, T. Makcinskas, and F. Ricci. Group recommendations with rank aggregation and collaborative filtering. In *RecSys*, pages 119–126, 2010.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[6] L. A. M. C. Carvalho and H. T. Macedo. Users' satisfaction in recommendation systems for groups: an approach based on noncooperative games. In *WWW (Companion Volume)*, pages 951–958, 2013.

[7] A. Crossen, J. Budzik, and K. J. Hammond. Flytrap: intelligent group music recommendation. In *IUI*, pages 184–185, 2002.

[8] J. Gorla, N. Lathia, S. Robertson, and J. Wang. Probabilistic group recommendation via information matching. In *WWW*, pages 495–504, 2013.

[9] T. L. Griffiths and M. Steyvers. Finding scientific topics. *Proceedings of the National academy of Sciences of the United States of America*, 101(Suppl 1):5228–5235, 2004.

[10] A. Jameson and B. Smyth. Recommendation to groups. In *The adaptive web*, pages 596–627. Springer, 2007.

[11] Y. Koren, R. M. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *IEEE Computer*, 42(8):30–37, 2009.

[12] 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 *KDD*, pages 1032–1040, 2012.

[13] X. Liu, Y. Tian, M. Ye, and W.-C. Lee. Exploring personal impact for group recommendation. In *CIKM*, pages 674–683, 2012.

[14] J. F. McCarthy. Pocket restaurant finder: A situated recommender systems for groups. In *Proceeding of Workshop on Mobile Ad-Hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems*, 2002.

[15] J. F. McCarthy and T. D. Anagnost. Musicfx: An arbiter of group preferences for computer aupported collaborative workouts. In *CSCW*, pages 363–372, 1998.

[16] K. McCarthy, M. Salamó, L. Coyle, L. McGinty, B. Smyth, and P. Nixon. Cats: A synchronous approach to collaborative group recommendation. In *FLAIRS Conference*, volume 2006, pages 86–91, 2006.

[17] M. R. McLaughlin and J. L. Herlocker. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In *SIGIR*, pages 329–336, 2004.

[18] M. O'Connor, D. Cosley, J. A. Konstan, and J. Riedl. Polylens: A recommender system for groups of user. In *ECSCW*, pages 199–218, 2001.

[19] M. O'Connor and J. Herlocker. Clustering items for collaborative filtering. In *Proceedings of the ACM SIGIR workshop on recommender systems*, volume 128. UC Berkeley, 1999.

[20] S. Pizzutilo, B. De Carolis, G. Cozzolongo, and F. Ambruoso. Group modeling in a public space: methods, techniques, experiences. In *Proceedings of the 5th WSEAS International Conference on Applied Informatics and Communications*, pages 175–180, 2005.

[21] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *WWW*, pages 285–295, 2001.

[22] S. Seko, T. Yagi, M. Motegi, and S. yo Muto. Group recommendation using feature space representing behavioral tendency and power balance among members. In *RecSys*, pages 101–108, 2011.

[23] X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. *Adv. Artificial Intelligence*, 2009, 2009.

[24] L. H. Ungar and D. P. Foster. Clustering methods for collaborative filtering. In *AAAI Workshop on Recommendation Systems*, number 1, 1998.

[25] C. Wang and D. M. Blei. Collaborative topic modeling for recommending scientific articles. In *KDD*, pages 448–456, 2011.

[26] M. Ye, X. Liu, and W.-C. Lee. Exploring social influence for recommendation: a generative model approach. In *SIGIR*, pages 671–680, 2012.

[27] M. Ye, P. Yin, W.-C. Lee, and D. L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*, pages 325–334, 2011.

[28] Z. Yu, X. Zhou, Y. Hao, and J. Gu. Tv program recommendation for multiple viewers based on user profile merging. *User Model. User-Adapt. Interact.*, 16(1):63–82, 2006.

[29] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann. Time-aware point-of-interest recommendation. In *SIGIR*, pages 363–372, 2013.

[30] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann. Who, where, when and what: discover spatio-temporal topics for twitter users. In *KDD*, pages 605–613, 2013.

[31] W. Zhang, J. Wang, and W. Feng. Combining latent factor model with location features for event-based group recommendation. In *KDD*, pages 910–918, 2013.