GIST: A Generative Model with Individual and Subgroup-based Topics For Group Recommendation

Ke Ji[a,b], Zhenxiang Chen[a,b], Runyuan Sun[a,b], Kun Ma[a,b], Zhongjie Yuan[c],
Guandong Xu[d]

[a]*School of Information Science and Engineering, University of Jinan, Jinan 250022, China*
[b]*Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China*
[c]*Shandong Normal University, Jinan 250014, China*
[d]*Data Science and Machine Intelligence Lab, Advanced Analytics Institute, University of Technology Sydney, Australia*

Abstract

In this paper, a Topic-based probabilistic model named *GIST* is proposed to infer group activities, and make group recommendations. Compared with existing individual-based aggregation methods, it not only considers individual members' interest, but also consider some subgroups' interest. Intuition might seem that when a group of users want to take part in an activity, not every group member is decisive, instead, more likely the subgroups of members having close relationships lead to the final activity decision. That motivates our study on jointly considering individual members' choices and subgroups' choices for group recommendations. Based on this, our model uses two kinds of unshared topics to model individual members' interest and subgroups' interest separately, and then make final recommendations according to the choices from the two aspects with a weight-based scheme. Moreover, the link information in the graph topology of the groups can be used to optimize the weights of our model. The experimental results on real-life data show that the recommendation accuracy is significantly improved by *GIST* comparing with the state-of-the-art methods.

Keywords: Group recommendation, Group activity, Decision making, Topic model, Recommender systems.

1. Introduction

Due to the ability of automatically generating a list of personalized items for users based on their past behaviors, recommender systems Adomavicius & Tuzhilin (2005) have become an indispensable technique of solving information

Email addresses: `ise_jik@ujn.edu.cn` (Ke Ji), `ise_chenzx@ujn.edu.cn` (Zhenxiang Chen), `ise_sunry@ujn.edu.cn` (Runyuan Sun), `ise_mak@ujn.edu.cn` (Kun Ma), `lovejie1189@vip.qq.com` (Zhongjie Yuan), `Guandong.Xu@uts.edu.au` (Guandong Xu)
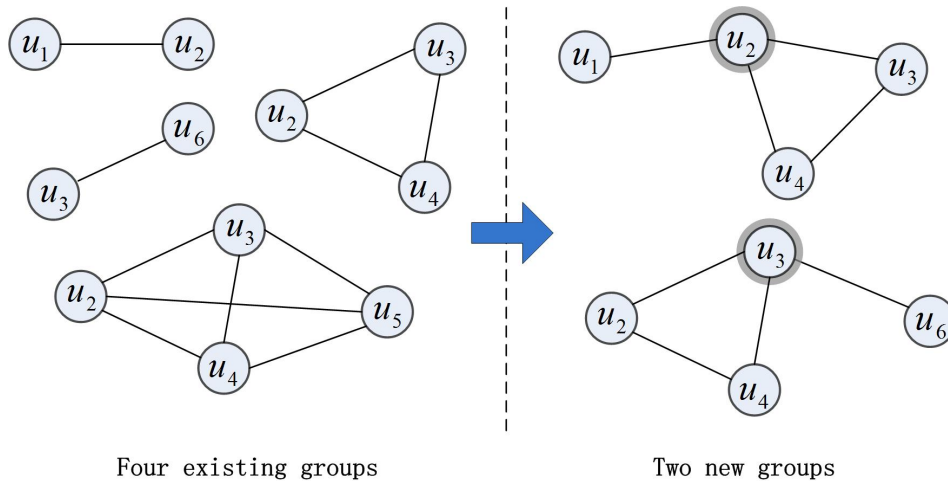
Figure 1: A toy example of constructing the graph topology of the new group ($\{u_1, u_2, u_3, u_4\}$ or $\{u_2, u_3, u_4, u_6\}$) based on the known link information of some existing groups ($\{u_1, u_2\}$, $\{u_3, u_6\}$, $\{u_2, u_3, u_4\}$ and $\{u_2, u_3, u_4, u_5\}$)

overload problem in the websites (*e.g.*, Amazon, Netflix, etc). Traditional recommendation approaches are designed to model individual activities, however people are not just isolated entities, have social friendship, and so would often take part in activities with others (*e.g.*, seeing movies with girlfriends, having dinners with families and going hiking with friends), which raise the need for special recommendation technique to solve group recommendation Jameson & Smyth (2007); Masthoff (2011). The change of the target object from an individual user to a group of users brings new challenges to recommender systems. Further, as more and more group activity records are available in some websites (*e.g.*, Facebook, Meetup and Foursquare), much research has been done to study group recommendation.

The most difficult problem facing group recommendation is how to define the group interest from a group whose members have different interest. A general solution to this problem is to use aggregation strategy that makes a trade-off on the choices of individual members for recommending items to a group. Previous aggregation methods can be classified into two approaches Berkovsky & Freyne (2010): preference aggregation Yu et al. (2006); McCarthy & Anagnost (1998) and score aggregation Dwork et al. (2001); Masthoff (2004). Both the two methods are memory-based approach, where the aggregation function can be average or least misery. Recently, in order to better identify personal impact in group activity, some model-based methods Ye et al. (2012); Salehi-Abari & Boutilier (2015); Gao et al. (2016); Liu et al. (2012); Yuan et al. (2014) incorporating content information, social influence or latent topics are proposed to bring great accuracy improvement over the memory-based methods.

However, very little of the above-mentioned work systematically analyzes why a group comes into being and what role each member plays in group

activities. It is easy to understand when a group of users take part in an activity together, they would build relationships with one another. Figure 1 shows the examples for the graph topology of some existing groups ($\{u_1, u_2\}$, $\{u_3, u_6\}$, $\{u_2, u_3, u_4\}$ and $\{u_2, u_3, u_4, u_5\}$) and new groups ($\{u_1, u_2, u_3, u_4\}$ and $\{u_2, u_3, u_4, u_6\}$). In each group example, the circular nodes represent users, and the edges between the nodes represent their relationships. The graph of each new group is built based on the links in the existing groups. By observing the topology structure of groups in the figure, we can find some useful information that has not been exploited by previous work:

- In group $\{u_1, u_2, u_3, u_4\}$, user u_2 has the maximum number 3 of links with other members, therefore u_2 might have greater influence in the group. But in another group $\{u_2, u_3, u_4, u_6\}$, u_2 has 2 links with with other members, less than 3 links for u_3, therefore u_3 might have greater influence than u_2 in the group. This phenomenon really shows that each user has different influence in different groups.

- In groups $\{u_1, u_2, u_3, u_4\}$ and $\{u_2, u_3, u_4, u_6\}$, there is a common subgroup $\{u_2, u_3, u_4\}$ that is a strongly connected component in the graph topologies of the two groups. Meanwhile, the rest user u_1 or u_6 has only 1 link with the members of the subgroup $\{u_2, u_3, u_4\}$. They are the equivalent of the comparatively isolated member compared to the subgroup $\{u_2, u_3, u_4\}$, and can be seen as a newcomer in the group. Intuitively, the subgroups of members having close relationships might play the more significant leading role in activity decision rather than the newcomers do.

In addition to graph topology, we provide a possibility to dig deeper to the group interest. We use "out to dinner" as an example to demonstrate interest difference between individual member and group. When there is only a single user, he might prefer to a simplistic food, like a hamburger or a bowl of noodles. When there are two users, they might chose a fancy restaurant and order two signature dishes. When there are more users together, they might go to order some dishes and one soup at a bigger restaurant or have a buffet as a leisure party. So these cases indicate that, compared to individual members' interest, subgroups' interest is probably more inherent to group interest. The discovery of this information inspires us to better model group activities by incorporating subgroups' choices.

Based on the above observations, we propose a generative model with individual and subgroup-based topics ($GIST$) for group recommendation. Given a group of users, we first build a graph topology corresponding to the group from existing groups, and extract strongly connected subgroups from the graph. Then, depending on the link information of the graph, we assign weights to all members and subgroups. Note that each member and subgroup are assumed to have a multinomial distribution over unshare latent topics, and the activity choices of each member or subgroup is influenced by how their topics interest them. The final activity decision is made by aggregating the members' choices and subgroups' choices: group activity would tend to individual mem-

bers' choices if the graph is sparse, tend to subgroups' choices otherwise. We evaluate our model by experiments on the dataset from Meetup. The results demonstrate the performance advantage of incorporating subgroups' choices into group recommendation.

The rest of this paper is organized as follows. We first give an overview of related work in Section 2. Then, we describe our model in detail in Section 3. We report the experimental results and analysis in Section 4. Finally, we provide a conclusion in Section 5.

2. Related Work

In this section, we review some related work, including (1) recommender system, (2) group recommendation, and (3) probabilistic topic model.

2.1. Recommender Systems

Due to the powerful capability of solving information overload, recommender systems have attracted a lot of attention in the past decade. Many excellent recommendation techniques Adomavicius & Tuzhilin (2005) have been proposed. Among them, *collaborative filtering* (CF) Su & Khoshgoftaar (2009) is the most popular and successful approach to building recommender systems. Typically, CF methods take the matrix of users' ratings on items as the input and predict a user's missing interest in an item by ratings from other users or items. Two primary approaches to CF are memory-based and model-based. The difference between them is that memory-based approaches Jin et al. (2004); Marlin (2004); Wang et al. (2006) predict the missing ratings based on ratings from similar users or items which can be found using the similarity metrics (PCC, VCC Breese et al. (1998)), whereas model-based approaches Salakhutdinov & Mnih (2008); Srebro & Jaakkola (2003); Hofmann (2004) explore the training data to train a model which can make fast prediction after all the parameters are learnt. Many experimental results and conclusions argue that model-based approaches can lead to somewhat more accurate results, while memory-based approaches may have some practical advantages. Recent development of context-aware and social-aware CF methods achieves better prediction and solves the problems of sparsity, cold-start, etc. However, the above methods are proposed to generate individual recommendations, cannot make effective recommendations for groups.

