An Enhanced Probabilistic Fairness-Aware Group Recommendation by Incorporating Social Activeness

Yang Xiao\textsuperscript{a}, Qingqi Pei\textsuperscript{a,}\textsuperscript{*}, Lina Yao\textsuperscript{b}, Shui Yu\textsuperscript{c}, Lei Bai\textsuperscript{b}, Xianzhi Wang\textsuperscript{c}

\textsuperscript{a}State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, 710071, China
\textsuperscript{b}School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia
\textsuperscript{c}School of Software, University of Technology Sydney, Sydney, NSW 2007, Australia

Abstract

Compared with individual recommendation, recommending services to a group of users is more complicated because of various users’ preference should be considered and introduces new challenging such as fairness, which has never been well studied in current works. In this paper, we propose a novel recommendation scheme called PFGR, which combines a probabilistic model with coalition game strategy, to ensure the accuracy and fairness between groups of users. Given a group of users and a set of services, PFGR models a generative process for service selection in light of several observations: 1) each group is related with several topics; 2) users’ decisions on the service selection depends on their expertise, the opinions of members they are familiar with, and group influence; 3) each group contains active users and inactive user, whose activeness contributes to the existence of group. PFGR first estimates the preference of each user on a candidate service via combining user’s expertise, inherent connection, and group influence. Then, it determines a group’s decision on a service by aggregating the preference of group members using adaptive weights. Finally, PFGR considers users’ activeness and employs a strategy based on coalition game to produce a ranked list which is fair to each group member as much as possible. Experimental results on three real-world datasets validate that PFGR can achieve higher Hit Rate and Average Reciprocal Hit Rank than state-of-the-art approaches, which indicates that PFGR attains both the precision and fairness of recommendation.

Keywords: Group recommendation, User activeness, Probabilistic model, Fairness, Coalition game

1. Introduction

Traditional recommender systems (RSs) aim to provide appropriate services for a single user based on her preferences. Such RSs have been deployed in a wide range of areas such as music (Yahoo), restaurants (Foursquare), and hiking (Meetup). However, many contexts requires recommending to a group of users (i.e., group recommendation) while various preferences of all the group members should be considered. For example, in cases of selecting a picnic location for a group of friends, recommending a restaurant for a company’s annual meeting, arranging attractions for a group of tourists, the traditional individual recommendation methods no longer fit.

Group recommendation is more complicated than individual recommendation. Since group members may have different preferences \cite{1,2}, a service preferred by one user may not satisfy another user’s taste. Moreover, each user hopes her preferred service to appear at a top position in the service list recommended to her group. According to the studies in the fair division of sources \cite{3,4,5}, a recommended services list is fair to a user if and only if her preferred service is ranked at a top position \cite{6}. Therefore, it is of paramount importance to recommend a ranked service list that is fair to every user, i.e., fairness. An ideal recommendation approach for group not only guarantee the accuracy but also efficiently solve fairness issue.

Most current studies on group recommendation \cite{7,8,9,10,11,12} determines the services that satisfy the group members’ preferences via modelling users’ implicit peer influence \cite{1}. However, they cannot solve the fairness issue because they commonly lack a proper method to balance the various preferences. Other studies \cite{6,13,14,2,15} convert the fairness issue into a comparison sequencing problem and design a preference-based sequencing strategy to rank the recommended services. Although this strategy can ensure fairness to some extent, it cannot tackle the scenarios where group members have conflicting preferences. As it is intractable to compare users’ preference (e.g., distinguish the optimal options from spicy and light food preferences), the recommended list derived by this strategy can only guarantees a part of users’ preferences instead of all the users’ preferences. Therefore, sequencing strategy based on preference is improper.

Fortunately, the social regularization principle \cite{16} pro-
vides a interesting viewpoint: the more contribution you pay, the more priority or return you win [17]. For a group, its formation and sustainability heavily depends on its members’ activeness, which refers to as the frequency of users’ interactions including sharing information or extending the social circle [13]. Inspired by the social regularization principle, it is more intuitive and proper to consider users’ activeness when ranking services, i.e., a user’s preference should be satisfied in priority if she contributes to the group more actively. Different from dealing with users’ preferences, we can easily quantify users’ activeness via simple statistic methods [19] and handle conflicting user preferences. For example, we can count up how many friends a user has or how much shopping information she shares.