2.2. Group Recommendation

Group activity refers to an essential scenario of people's social life, where a set of users take part in an activity together. For example, tourism, movie, restaurant and KTV are more suited to group rather than to individual user. An increasing number of group activities in the websites has stimulated much recent research work on group recommendation. A common idea behind group recommendation is to aggregate all members' choices. Existing aggregation strategies

fall into two major categories: preference aggregation Yu et al. (2006); Mc-Carthy & Anagnost (1998) and score aggregation Dwork et al. (2001); Masthoff (2004). The first merges each member's profile into a profile for the group, and makes recommendations based on the group profile. The second generates each member's prediction, and merges all members' predictions into a prediction. The two strategies are memory-based approaches. Recently, some model-based methods have been proposed to achieve better prediction over memory-based approaches. At first, the assume of "social influence" that a member will trust his friends' behavior in the item selection is used to model group activity decision: Ye et al. (2012) proposed a social influenced selection (SIS) model to mine the social influence between linked friends and the personal preference; Salehi-Abari & Boutilier (2015) proposed a preference-oriented social network model which uses the friends' homogeneity to improve the case that the preferences of some group members are unobserved; Gao et al. (2016) proposed a new Bayesian latent factor model that embeds social group influence into matrix factorization framework. But a group does not always contains the explicit friendship between the members, limiting their range of applicability. Liu et al. (2012) introduced the notion of personal impact to a group decision, and proposed a probabilistic model, in which each member's personal preference is modeled as a mixture of latent topics, and the item selection considers each user's personal preferences and personal influences. Yuan et al. (2014) gave the intuition that the group members who are expert in topics relevant to the content factors of the group are usually more influential. Based on Yuan et al. (2014), they proposed a probabilistic model which incorporates both users' section history and personal considerations of content factors. Beyond that, some experimental work Basu Roy et al. (2015); Carvalho & Macedo (2013); Delic et al. (2016); Guy et al. (2016) analyzed preliminarily the composition and activity of the groups. The above-mentioned studies are designed only from the point of individual members' interest, without considering the formation of the groups and interest difference between individual member interest and group.

2.3. Probabilistic Topic Model

Probabilistic topic model Blei (2012) is a statistical method for discovering the latent topic structure that occurs in a large collection of documents. This technique can extract and infer the relations of the contextual-usage of words and has been applied to many tasks (e.g., classification, novelty detection, summarization, similarity judgment, etc.). The common topic models used in practice have singular value decomposition (SVD) Brand, probabilistic matrix factorization (PMF) Salakhutdinov & Mnih (2008), probabilistic Latent Semantic Analysis ($pLSA$) Hofmann (1999), latent Dirichlet allocation (LDA) Blei et al. (2003) and author topic (AT) Rosen-Zvi et al. (2004). SVD and PMF-based methods have the similar purpose that uses the product of several low-rank matrices to approximate the original matrixed data. $pLSA$, LDA and AT are all generative statistical models of taking each word in a document as a sample from a mixture model, where the mixture components are multinomial random variables that can be viewed as representations of latent topics. By
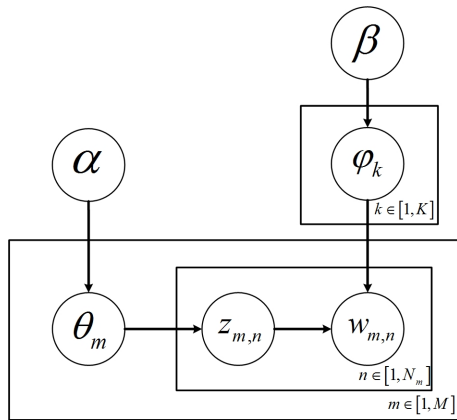
Figure 2: Latent Dirichlet allocation-*LDA*

the latent topics, they allow sets of observed variables to be explained by hidden variables. In particular, *pLSA* models the probability of each co-occurrence (document-word) as a mixture of conditionally independent multinomial distributions: a latent topic must be chosen conditionally to a document in advance and a word is then generated according to the topic. In essence, *LDA* is a generalisation of the *pLSA*, can be seen as the Bayesian version of *pLSA*. As shown in Figure 2, it has an assumption that Dirichlet priors α and β are put over two multinomial distributions: document-topic distribution θ and topic-word distribution φ. Then, for the simple case of LDA, the generation process of a word w in a document m is as follows: vectors θ_i and φ_j are sampled from Dirichlet distributions $\theta_i \sim Dir\left(\theta_i \mid \alpha\right)$ and $\varphi_j \sim Dir\left(\varphi_j \mid \beta\right)$, and afterwards topic z is sampled from m's θ_m: $z \sim Mult\left(z \mid \theta_m\right)$, then word w is sampled from z's φ_z: $w \sim Mult\left(w \mid \varphi_z\right)$. *AT* is also proposed to characterize documents, but in particular is aimed at co-authored document text. It extends *LDA* to discover the relationships between authors, documents, and words by allowing the mixture weights for different topics to be determined by the authors of the document. In addition to being proposed for documents and words, Topic modeling algorithms can be adapted to many kinds of data in the same way. They have already been exploited for recommendation systems in different domains, like tourism recommendations Hao et al. (2010), paper recommendations Wang & Blei (2011) and location recommendations Kurashima et al. (2013). Among other applications, they have been used to find patterns in genetic data Lee et al. (2015), images Zhang et al. (2015), and social networks Rajani et al. (2014).

Recently, *LDA* and *AT* models have been successfully applied to develop group recommendation models Liu et al. (2012); Yuan et al. (2014) with varying degrees of performance improvement. In this paper, we also explore Probabilistic topic model for the possible improvement of group recommendation. Our model is designed based on *LDA* and *AT* models.

3. Our Model

We first formalize group recommendation problem and define notations in Section 3.1. Then, our *GIST* model is presented in detail in Section 3.2. Next, we provide the solution to parameter inference during the optimization process. Finally, we give the method of how to incorporate the topology information to optimize our model in Section 3.4.

3.1. Problem Setting

Let $\mathcal{U}$, $\mathcal{I}$ and $\mathcal{G}$ represent the user set, item set and group activity set, respectively. Let G_g be a group consisting the members $U_g = \left\{u_{g,1}, u_{g,2}, \ldots, u_{g,|ug|}\right\}$, where $u_{g,*}$ is one element of $\mathcal{U}$ and $|ug|$ is the number of the members. We define an undirected link graph (ν_g, ε_g) for group G_g, where the node set ν_g represents the members and the edge set ε_g represents the links between the members. Let $\mathcal{S}$ represent the subgroup set, and $S_g = \left\{s_{g,1}, s_{g,2}, \ldots, s_{g,|sg|}\right\}$ represent a set of strongly connected subgroups extracted from G_g's graph (ν_g, ε_g), where $S_{g,*}$ is one element of $\mathcal{S}$ and $|sg|$ is the number of the subgroups. Then, G_g can be written as $G_g = \{U_g, S_g\}$. With the items selected by the groups, a collection of group activities is $\mathcal{G} = \left\langle \vec{G}, \vec{i} \right\rangle = \{\langle G_1, i_1 \rangle, \langle G_2, i_2 \rangle, \ldots, \langle G_N, i_N \rangle\}$, representing N group-item pairs, where each pair $\langle G_g, i_g \rangle$ denotes that item $i_g \in \mathcal{I}$ is selected by some users U_g whose link graph (ν_g, ε_g) contains the strongly connected subgroups S_g. Beyond the basic definitions, Table 1 shows some other symbol definitions, and the following section will provide further details in their presence.

With the notations, given group G_g, our goal is to aggregate the choices of individual members U_g and the choices of subgroups S_g for group recommendations.

3.2. GIST Model for Group Recommendation

As mentioned in the introduction, our approach to group recommendation relies on having the graph topology of the members. Here, we use the undirected graph to represent them, and realize the obtaining of the graphs with the help of a free Java graph library-JGraphT[1].

In addition to consider individual members' interest, the subgroup information is incorporated to more accurately understand how an activity decision is reached in a group of users. Our idea is to fuse individual members' interest and subgroups' interest into a joint model, and the final group recommendation of mixing between the two aspects is based on the link information of the graph topology. In order to achieve the idea, we propose a generative model with individual and subgroup-based topics which emploies a conventional bag-of-word representation with Dirichlet-Multinomial model.

---

[1] JGraphT is designed to be simple and type-safe via Java generics, and provides mathematical graph-theory objects and algorithms.

Table 1: Symbols

Symbol	Description
$\mathcal{U}, \mathcal{I}, \mathcal{G}$	user set, item set and group activity set
$\mathcal{S}$	the set of strongly connected subgroups in $\mathcal{G}$
k	the number of latent topics
G_g, N	a group , the number of group activities $\mathcal{G}$
U_g	the member set in group G_g
(ν_g, ε_g)	G_g's link graph
S_g	the strongly connected subgroups in graph (ν_g, ε_g)
i_g	the item selected by group G_g
z_g^U, z_g^U	individual topic and subgroup topic
c_g	the switch of the individual-subgroup pair
$u_{(\cdot)} z_{(\cdot)} s_{(\cdot)} i_{(\cdot)}$	the set of users, topics, subgroups and items generated when $c_g = (\cdot)$
π_g^U	distribution of the individual influence for U_g
π_g^S	distribution of the subgroup influence for S_g
$\theta_u^{U Z_0}$	topic distribution for user u
$\theta_s^{S Z_1}$	topic distribution for strongly connected group s
$\varphi_z^{Z_0 I}$	item distribution for individual topic z
$\varphi_z^{Z_1 I}$	item distribution for subgroup topic z
λ_g	Bernoulli distribution for group g
α, ρ, β, η	Dirichlet prior for $\theta_u^{U Z_0}$, $\theta_s^{S Z_1}$ $\varphi_z^{Z_0 I}$ and $\varphi_z^{Z_1 I}$
γ	Beta prior for λ_g
$\vec{n}_{u,z,\neg g}^{U Z_0}$	user u's count for topic z, excluding the group-item pair $\langle G_g, i_g \rangle$
$\vec{n}_{z,i,\neg g}^{Z_0 I}$	individual topic z's count for item i, excluding the group-item pair $\langle G_g, i_g \rangle$
$\vec{n}_{s,z,\neg g}^{S Z_1}$	subgroup s's count for topic z, excluding the group-item pair $\langle G_g, i_g \rangle$
$\vec{n}_{z,i,\neg g}^{Z_1 I}$	subgroup topic z's count for item i, excluding the group-item pair n
$\vec{n}_{(0),\neg g}^{C}$, $\vec{n}_{(1),\neg g}^{C}$	the collection $\mathcal{G}$'s switch counts, excluding the group-item pair $\langle G_g, i_g \rangle$

In the following, we first present our approaches to individual members' interest and subgroups' interest.

Individual member. For having social relationship, users have influence on other members of the group, and so every individual member's choice would influence the activity decision within the group. As mentioned in Section 1, when a user is in different groups, he will behave different influence, for instance if very familiar with the members, he is more likely to influence the group activity, on the contrary if as a stranger, he may have limited influence. Well, we suppose that the more link a user has with the other members in the graph, the more chance he has to decide the group activity. Given the graph topology of the group members $U_g = \{u_{g,1}, u_{g,2}, \ldots, u_{g,|ug|}\}$, we define a weight vector $\pi_g^U = \left(\pi_{g,1}^U, \pi_{g,2}^U, \ldots, \pi_{g,|ug|}^U\right)$ over U_g that has the normalization constraint $\sum_{i=1}^{|ug|} \pi_{g,i}^U = 1$. The weight $\pi_{g,u}^U$ is assigned based on the proportion of user u's link number $N_{g,u}^U$ with other members of U_g:

$$\pi_{g,u}^U = \frac{N_{g,u}^U + 1}{\sum_{i=1}^{|ug|} \left(N_{g,i}^U + 1\right)} \tag{1}$$

where we add "1" to $N_{g,u}^U$ in order to avoid ignoring the interest of members that contain no link. And $N_{g,u}^U$ is represented as:

$$N_{g,u}^U = \sum_{u' \in U_g} N\left(u, u'\right) \tag{2}$$

where $N\left(u, u'\right)$ counts the link number between u and u' in existing groups. The left part of Figure 3 shows an example of the calculation.

Subgroup. As mentioned in Section 1, an individual member's interest may be different from the group interest, and some subgroups' interest seems more consistent with the group interest. When extracting the subgroups from the graph topology, we need only focus on the strongly connected components without having to enumerate all the possible subgroups due to the reason that only the subgroups whose members are already familiar with each other and engage in various of activities together would play a leading role in the group. Note that strongly connected components must satisfy such "fully connected" condition: every member is reachable from every other member Newman et al. (2006); Newman (2010). In this paper, we consider two subgroup extraction strategies:

1. n-member: this strategy finds the strongly connected subgroups which contain the fixed n number of members.
2. Maximal-member: this strategy finds the strongly connected subgroups which contain the maximum number of members.

Given the strongly connected subgroups $S_g = \{s_{g,1}, s_{g,2}, \ldots, s_{g,|sg|}\}$ of a group G_g, we define a weight vector $\pi_g^S = \left(\pi_{g,1}^S, \pi_{g,2}^S, \ldots, \pi_{g,|sg|}^S\right)$ over S_g that has the
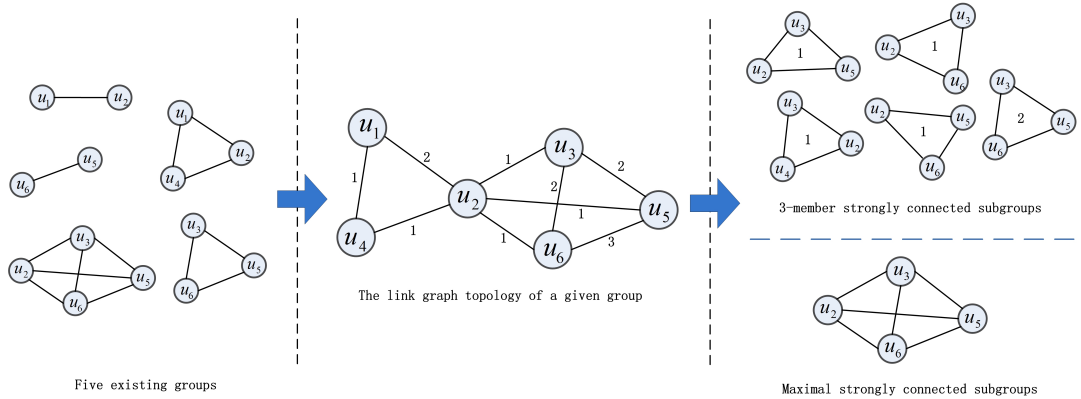
Figure 3: A toy example of constructing the graph topology of the new group $\{u_1, u_2, u_3, u_4, u_5, u_6\}$, and then extract the strongly connected subgroups from the graph using n-member or Maximal-member strategy.

normalization constraint $\sum_{i=1}^{|sg|} \pi_{g,i}^S = 1$. The weight $\pi_{g,u}^U$ is assigned based on the proportion of occurrence number $N_{g,s}^S$ of subgroup $S_{g,s}$ in existing group activities:

$$\pi_{g,s}^S = \frac{N_{g,s}^S}{\sum_{i=1}^{|sg|} N_{g,i}^S} \tag{3}$$

The right part of Figure 3 shows an example of the calculation.

Note that by importing users' historical relationship information into JGraphT, we can conveniently use its APIs to extract the members and subgroups. However, how to use the APIs is out of the scope of this paper, please visit this site [2] to get their detailed usage about JGraphT.

Next, we present our approach to modeling how an item is probabilistically selected by group G_g with the known weight π_g^U of members U_g and weight π_g^S of subgroups S_g. On the whole, the probability of an item selected by a group reflects the mixture of individual members' interest and subgroups' interest. More specifically, the interest of each member or subgroup is modeled as mixture of several latent topics, and each topic has a multinomial distribution over item set $\mathcal{I}$. Note that individual members' interest and subgroups' interest are from two different sources - the equivalent of two different corpora, so we use two kinds of unshared topics Z_0 and Z_1 to model the two interest separately. A switch $c_g \in \{0, 1\}$ drawn from a Bernoulli distribution λ_g with a Beta prior $\gamma = \{\gamma_0, \gamma_1\}$ is defined to decide which interest (individual or subgroup) dominates the item selection: if $c_g = 0$, the item is sampled based on individual members' interest, otherwise the item is sampled based on subgroups' interest. For individual members' interest, the selection process consists of three steps: 1) a user u is drawn from the set U_g according to weight π_g^U; 2) probabilistically
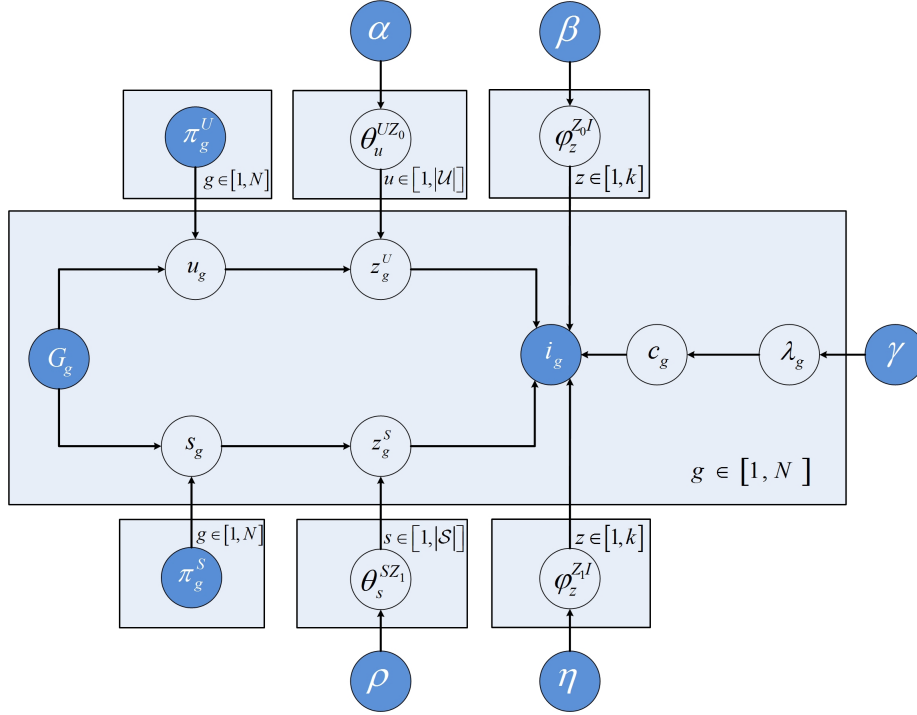
_____

[2]http://jgrapht.org/

10

Figure 4: Our proposed model - *GIST*

select a topic $z_g^U = z$ from z_0 according to user u's topic distribution $\theta_u^{UZ_0}$; 3) probabilistically select an item from $\mathcal{I}$ according to topic z's item distribution $\varphi_z^{Z_0 I}$. And subgroups' interest has the same selection process: 1) a subgroup s is drawn from S_g according to weight π_g^S; 2) probabilistically select a topic $z_g^S = z$ from z_1 according to subgroup s's topic distribution $\theta_s^{SZ_1}$; 3) probabilistically select an item from $\mathcal{I}$ according to topic z's item distribution $\varphi_z^{Z_1 I}$.

Now we give the complete graphical representation of our *GIST* model in Figure 4, where blue circles indicate observable variables and white circles indicate hidden variables. In summary, the generative process for all the collection of group activities $\mathcal{G}$ is as follows:

- Draw switch distribution $\lambda_g \sim Beta\,(\gamma)$.