We borrow the fairness definition from [6, 13] and propose a novel two-stage group recommendation model called PFGR. PFGR couples user’s various preferences and activeness, which has seldom been studied by previous work. PFGR consists of two parts: multi-facet probabilistic graph model (MFPG) and activeness-based coalition game strategy (ACG). During recommendation, PFGR first applies MFPG to produce the services which satisfy all the members’ preferences by modeling several observations (see Section 3.5) obtained from the real life. Then, it utilizes ACG to rank these services to attain a trade-off among various preferences.

Specifically, MFPG is a probabilistic generative model that aims to select the services preferred by a group. It is developed on latent Dirichlet allocation (LDA), which has been proven successful in modeling implicit interactions [20, 21]. Compared with other group recommendation model based on LDA [1, 9], MFPG considers more implicit interactions such as users’ social links, preferences, influence, and common-interest. In particular, considering users’ implicit interactions can help group members to better select their desired services. ACG is inspired by the coalition game theory, which has two advantages when compared with current sequencing strategy based on the greedy algorithm [6, 13] or the non-cooperative game theory [14, 15]: 1) instead of considering a single user’s preference, the coalition game innately considers users’ peer influence (e.g., common-interest, social links) and therefore conforms to the fact that a user’s selection may be affected by others; 2) the coalition game theory considers the balance between several coalitions. That makes it easier to find the equilibrium among a large number of users in a dynamic environment where each user’s preferences may change over time.

We make the following contributions in this paper:

• We propose a novel two-stage group recommendation approach named PFGR which both guarantee the accuracy of recommendation and efficiently solve fairness issue. PFGR couples users’ preferences and activeness, which has not been well studied before.

• We design an activeness-based sequencing strategy to ranking services following the social regularization principle to promote the fairness in recommendation. This strategy can better solve conflicted preference contexts when compared with the traditional preference-based sequencing strategy.

• We conduct extensive experiments to validate the effectiveness of PFGR under various settings on three real data sets. The evaluation results show our scheme consistently outperforms state-of-the-art approaches when considering the fairness simultaneously.

The rest of this paper is organized as follows: Section 2 reviews the related work. Section 3 introduces the preliminaries and formulates the group recommendation problem. Section 4 presents the details of our proposal, including the MFPG model and ACG strategy. Section 5 reports our analysis of experiment results. Finally, Section VI concludes the paper.

2. Related Work

2.1. Group Recommendation

Generally, group recommendation methods can be divided into two categories: the preference aggregation method and the score aggregation method [22]. The former method first aggregates the profiles of the group members into one file, i.e., constructs a virtual user, and then make recommendations to this virtual user [23, 24]. The latter, on the contrary, first produces recommendations for each group member, then aggregates their recommendation results to this group [25]. In our work, the proposed approach belongs to the score aggregation method.

The score aggregation approaches usually employ two aggregation strategies: Average and Least Misery, which have been widely adopted in group recommendation [13, 7]. Recently, several score aggregation-based models have been proposed. In [9, 1], authors assume that each group has a multinomial distribution over latent topics and these topics attract a lot of users to join in. The service selection of a user depends on either the group influence or personal consideration. [7] designs a rank aggregation methods combining AVE with LM strategies. [7] first generate s each user’s rating predictions on candidate services and then aggregate this rating via AVE or LM strategy to get the final recommendation for the group. [6] and [13] apply a greedy algorithm to maximize the performance of group recommendation. Other schemes [10, 11] involve trust or social relationships in group recommendation. [10] considers social relationships strength in a group collaborative filtering context. [11] defines an empathetic social choice framework in which agents derive utility based on both their intrinsic preferences and the satisfaction of their neighbors.

Although these methods consider users’ implicit peer influence such as the peer influence or social links, they
can’t handle fairness issue because they lack a proper strategy to determine a balance trade-off among users’ various preferences.

2.2. Fairness in Group Recommendation

Several works focus on fairness in group recommendation. Some schemes [26, 27] treat the group decision as a voting campaign and use voting mechanism to find a proper recommendation. However, these schemes do not explicitly consider fairness in the models, and the definition of fairness in these works is obscure. Besides, these works do not involve the connection between users in group recommendation, either. Another works [15, 14, 2, 28] aim to find an equilibrium among various users’ preferences via considering social relationship. Although introducing social links can to some degrees solve fairness issue, these method can’t handle the conflicted contexts because of their strategies are preference-based.

Two similar works on fairness in group recommendation are [9, 13]. Different from other works, they explicitly define the fairness which conforms to fair division of resources [3, 4, 5]. In their works, fairness is defined as a fact that a user is satisfied with a service if and only if this service is ranked at the top-rated position in the final recommended list. More specifically, in [6], authors first define fairness based on proportionality and envy-freeness. Then they extend the definition into two practical scenarios where they add categories and spatial constraints and design a greedy strategy based on preference to maximize fairness in these two scenarios. However, scheme in [6] sometimes may cause greater unfairness. Consider a group with ten users and a recommended service list, seven of them are satisfied with this list while three of them dislike. According to [6], the fairness value is 0.7 which means most members of this group think this recommendation is fair while ignoring the remaining three users. This fairness is prejudiced when neglecting three users. [15] considers fairness from the perspective of social welfare. Authors first construct individual utility function for each user in a group and then propose two concepts of social welfare and fairness for modeling utility function and the balance between group members. Then they determine the average utility value of each service. Obviously, [15] is more fair than [6] because [15] considers all the users’ preferences. However, it can’t handle the conflicted contexts because the definition about social welfare confused the contradiction of preferences between members, while our social activeness can handle it with the frequency of interaction (easily to be quantified) considered.

In our work, we propose a more proper sequencing strategy based on activeness. Compared with preference, it is tractable to quantify users’ activeness via simple statistic method [19], and our strategy also conforms to social regularization principles [16] (More details are shown in Section 3.1).

### Table 1: Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G, S$</td>
<td>a group, services set</td>
</tr>
<tr>
<td>$</td>
<td>G</td>
</tr>
<tr>
<td>$\mu, \mu_i$</td>
<td>any user in G, the $i$th user in G</td>
</tr>
<tr>
<td>$s_i$</td>
<td>the $i$th service in $S$</td>
</tr>
<tr>
<td>$Z$</td>
<td>a set of latent topics including $K$ topics, i.e., $Z = {z_1, z_2, ..., z_K}$</td>
</tr>
<tr>
<td>$Z_{\mu}$</td>
<td>user $\mu$’s latent topic set, i.e., $Z_{\mu} = {z_{\mu 1}, ..., z_{\mu \tau}}$</td>
</tr>
<tr>
<td>$D$</td>
<td>a decision set containing four value, i.e., $D = {d</td>
</tr>
<tr>
<td>$T_{mat}$</td>
<td>social relationships matrix in G</td>
</tr>
<tr>
<td>$T_{\mu}$, $\mu_i$</td>
<td>the set of $\mu$’s social links, $\mu_i \in T_{\mu}$</td>
</tr>
<tr>
<td>$C_{\mu}$, $\mu_c$</td>
<td>the set of users with common interests with $\mu$, $\mu_c \in C_{\mu}$</td>
</tr>
<tr>
<td>$&lt; \mu, s_j &gt;$</td>
<td>user $\mu$ selects $s_j$</td>
</tr>
<tr>
<td>$\theta_{\mu}$, $\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>topic distribution of a group $G$</td>
</tr>
<tr>
<td>$\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>distribution of users specific expertise on topic $z$</td>
</tr>
<tr>
<td>$\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>distribution of users specific expertise on service $s$</td>
</tr>
<tr>
<td>$\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>distribution of group specific expertise on service $s_j$</td>
</tr>
<tr>
<td>$\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>distribution of users in $T_{\mu}$ specific expertise on service $s$</td>
</tr>
<tr>
<td>$\psi^z_{\mu}$, $\psi^s_{\mu}$, $\psi^z_{\psi}$, $\psi^s_{\psi}$, $\psi^c_{\psi}$</td>
<td>distribution of users in $C_{\mu}$ specific expertise on service $s$</td>
</tr>
<tr>
<td>$S_{red}$</td>
<td>a ranked services list after adjustment</td>
</tr>
<tr>
<td>$\alpha, \beta_1, \beta_2$</td>
<td>parameters of $\theta_{G}, \psi^z_{\mu}, \psi^s_{\mu}, \psi^z_{\psi}, \psi^s_{\psi}, \psi^c_{\psi}$</td>
</tr>
<tr>
<td>$\rho, \eta_1, \eta_2$</td>
<td>number of times topic $z$ is assigned to $G$</td>
</tr>
<tr>
<td>$\tau_{z,G}$</td>
<td>number of times user $\mu$ is derived from topic $z$</td>
</tr>
<tr>
<td>$\tau_{\mu,z}$</td>
<td>number of times service $s$ is derived from user $\mu$</td>
</tr>
<tr>
<td>$\tau_{\mu,s}$</td>
<td>number of times $s$ is derived from $\mu \in T_{\mu}$</td>
</tr>
<tr>
<td>$\tau_{c,s}$</td>
<td>number of times $s$ is derived from $\mu_c \in C_{\mu}$</td>
</tr>
<tr>
<td>$\tau_{z,s}$</td>
<td>number of times service $s$ is derived from topic $z$ in $G$</td>
</tr>
<tr>
<td>$\tau_{\mu,d}$</td>
<td>number of times $d$ is drawn from $\mu$</td>
</tr>
</tbody>
</table>