- For each user $u \in \mathcal{U}$, draw topic distribution $\theta_u^{UZ_0} \sim Dirichlet\,(\alpha)$

- For each strongly connected group $s \in \mathcal{S}$, draw topic distribution $\theta_s^{SZ_1} \sim Dirichlet\,(\rho)$

- For each user topic $z \in Z_0$, draw item distribution $\varphi_z^{Z_0 I} \sim Dirichlet\,(\beta)$

- For each subgroup topic $z \in Z_1$, draw item distribution $\varphi_z^{Z_1 I} \sim Dirichlet\,(\eta)$

11

- For each group G_g

  Compute π_g^U and π_g^S based the graph topology of group G_g

    1. Draw $c_g \sim Multinomial\,(\lambda_g)$
    2. if $c_g = 0$,
        (a) Draw user $u \sim Multinomial\,\left(\pi_g^U\right)$ from g's members $U_g = \left\{u_{g,1}, u_{g,2}, \ldots, u_{g,|ug|}\right\}$
        (b) Draw topic $z \sim Multinomial\,\left(\theta_u^{UZ_0}\right)$ from individual topic Z_0.
        (c) Draw item $i_g \sim Multinomial\,\left(\varphi_z^{Z_0 I}\right)$
    3. if $c_g = 1$,
        (a) Draw subgroup $s \sim Multinomial\,\left(\pi_g^U\right)$ from g's subgroups $S_g = \left\{s_{g,1}, s_{g,2}, \ldots, s_{g,|sg|}\right\}$.
        (b) Draw topic $z \sim Multinomial\,\left(\theta_s^{SZ_1}\right)$ from subgroup topic Z_1.
        (c) Draw item $i_g \sim Multinomial\,\left(\varphi_z^{Z_1 I}\right)$

3.3. Parameter Inference and Recommendation

The likelihood probability of the collection of $\mathcal{G}$ for our GIST model is

$$\Pr\left(\vec{u}, \vec{s}, \vec{z}, \vec{i}, \vec{c} \mid \vec{G}, \pi^U, \pi^S, \alpha, \beta, \rho, \eta, \gamma\right)$$
$$= \Pr\left(\vec{u}, \vec{s}, \vec{z}, \vec{i} \mid \vec{G}, \vec{c}, \pi^U, \pi^S, \alpha, \beta, \rho, \eta\right)\Pr\left(\vec{c} \mid \gamma\right) \qquad (4)$$

Because $\left\{\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}, \vec{\lambda}\right\}$ are unknown, they are removed by marginalization in Equation 4. We define the set of groups $\vec{G}_{(0)}$, the set of users $\vec{u}_{(0)}$ and the set of items $\vec{i}_{(0)}$ that sampled based on individual members' interest (i,e,. $c_g = 0$). The rest is the set of groups $\vec{G}_{(1)}$, the set of subgroups $\vec{s}_{(1)}$ and the set of items $\vec{i}_{(1)}$ that sampled based on subgroups' interest (i,e,. $c_g = 1$). By dividing the collection into the sets $\left\{\vec{u}_{(0)}, \vec{z}_{(0)}, \vec{i}_{(0)}\right\}$ and $\left\{\vec{s}_{(1)}, \vec{z}_{(1)}, \vec{i}_{(1)}\right\}$ with the number N_0 and N_1, Equation 4 can be rewritten as follows:

$$\Pr\left(\vec{u}, \vec{s}, \vec{z}, \vec{i}, \vec{c} \mid \vec{G}, \pi^U, \pi^S, \alpha, \beta, \rho, \eta, \gamma\right)$$
$$= \Pr\left(\vec{u}_{(0)}, \vec{z}_{(0)}, \vec{i}_{(0)} \mid \vec{G}_{(0)}, \pi^U, \alpha, \beta\right)\Pr\left(\vec{s}_{(1)}, \vec{z}_{(1)}, \vec{i}_{(1)} \mid \vec{G}_{,(1)}, \pi^S, \rho, \eta\right)\Pr\left(\vec{c} \mid \gamma\right) \quad (5)$$

And the three terms in Equation 5 are:

$$\Pr\left(\vec{u}_{(0)}, \vec{u}_{(0)}, \vec{i}_{(0)} \mid \vec{G}_{(0)}, \pi^U, \alpha, \beta\right)$$
$$= \Pr\left(\vec{u}_{(0)} \mid \vec{G}_{(0)}, \pi^U\right)\Pr\left(\vec{z}_{(0)} \mid \vec{u}_{(0)}, \alpha\right)\Pr\left(\vec{i}_{(0)} \mid \vec{z}_{(0)}, \beta\right)$$
$$= \prod_{n=1}^{N_0} \pi_{g, \vec{u}_{(0),n}}^U \prod_{u=1}^{|\mathcal{U}|} \frac{\triangle\left(\vec{n}_{u,*}^{UZ_0} + \alpha\right)}{\triangle\left(\alpha\right)} \prod_{z=1}^{k} \frac{\triangle\left(\vec{n}_{z,*}^{Z_0 I} + \beta\right)}{\triangle\left(\beta\right)} \qquad (6)$$

$$\Pr\left(\vec{s}_{(1)}, \vec{z}_{(1)}, \vec{i}_{(1)} \mid \vec{G}_{(1)}, \pi^S, \rho, \eta\right)$$

$$= \Pr\left(\vec{s}_{(1)} \mid \vec{G}_{(1)}, \pi^S\right) \Pr\left(\vec{z}_{(1)} \mid \vec{s}_{(1)}, \rho\right) \Pr\left(\vec{i}_{(1)} \mid \vec{z}_{(1)}, \eta\right)$$

$$= \prod_{n=1}^{N_1} \pi^U_{g,\vec{s}_{(1),n}} \prod_{s=1}^{|\mathcal{S}|} \frac{\triangle\left(\vec{n}^{SZ_1}_{s,*} + \rho\right)}{\triangle\left(\vec{\rho}\right)} \prod_{z=1}^{k} \frac{\triangle\left(\vec{n}^{Z_1 I}_{z,*} + \eta\right)}{\triangle\left(\eta\right)} \tag{7}$$

$$\Pr\left(\vec{c} \mid \gamma\right) = \frac{\triangle\left(\vec{n}^C + \gamma\right)}{\triangle\left(\gamma\right)} \tag{8}$$

Here, we introduce some notion for Equation 6, 7 and 8: $\vec{n}^{UZ_0}_{u,*}$ denotes a vector of user u's topic counts, $\vec{n}^{Z_0 I}_{z,*}$ denotes a vector of user topic z's item counts, $\vec{n}^{SZ_1}_{s,*}$ denotes a vector of subgroup s's topic counts, $\vec{n}^{Z_1 I}_{z,*}$ denotes a vector of subgroup topic z's item counts and $\vec{n}^C$ denotes a vector of the whole collection $\mathcal{G}$'s switch counts. Because of conjugate prior-likelihood pair, the second term (joint probability of the user u' topic selection) is $\frac{\triangle\left(\vec{n}^{UZ_0}_{u,*}+\alpha\right)}{\triangle(\alpha)}$ with prior parameter α and Dirichlet delta function Heinrich (2008) $\triangle\left(\cdot\right)$, for instance, $\triangle\left(\alpha\right) = \prod_{z=1}^{k} \Gamma\left(\alpha_z\right) / \Gamma\left(\prod_{z=1}^{k} \alpha_z\right)$. We give the detailed derivation in Equation A.1 and A.2 of the appendix. The similar joint probability is also in other terms of Equation 6, 7 and 8.

Our goal is to learn the hidden variable assignment, and estimate the unknown parameters $\left\{\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}, \vec{\lambda}\right\}$. Although conjugate prior-likelihood is a relatively simple model, exact inference for the parameters is intractable. The general solution to this problem is to use approximate inference algorithm. In this paper, we use collapsed Gibbs sampling, a special case of Markov-chain Monte Carlo (MCMC) simulation, to optimize our model. In particular, Gibbs sampling with our model is a two-step approach, where each round of sampling consists of two step: first sample the switch variable c_g for each group activity, and then sample user u_g and individual topic z_g^U or subgroup s_g and subgroup topic z_g^S based on c_g.

Given a group-item pair $\langle G_g, i_g \rangle$, we begin with sampling the switch assignment c_g. The conditional probability for c_g conditioned on all the other switch $\vec{c}_{\neg g}$ is

$$\Pr\left(c_g = 1 \mid \vec{G}, \vec{u}, \vec{s}, \vec{z}, \vec{c}_{\neg g}, \vec{i}\right) \propto \frac{\Pr\left(\vec{c} \mid \gamma\right)}{\Pr\left(\vec{c}_{\neg g} \mid \gamma\right)} \propto \frac{\vec{n}^C_{(1),\neg g} + \gamma_0}{\vec{n}^C_{(1),\neg g} + \vec{n}^C_{(0),\neg g} + \gamma_0 + \gamma_1} \tag{9}$$

After sampling switch assignment for all group-item pairs, we draw user u_g and user topic or subgroup and subgroup topic for each pair. Note that if the item i_g is drawn based on individual members', *i.e.*, $c_g = 0$, we sample user u_g from U_g and individual topic z_g according to the following conditional probability conditioned on all the other $\vec{u}_{(0),\neg g}$ and $\vec{z}_{(0),\neg g}$:

$$\Pr\left(u_g = u, z_g^U = z \mid \vec{G}_{(0)}, \vec{z}_{(0),\neg g}, \vec{u}_{(0),\neg g}, \vec{i}_{(0)}\right)$$

$$\propto \quad \frac{\Pr\left(\vec{u}_{(0)} \mid \vec{G}_{(0)}\right)}{\Pr\left(\vec{u}_{(0)\neg g} \mid \vec{G}_{(0),\neg g}\right)} \cdot \frac{\Pr\left(\vec{z}_{(0)} \mid \vec{u}_{(0)}\right)}{\Pr\left(\vec{z}_{(0),\neg g} \mid \vec{u}_{(0),\neg g}\right)} \cdot \frac{\Pr\left(\vec{i}_{(0)} \mid \vec{z}_{(0)}\right)}{\Pr\left(\vec{i}_{(0),\neg g} \mid \vec{z}_{(0),\neg g}\right)}$$

$$\propto \quad \pi_{g,u}^U \cdot \frac{\vec{n}_{u,z,\neg g}^{UZ_0} + \alpha_z}{\sum_{z'} \vec{n}_{u,z',\neg g}^{UZ_0} + \alpha_{z'}} \cdot \frac{\vec{n}_{z,i,\neg g}^{Z_0 I} + \beta_i}{\sum_{i'} \vec{n}_{z,i',\neg g}^{Z_0 I} + \beta_{i'}} \tag{10}$$

Otherwise, if the item i_g is drawn based on subgroup choice, *i.e.*, $c_g = 1$, we sample subgroup s_g from S_g and subgroup topic z_g from according to the following conditional probability on all the other $\vec{s}_{(1),\neg g}$ and $\vec{z}_{(1),\neg g}$:

$$\Pr\left(s_g = s, z_g^S = z \mid \vec{G}_{(1)}, \vec{s}_{(1),\neg g}, \vec{z}_{(1),\neg g}, \vec{i}_{(1)}\right)$$

$$\propto \quad \frac{\Pr\left(\vec{s}_{(1)} \mid \vec{G}_{(1)}\right)}{\Pr\left(\vec{s}_{(1)\neg g} \mid \vec{G}_{(1),\neg g}\right)} \cdot \frac{\Pr\left(\vec{z}_{(1)} \mid \vec{s}_{(1)}\right)}{\Pr\left(\vec{z}_{(1),\neg g} \mid \vec{s}_{(1),\neg g}\right)} \cdot \frac{\Pr\left(\vec{i}_{(1)} \mid \vec{z}_{(1)}\right)}{\Pr\left(\vec{i}_{(1),\neg g} \mid \vec{z}_{(1),\neg g}\right)}$$

$$\propto \quad \pi_{g,s}^S \cdot \frac{\vec{n}_{s,z,\neg g}^{SZ_1} + \rho_z}{\sum_{z'} \vec{n}_{s,z',\neg g}^{SZ_1} + \rho_{z'}} \cdot \frac{\vec{n}_{z,i,\neg g}^{Z_1 I} + \eta_i}{\sum_{i'} \vec{n}_{z,i',\neg g}^{Z_1 I} + \eta_{i'}} \tag{11}$$

We first initialize the parameters in random way, and then take a sufficient number of repeatly sampling $\vec{u}, \vec{s}, \vec{z}, \vec{c}$ through a burn-in period to obtain a steady approximation. The detailed sampling algorithm is described in Algorithm 1. After the sampling process, the posterior estimates for $\left\{\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}, \vec{\lambda}\right\}$ can be computed as follows:

$$\widehat{\theta}_{u,z}^{UZ_0} = \frac{\vec{n}_{u,z}^{UZ_0} + \alpha_z}{\sum_{z' \in z_0}\left(\vec{n}_{u,z'}^{UZ_0} + \alpha_{z'}\right)} \tag{12}$$

$$\widehat{\theta}_{s,z}^{SZ_1} = \frac{\vec{n}_{s,z}^{SZ_1} + \rho_z}{\sum_{z' \in z_1}\left(\vec{n}_{s,z'}^{SZ} + \rho_{z'}\right)} \tag{13}$$

$$\widehat{\varphi}_{z,i}^{Z_0 I} = \frac{\vec{n}_{z,i}^{Z_0 I} + \beta_i}{\sum_{i' \in \mathcal{I}}\left(\vec{n}_{z,i'}^{Z_0 I} + \beta_{i'}\right)} \tag{14}$$

$$\widehat{\varphi}_{z,i}^{Z_1 I} = \frac{\vec{n}_{z,i}^{Z_1 I} + \eta_i}{\sum_{i' \in \mathcal{I}}\left(\vec{n}_{z,i'}^{Z_1 I} + \eta_{i'}\right)} \tag{15}$$

$$\widehat{\lambda}_g = \frac{\vec{n}_{(1)}^C + \gamma_0}{\vec{n}_{(1)}^C + \vec{n}_{(0)}^C + \gamma_0 + \gamma_1} \tag{16}$$

Algorithm 1: Two-step Gibbs sampling for GIST Model

Input: a collection of group activities $\mathcal{G} = \left\langle \vec{G}, \vec{i} \right\rangle$ given by the group G_g
for the items i_g. prior parameters α, β, η, ρ and γ. MAX_{iter} L.

Output: estimated parameters $\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}$ and $\vec{\lambda}$

1 Initialization:

2 for *all group-item pair $g \in [1, N]$* **do**

3   compute π_g^U and π_g^S based the graph topology information of group G_g;

4   Randomly sample switch c_g from $[0, 1]$;

5   Randomly sample u_g and z_g^U from $U_g = \left\{ u_{g,1}, u_{g,2}, \ldots, u_{g,|ug|} \right\}$ and user topic z_0;

6   Randomly sample s_g and z_g^S from $S_g = \left\{ s_{g,1}, s_{g,2}, \ldots, s_{g,|sg|} \right\}$ and subgroup topic z_1;

7   Update counters $\vec{n}^{UZ_0}$, $\vec{n}^{Z_0 I}$, $\vec{n}^{SZ_1}$, $\vec{n}^{Z_1 I}$ and $\vec{n}^C$;

8 end

9 Sampling period:

10 for *iter $\leftarrow 1, 2, \ldots L$* **do**

11   **for** *all group-item pair $g \in [1, N]$* **do**

12   Remove c_g from count $\vec{n}^C$;

13   sample switch $c_g \sim \Pr\left(c_g = c \mid \vec{G}, \vec{u}, \vec{s}, \vec{z}, \vec{c}_{\neg g}, \vec{i} \right)$;

14   Update counters $\vec{n}^C$;

15   **if** $c_g = 0$ **then**

16   Remove u_g and z_g^U from counts $\vec{n}^{UZ_0}$ and $\vec{n}^{Z_0 I}$;

17   sample u_g and z_g^U
$\sim \Pr\left(u_g = u, z_g^U = k \mid \vec{G}_{(0)}, \vec{z}_{(0), \neg g}, \vec{u}_{(0), \neg g}, \vec{i}_{(0)} \right)$;

18   use the new u_g and z_g^U to update counts $\vec{n}^{UZ_0}$ and $\vec{n}^{Z_0 I}$;

19   **else**

20   Remove s_g and z_g^S from counts $\vec{n}^{SZ_1}$ and $\vec{n}^{Z_1 I}$;

21   sample s_g and z_g^S
$\sim \Pr\left(s_g = s, z_g^S = k \mid \vec{G}_{(1)}, \vec{z}_{(1), \neg g}, \vec{s}_{(1), \neg g}, \vec{i}_{(1)} \right)$;

22   use the new u_g and z_g^U to update counts $\vec{n}^{SZ_1}$ and $\vec{n}^{Z_1 I}$;

23   **end**

24   **end**

25   **if** *converged or iter $= L$* **then**

26   read out $\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}$ and $\vec{\lambda}$ by Equation 12-16;

27   **end**

28 end

29 return $\theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}$ and $\vec{\lambda}$

When making recommendations for a target group of users $U_g = \{u_{g,1}, u_{g,2}, \ldots, u_{g,|ug|}\}$, we first construct the graph topology of the group based on existing group activities, and find strongly connected subgroups $S_g = \{s_{g,1}, s_{g,2}, \ldots, s_{g,|sg|}\}$ from the graph using the extraction strategy (*n-member* or *Maximal-member*). Then, depending on the link information of the graph, we assign weight π_g^U and π_g^S to users U_g and subgroups S_g. Therefore, following the generative process in Section 3.2, the probability that a group with U_g and S_g selects an item i is derived as Equation 17. We give the detailed derivation in Equation A.3 of the appendix.

$$\Pr\left(i \mid U_g, S_g\right) = \lambda_g \sum_{u \in U_g} \sum_{z \in Z_0} \pi_{g,u}^U \theta_{u,z}^{U Z_0} \varphi_{z,i}^{Z_0 I} + (1 - \lambda_g) \sum_{s \in S_g} \sum_{z \in Z_1} \pi_{g,s}^S \theta_{s,z}^{S Z_1} \varphi_{z,i}^{Z_1 I} \tag{17}$$

3.4. Incorporation of Topology Information

In our model, the parameter λ_g in Bernoulli distribution controls the switch c_g of deciding which interest (individual member or subgroup) may play a leading role in the activity decision.

When the graph is sparse, there are few links between the members, meaning that they are not familiar with each other and group activity should more tend to individual members' choices (i.e., λ_g should be a smaller value, tend to 0); on the contrary, when the graph is dense, there are some "complete" subgroups whose members are connected to every other member, meaning that they often spend time together and group activity should more tend to the subgroups' choices (i.e., λ_g should be a bigger value, tend to 1).

On above evidence, the link information in the graph is likely to help optimize λ_g. The Beta prior $\gamma = \{\gamma_0, \gamma_1\}$ for λ_g allows us to incorporate additional information into our model. In order to incorporate the link information with prior γ, we introduce the most widely used concept in graph theory - "density" Neetil & de Mendez (2012), which describes the linkage among the nodes in a graph. Given group G_g's graph (ν_g, ε_g), the density is

$$d_g = \frac{|\varepsilon_g|}{|\nu_g| \left(|\nu_g| - 1\right) / 2} \tag{18}$$

where $|\nu_g|$ and $|\varepsilon_g|$ are the number of the node set ν_g and edge set ε_g. Then, we define a new prior $\gamma_g = \{\gamma_{g,0}, \gamma_{g,1}\}$ by making $\gamma_{g,0}$ and $\gamma_{g,1}$ proportional to d_g and $1 - d_g$ in Equation 19.

$$\gamma_{g,0} \propto 1 - d_g, \ \gamma_{g,1} \propto d_g \tag{19}$$

In this way, the posterior estimation for λ_g in Equation 16 can be rewritten as Equation 20. This causes a benefit that when making recommendations for different groups, λ_g can be updated dynamically based on the change of $\gamma_{g,0}$ and $\gamma_{g,1}$.

$$\widehat{\lambda}_g = \frac{\vec{n}_{(1)}^C + \gamma_{g,0}}{\vec{n}_{(1)}^C + \vec{n}_{(0)}^C + \gamma_{g,0} + \gamma_{g,1}} \tag{20}$$

Table 2: Statistics of the meetup datasets

Meetup Group Tag		Meetup Event Venue	
Description	Value	Description	Value
# users	124,047	# users	329,590
# Groups	1,613	# Events	115,634
# Tags	2,118	# Venues	47,606

4. Experiments

4.1. Datasets

In this paper, we evaluate our model on a real-world Meetup dataset of Event-based Social Network (EBSN). As a new type of social network, Meetup EBSN is the world's largest network of local groups. It is a convenient online platform for people getting together to learn something, do something or share something, where more than 9,000 groups get together in local communities each day, each one with the goal of improving themselves or their communities. This dataset crawled from Meetup.com from Oct 2011 to Jan 2012 is provided by and can be downloaded from the site[3]. Following previous work, we also use the dataset to do the two types of group recommendation tasks: 1) as each meetup group can specify some tags to represent interest of the group, we may recommend some tags to groups based on existing group-tag pairs; 2) as meetup allows users to form social groups to event venues, we may recommend venues to groups based existing group-venue participations.

The statistics of the datasets are summarized in Table 2. Each dataset is divided into three parts (validation set, training set and test set): select 10% ratings randomly as the validation set, and partition the rest 90% ratings into training set and test set. We turn the parameters according to the experiments on the validation set, and compare the performance of all the baselines and our model according to the experiments on the test set.

4.2. Metrics

In our experiments, we use two popular ranking metrics-*Precision@N* and *Recall@N* to evaluate the models.

Precision@N is the fraction of the top-N recommended items that are selected by a group.

$$Precision@N = \frac{|\{top - N\ items\} \cap \{items\ that\ are\ selected\}|}{|\{top - N\ items\}|} \tag{21}$$

Recall@N is the fraction of the items that are selected by a group that are successfully recommended in the top-N items.

$$Recall@N = \frac{|\{top - N\ items\} \cap \{items\ that\ are\ selected\}|}{|\{items\ that\ are\ selected\}|} \tag{22}$$

---

[3]http://largenetwork.org/ebsn

17

Precision@N considers only the top most results returned by the system, while *Recall@N* considers all the items selected by a group. We average *Precision@N* and *Recall@N* of all groups in the test set as the final prediction. The larger *Precision@N* or *Recall@N* value means better recommendation performance.

4.3. Comparison and Implementation

In order to show the improvement of our proposed model, we implement the following baselines for the comparison with *GIST*.

- **User-based CF with averaging strategy (*CF-AVG*):** Given a group of users, this method first estimates the item selection score $\Pr(i \mid u)$ of each member in the group by performing *User-CF*. The score of group to item is equal to the average of the score for all members, i.e., $\Pr(i \mid G_g) = \frac{\sum_{u \in u_g} \Pr(i \mid u)}{|u_g|}$

- **User-based CF with least-misery strategy (*CF-LM*):** Given a group of users, this method first estimates the item selection score $\Pr(i \mid u)$ of each member by performing *User-CF*. The score of group to item is equal to the smallest of the score for all members, i.e., $\Pr(i \mid G_g) = \min_{u \in u_g} \{\Pr(i \mid u)\}$.

- **Personal impact topic model (*PIT*):** is a probabilistic model, proposed by Liu et al. (2012), which introduces the notion of personal impact to differentiate contributions of group members to a group decision. In the model, each user's personal preference is modeled as a mixture of latent topics, and the group item selection considers users' personal preferences and their personal influences.

- **COnsensus Model (*COM*):** is a probabilistic model, proposed by Yuan et al. (2014), which is designed based on these intuitions that users in a group may have different influences, and those who are expert in topics relevant to the group are usually more influential. When making recommendations, *COM* estimates the preference of a group to an item by aggregating the preferences of the group members with different weights.

Meanwhile, we implement different configurations of our model with *n-member* or *Maximal-member* to investigate impacts of different subgroup extraction strategies on recommendation accuracy.
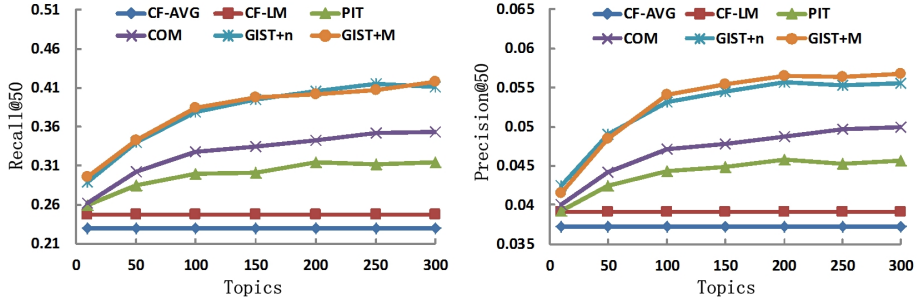
- **GIST +n:** this model uses the *n-member* subgroup extraction strategy.

- **GIST +M:** this model uses the *Maximal-member* subgroup extraction strategy.

JAMA[4] is an open matrix package for Java, developed at NIST (short for National Institute of Standards and Technology) and the University of Maryland.

---

[4]http://math.nist.gov/javanumerics/jama/
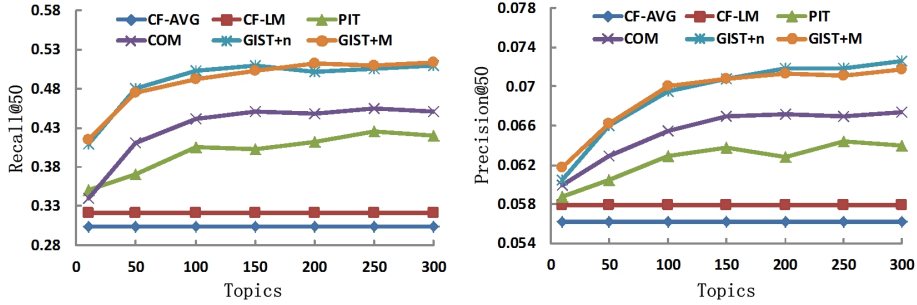
(a) Performance on Tag dataset

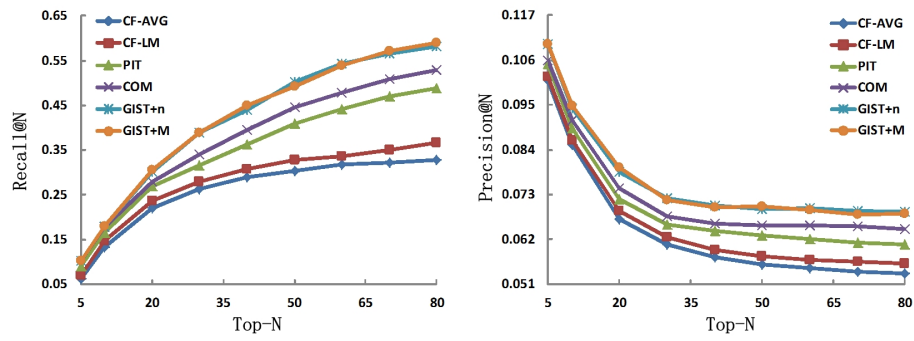(b) Performance on Venue dataset

Figure 5: Performance on different number of topics k (Top-50 recommendation).

It provides the fundamental operations of numerical linear algebra, such as matrix addition and multiplication, matrix norms and selected element-by-element array operations, etc. All models in our experiments are implemented using this library.

4.4. Performance analysis

To perform a comprehensive evaluation, we design the following experiments to test the models.

Performance on different number of topics k. In our model, the topics for users and subgroups are unshared, so their number can be different. But, they need the cross validation to reach optimal setting. It is an emerging problem of how quickly to search the optimal unshared topic parameters in topic model. But, so far, there was hardly any work to find a good way to solve it. In order to reduce the complexity of parameter searching process, their topics are initially set to the same number k, like the setting in Yuan et al. (2014). Though this way is not optimal, it is a more appropriate choice at the moment. k has an important influence on the topic models (*PIT*, *COM* and *GIST*), but has no influence on the CF-based aggregation methods (*CF-AVG* and *CF-LM*): if k is too small, it is not sufficient to discover the latent topic structure; if

19

(a) Meetup Tag dataset
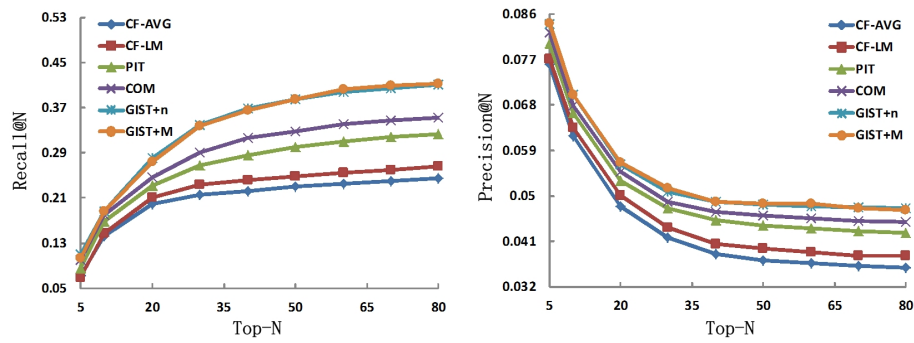
(b) Meetup Venue dataset

Figure 6: Performance on different number of recommendations N (topics $k = 200$).

k is too big, the complexity of parameter inference will considerably increase. Figure 5 shows the experiment results with different k. We first fix the number of recommendations at 50, and conduct experiments with k ranging from 10 to 300. From the figure, we observe that the performance of the topic models continues to improve with increasing the number of k, but when $k > 200$, the topic models get small performance improvement, which demonstrates the stability of the models to the topic number. Based on a balance between recommendation accuracy and computation cost, we choose $k = 200$ as the topic number. And, in particular, when the topics are enough to the models, our models ($GIST+n$ and $GIST+M$) have almost the same accuracy, both achieves superior performance than other state-of-the-art models for Top-50 recommendation.

Performance on different number of recommendations N. After setting the number of topics $k = 200$, we conduct the comparison experiments of all the models for $Top - N$ recommendation. $Recall@N$ and $Precision@N$ values of each model with N varying from 5 to 80 are shown in Figure 6. We observe that when increasing the number of recommendations N, $Recall$ increase and $Precision$ decrease regularly. After $N > 50$, the curve of two metrics tend to flatten out and the advantage of our models is much more obvious.

In order to provide a direct comparison, the ranking accuracy of all the models on $Top - 50$ recommendation is listed in Figure 7. The results show that for $Recall@50$ and $Precision@50$, our models ($GIST+n$ and $GIST+M$) have almost the same accuracy, and both greatly improve the accuracy of the baselines (CF-AVG, CF-LM, PIT and COM).

Performance on different size of groups. In order to make a more detailed analysis, we evaluate how the models perform for different size of groups based on how many members a group has. For Meetup Tag dataset, groups are categorized into six classes: "1-7", "8-14", "15-21", "22-28", "29-35" and ">=36", denoting the number of members in a group, and for Meetup Venue dataset, groups are also categorized into six classes: "1-7", "8-14", "15-21", "22-28", "29-35" and ">=36". $Recall@50$ and $Precision@50$ values of each method on the six classes of the two datasets are shown in Figure 8. We can observe that our models ($GIST+n$ and $GIST+M$) outperform the baselines for all the classes, especially when the number of group members is large, because the subgroup information increases with the group size. Among our models, $GIST+M$ outperforms $GIST+n$ on small size of groups; on the contrary, $GIST+n$ achieves better performance on large size of groups. This indicates that $Maximal$-$member$ strategy is suitable for small and simple groups, whereas n-$member$ strategy is suitable for large and complex groups.

Impact of parameter λ_g. In our model, the weights λ_g and $\lambda_g - 1$ corresponding to individual and subgroup interests can be updated dynamically based on the change of the prior $\gamma_g = \{\gamma_{g,0}, \gamma_{g,1}\}$. By incorporating the density d_g of the link graph of the members into γ_g, λ_g can be optimized to better make adjustments on different relevance of individual and subgroup interests with the group interest. To highlight the optimized function for λ_g, we design two sets of experiments: 1) use Equation 19 to generate the optimal λ_g for every group, i.e., λ_g is constantly changing in different groups; 2) fix $\lambda_g = \frac{1}{2}$
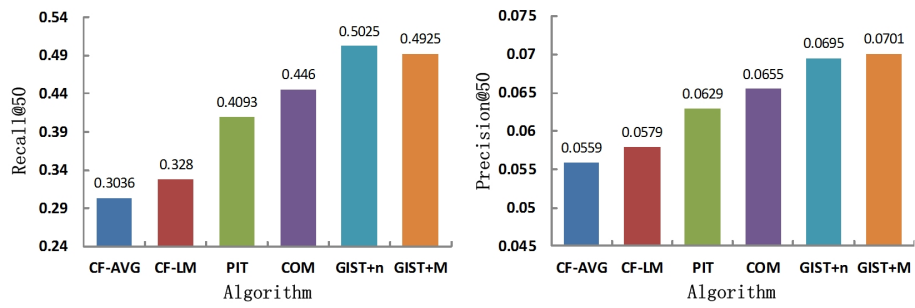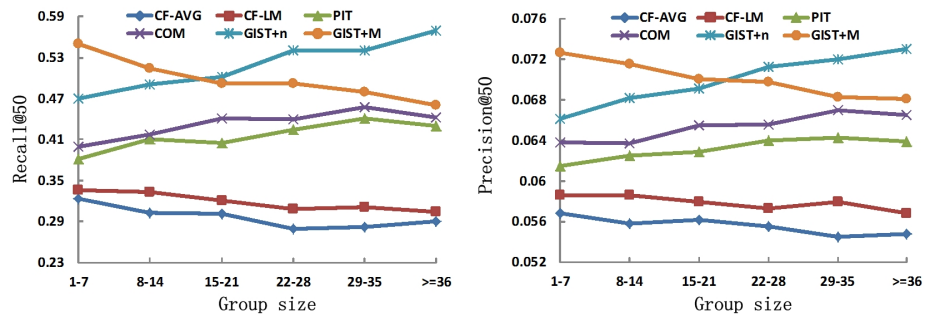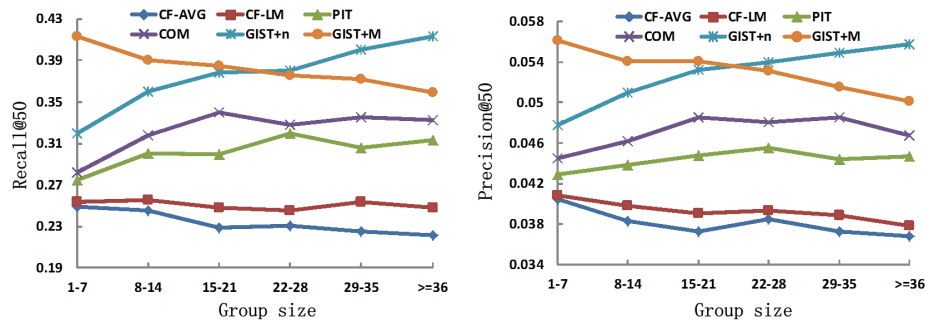
(a) Meetup Tag dataset
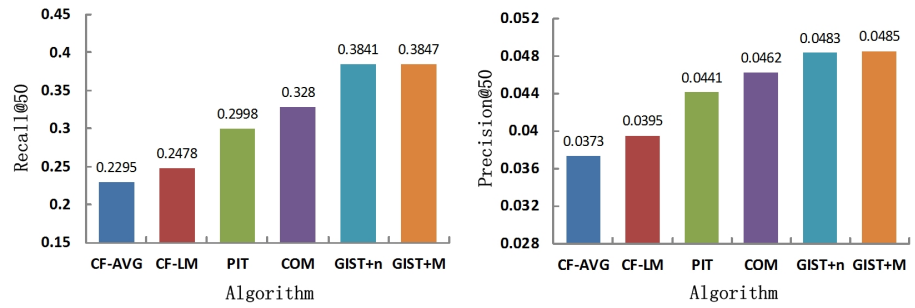
(b) Meetup Venue dataset

Figure 7: Performance comparison ($Top-50$ recommendation, topics $k = 200$).

22

(a) Meetup Tag dataset

(b) Meetup Venue dataset

Figure 8: Performance on different size of groups($Top-50$ recommendation, topics $k = 200$).

Table 3: Performance of $GIST+n$ and $GIST+M$ with $\lambda_g = \frac{1}{2}$ and the optimal λ_g ($Top-50$ recommendation, topics $k = 200$)

Dataset	Method	Recall@50		Precision@50	
		$\lambda_g = \frac{1}{2}$	optimal λ_g	$\lambda_g = \frac{1}{2}$	optimal λ_g
Tag	GIST+n	0.4864	**0.5025**	0.0672	**0.0695**
	GIST+M	0.4783	**0.4925**	0.068	**0.0701**
Venue	GIST+n	0.3769	**0.3841**	0.0453	**0.0483**
	GIST+M	0.3778	**0.3847**	0.0461	**0.0485**

for all the groups, i.e., individual members' interest and subgroups' interest are treated as equals in all groups. Table 4 shows the results of the experiments on $Top-50$ recommendation with topic $k = 200$. Compared with $\lambda_g = \frac{1}{2}$, our models ($GIST+n$ and $GIST+M$) with the optimal λ_g have better performance. This demonstrates that rather than directly giving a fixed λ_g, the optimized function can reasonably define λ_g to model the final group decision based on the group interest from individual members and subgroups.

5. Conclusions

In this paper, we study group recommendation by exploiting the members' link information to infer group activity. A key characteristic of our work is to extract some strongly connected subgroups from the graph topology constructed based on the links between the members. Then, a Topic-based probabilistic model ($GIST$) is proposed to fuse individual members' interest and subgroups' interest into a joint model, and make group recommendations by mixing the two aspects with a weight-based scheme. Specifically, the major difference from the other generative models is that we use two kinds of unshared topics to represent individual members' interest and subgroups' interest separately, and incorporate the topology information to optimize the weights. Experimental results show that our model outperforms the state-of-the-art methods.

Through a study on introducing the subgroup information, we conclusively show that the group topology contributes to modeling the generative process of group activities. Our work has preliminarily revealed that the links between the members play a major role in group recommendation. But meanwhile, there is the weaknesses of their system that how to set the suitable number of unshare topics at the low computational complexity.

In our future work, it is necessary to investigate how quickly to search the optimal parameters in unshared topic. In addition, there are more complex features in the topology, such as centricity, coverage, information diffusion. We will investigate how to incorporate such information.

Acknowledgement

Appendix A.

Equation 4 marginalise out the likelihood parameters $\left\{ \theta^{UZ_0}, \theta^{SZ_1}, \varphi^{Z_0 I}, \varphi^{Z_1 I}, \vec{\lambda} \right\}$ in closed form, and thus express the likelihood of observations directly in terms of hyperparameters $\{\alpha, \beta, \rho, \eta, \gamma\}$. The marginalization is reflected in some parts of Equation 6, 7 and 8. As an example, we give a detailed presentation for $\Pr\left(\vec{z}_{(0)} \mid \vec{u}_{(0)}, \alpha\right)$ in Equation A.1. To facilitate understanding, we simplify $\vec{n}_{u,z}^{UZ_0}$ to $\vec{n}_{u,z}$. The other parts have similar form.

$$
\begin{aligned}
& \Pr\left(\vec{z}_{(0)} \mid \vec{u}_{(0)}, \alpha\right) \\
= \ & \prod_{u=1}^{|\mathcal{U}|} \int \Pr\left(\vec{z}_{(0)} \mid \theta_u^{UZ_0}\right) \Pr\left(\theta_u^{UZ_0} \mid u, \alpha\right) d\theta_u^{UZ_0} = \prod_{u=1}^{|\mathcal{U}|} \frac{\triangle\left(\vec{n}_{u,*} + \alpha\right)}{\triangle\left(\alpha\right)} \quad \text{(A.1)}
\end{aligned}
$$

And a brief deduction to $\int \Pr\left(\vec{z}_{(0)} \mid \theta_u^{UZ_0}\right) \Pr\left(\theta_u^{UZ_0} \mid u, \alpha\right) d\theta_u^{UZ_0}$ is shown in Equation A.2.

$$
\begin{aligned}
& \int \Pr\left(\vec{z}_{(0)} \mid \theta_u^{UZ_0}\right) \Pr\left(\theta_u^{UZ_0} \mid \alpha\right) d\theta_u^{UZ_0} \\
= \ & \int \prod_{n=1}^{N_0} Mult\left(\vec{z}_{(0),n} \mid \theta_u^{UZ_0}\right) Dir\left(\theta_u^{UZ_0} \mid \alpha\right) d\theta_u^{UZ_0} \\
= \ & \int \prod_{z=1}^{k} \theta_{u,z}^{UZ_0^{(\vec{n}_{u,z})}} \frac{1}{\triangle\left(\alpha\right)} \prod_{z=1}^{k} \theta_{u,z}^{UZ_0^{(\alpha_z - 1)}} d\theta_u^{UZ_0} \\
= \ & \frac{1}{\triangle\left(\alpha\right)} \int \prod_{z=1}^{k} \theta_{u,z}^{UZ_0^{(\vec{n}_{u,z} + \alpha_z - 1)}} d\theta_u^{UZ_0} \\
= \ & \frac{\triangle\left(\vec{n}_{u,*}^{UZ_0} + \alpha\right)}{\triangle\left(\alpha\right)} \quad \text{(A.2)}
\end{aligned}
$$

Finally, the derivation process of the probability in Equation 17 is shown in Equation A.3.

$$\Pr\left(i \mid U_g, S_g\right)$$

$$= \sum_{c_g \in \{0,1\}} \Pr\left(i, c_g \mid U_g, S_g\right) = \sum_{c_g \in \{0,1\}} \Pr\left(i \mid c_g, U_g, S_g\right) \Pr\left(c_g\right)$$

$$= \Pr\left(i \mid c_g = 0, U_g, S_g\right) \Pr\left(c_g = 0\right) + \Pr\left(i \mid c_g = 1, U_g, S_g\right) \Pr\left(c_g = 1\right)$$

$$= \lambda_g \Pr\left(i \mid c_g = 0, U_g, S_g\right) + (1 - \lambda_g) \Pr\left(i \mid c_g = 1, U_g, S_g\right)$$

$$= \lambda_g \sum_{u \in U_g} \Pr\left(i, u \mid U_g\right) + (1 - \lambda_g) \sum_{s \in S_g} \Pr\left(i, s \mid S_g\right)$$

$$= \lambda_g \sum_{u \in U_g} \Pr\left(i \mid u\right) \Pr\left(u \mid U_g\right) + (1 - \lambda_g) \sum_{s \in S_g} \Pr\left(i \mid s\right) \Pr\left(s \mid S_g\right)$$

$$= \lambda_g \sum_{u \in U_g} \pi_{g,u}^{U} \Pr\left(i \mid u\right) + (1 - \lambda_g) \sum_{s \in S_g} \pi_{g,s}^{S} \Pr\left(i \mid s\right)$$

$$= \lambda_g \sum_{u \in U_g} \pi_{g,u}^{U} \sum_{z \in Z_1} \Pr\left(i, z \mid u\right) + (1 - \lambda_g) \sum_{s \in S_g} \pi_{g,s}^{S} \sum_{z \in Z_0} \Pr\left(i, z \mid s\right)$$

$$= \lambda_g \sum_{u \in U_g} \pi_{g,u}^{U} \sum_{z \in Z_1} \Pr\left(i \mid z, u\right) \Pr\left(z \mid u\right) + (1 - \lambda_g) \sum_{s \in S_g} \pi_{g,s}^{S} \sum_{z \in Z_0} \Pr\left(i \mid z, s\right) \Pr\left(z \mid s\right)$$

$$= \lambda_g \sum_{u \in U_g} \sum_{z \in Z_0} \pi_{g,u}^{U} \theta_{u,z}^{UZ_0} \varphi_{z,i}^{Z_0 I} + (1 - \lambda_g) \sum_{s \in S_g} \sum_{z \in Z_1} \pi_{g,s}^{S} \theta_{s,z}^{SZ_1} \varphi_{z,i}^{Z_1 I} \tag{A.3}$$

References

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, *17*, 734–749.

Basu Roy, S., Lakshmanan, L. V., & Liu, R. (2015). From group recommendations to group formation. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data* SIGMOD '15 (pp. 1603–1616). New York, NY, USA: ACM.

Berkovsky, S., & Freyne, J. (2010). Group-based recipe recommendations: Analysis of data aggregation strategies. In *Proceedings of the Fourth ACM Conference on Recommender Systems* RecSys '10 (pp. 111–118). New York, NY, USA: ACM.

Blei, D. M. (2012). Probabilistic topic models. *Commun. ACM*, *55*, 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, *3*, 993–1022.

Brand, M. (). Fast online svd revisions for lightweight recommender systems. In *Proceedings of the 2003 SIAM International Conference on Data Mining* chapter 4. (pp. 37–46).

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence* UAI'98 (pp. 43–52). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Carvalho, L. A. M. a. C., & Macedo, H. T. (2013). Users' satisfaction in recommendation systems for groups: An approach based on noncooperative games. In *Proceedings of the 22Nd International Conference on World Wide Web* WWW '13 Companion (pp. 951–958). New York, NY, USA: ACM.

Delic, A., Neidhardt, J., Nguyen, T. N., Ricci, F., Rook, L., Werthner, H., & Zanker, M. (2016). Observing group decision making processes. In *Proceedings of the 10th ACM Conference on Recommender Systems* RecSys '16 (pp. 147–150). New York, NY, USA: ACM.

Dwork, C., Kumar, R., Naor, M., & Sivakumar, D. (2001). Rank aggregation methods for the web. In *Proceedings of the 10th International Conference on World Wide Web* WWW '01 (pp. 613–622). New York, NY, USA: ACM.

Gao, L., Wu, J., Qiao, Z., Zhou, C., Yang, H., & Hu, Y. (2016). Collaborative social group influence for event recommendation. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* CIKM '16 (pp. 1941–1944). New York, NY, USA: ACM.

Guy, I., Ronen, I., Kravi, E., & Barnea, M. (2016). Increasing activity in enterprise online communities using content recommendation. *ACM Trans. Comput.-Hum. Interact.*, *23*, 22:1–22:28.

Hao, Q., Cai, R., Wang, C., Xiao, R., Yang, J.-M., Pang, Y., & Zhang, L. (2010). Equip tourists with knowledge mined from travelogues. In *Proceedings of the 19th International Conference on World Wide Web* WWW '10 (pp. 401–410). New York, NY, USA: ACM.

Heinrich, G. (2008). Parameter estimation for text analysis. *Technical Report*, .

Hofmann, T. (1999). Probabilistic latent semantic indexing. In *Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* SIGIR '99 (pp. 50–57). New York, NY, USA: ACM.

Hofmann, T. (2004). Latent semantic models for collaborative filtering. *ACM Trans. Inf. Syst.*, *22*, 89–115.

Jameson, A., & Smyth, B. (2007). The adaptive web. chapter Recommendation to Groups. (pp. 596–627). Berlin, Heidelberg: Springer-Verlag.

Jin, R., Chai, J. Y., & Si, L. (2004). An automatic weighting scheme for collaborative filtering. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* SIGIR '04 (pp. 337–344). New York, NY, USA: ACM.

Kurashima, T., Iwata, T., Hoshide, T., Takaya, N., & Fujimura, K. (2013). Geo topic model: Joint modeling of user's activity area and interests for location recommendation. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining* WSDM '13 (pp. 375–384). New York, NY, USA: ACM.

Lee, K., Kim, Y. H., Baek, S. H., & Song, M. (2015). Analyzing subject-method network of bioinformatics and biology. In *Proceedings of the ACM Ninth International Workshop on Data and Text Mining in Biomedical Informatics* DTMBIO '15 (pp. 27–27). New York, NY, USA: ACM.

Liu, X., Tian, Y., Ye, M., & Lee, W.-C. (2012). Exploring personal impact for group recommendation. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management* CIKM '12 (pp. 674–683). New York, NY, USA: ACM.

Marlin, B. (2004). Modeling user rating profiles for collaborative filtering. In S. Thrun, L. Saul, & B. Schölkopf (Eds.), *Advances in Neural Information Processing Systems 16*. Cambridge, MA: MIT Press.

Masthoff, J. (2004). Group modeling: Selecting a sequence of television items to suit a group of viewers. *User Modeling and User-Adapted Interaction*, *14*, 37–85.

Masthoff, J. (2011). Group recommender systems: Combining individual models. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 677–702). Springer US.

McCarthy, J. F., & Anagnost, T. D. (1998). Musicfx: An arbiter of group preferences for computer supported collaborative workouts. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* CSCW '98 (pp. 363–372). New York, NY, USA: ACM.

Neetil, J., & de Mendez, P. O. (2012). *Sparsity: Graphs, Structures, and Algorithms*. Springer Publishing Company, Incorporated.

Newman, M. (2010). *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc.

Newman, M., Barabasi, A.-L., & Watts, D. J. (2006). *The Structure and Dynamics of Networks: (Princeton Studies in Complexity)*. Princeton, NJ, USA: Princeton University Press.

Rajani, N. F. N., McArdle, K., & Baldridge, J. (2014). Extracting topics based on authors, recipients and content in microblogs. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval* SIGIR '14 (pp. 1171–1174). New York, NY, USA: ACM.

Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004). The author-topic model for authors and documents. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence* UAI '04 (pp. 487–494). Arlington, Virginia, United States: AUAI Press.

Salakhutdinov, R., & Mnih, A. (2008). Probabilistic matrix factorization. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in Neural Information Processing Systems 20* (pp. 1257–1264). Cambridge, MA: MIT Press.

Salehi-Abari, A., & Boutilier, C. (2015). Preference-oriented social networks: Group recommendation and inference. In *Proceedings of the 9th ACM Conference on Recommender Systems* RecSys '15 (pp. 35–42). New York, NY, USA: ACM.

Srebro, N., & Jaakkola, T. (2003). Weighted low-rank approximations. In *ICML* (pp. 720–727).

Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Adv. in Artif. Intell.*, *2009*, 4:2–4:2.

Wang, C., & Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* KDD '11 (pp. 448–456). New York, NY, USA: ACM.

Wang, J., de Vries, A. P., & Reinders, M. J. T. (2006). Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* SIGIR '06 (pp. 501–508). New York, NY, USA: ACM.

Ye, M., Liu, X., & Lee, W.-C. (2012). Exploring social influence for recommendation: A generative model approach. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* SIGIR '12 (pp. 671–680). New York, NY, USA: ACM.

Yu, Z., Zhou, X., Hao, Y., & Gu, J. (2006). Tv program recommendation for multiple viewers based on user profile merging. *User Modeling and User-Adapted Interaction*, *16*, 63–82.

Yuan, Q., Cong, G., & Lin, C.-Y. (2014). Com: A generative model for group recommendation. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* KDD '14 (pp. 163–172). New York, NY, USA: ACM.

Zhang, X., Li, Z., Wang, S., Yang, Y., & Lv, X. (2015). Location prediction of social images via generative model. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval* ICMR '15 (pp. 275–282). New York, NY, USA: ACM.