### 3. Preliminaries

In this section, we first introduce some preliminaries and problem formulation, then provide several observations concluded from the real world. The main notations used in this paper are listed in Table 1.

#### 3.1. Users’ activeness in group

In sociology, a group can be defined as two or more people who interact with one another, share similar characteristics, and collectively have a sense of unity [29, 30]. According to the definition, we know that the form and sustainability of a group depends on the frequency of users’
interaction including sharing information (e.g., shopping experience, service promotion, etc.) or extending social circle [13, 2]. In this work, we define the frequency of users’ interactions as activity based on [31]. Generally, a group contains active members and inactive members. Active members often share more information including shopping experience, interesting news, etc. or attract new user to join the group, such interactions pay more contribution to the existence of group than the inactive. Hence, users’ activity must be taken into account when making recommendation for groups. Specifically, service preferred by active users should be in priority ranked at a top position when ranking service because of their more contribution.

3.2. Coalition game theory

Coalition game theory has been validated to be efficient in resource distribution, decision making and widely utilized in economic and engineering areas [32, 33, 34, 35]. A coalition game is a game with competition between groups of players due to the possibility of external enforcement of cooperative behavior [30, 37, 38]. The game is thus a competition between coalitions of players rather than a competition between individual players.

Formally, the coalitional game contains a set of \( n \) players which can be divided into \( C \) coalitions \((C < n)\) and a characteristic function \( v : 2^N \rightarrow \mathbb{R}\), where the characteristic function of the game assigns to each possible coalition a numeric value that intuitively represent the utility or payoff which can be distributed among coalition members. The final target of coalition game is to optimize the sum of utility value.

For recommendation, a group contains active users and inactive user, which can be constructed two coalition according to coalition game theory. In our work, we divide the coalition based on users’ activity. The utility value \( v \) in our work represents a ranked service list. Different from coalition game theory, our target is to determine a proper ranked service list where the position of each service can satisfy users’ preference as much as possible.

3.3. Social links

According to [39, 40], social links is defined as the connections that exit between people who have recurring interactions that are perceived by the participants to have personal meaning. This definition contains relationships between friends, neighbors, workmates, etc. In RSs, current works aims to consider two kind of relationships, i.e., trust relationship and friendship, which has been validated to significantly improve the recommendation performance in practice [41, 14, 28, 42, 43].

Generally, trust relationship and friendship is modelled as a graph and represented as a 0-1 matrix \( T_{\text{mat}} \), i.e., \( \forall \mu, \mu_j \in G, \text{if there exists social relationship between them, } T_{\text{mat}} = 1 \), otherwise 0. Note that the difference between trust relationship and friendship is that the former is modelled as a directed graph while the later is modelled as an undirected graph [44]. In real social platforms such as Epinions and Douban, social relationship are precisely expressed. In our experiments, social relationship is directly obtained in the data sets.

3.4. Problem statement

Given a set of services \( S = \{s_1, s_2, ..., s_n\} \) to be recommended, \( G \) is a group which contains \( m \) users. For \( \forall \mu (\mu \in G) \), we can obtain his preference according to purchased services. Besides, there exist some users who are connected with \( \mu \) via social links. Here we use \( T_\mu \) to denote the set of \( \mu \)'s social links, \( T_\mu = \{\mu_{t_1}, \mu_{t_2}, ..., \mu_{t_k}\} (\forall \mu_{t_k} \in G) \). We hope the recommended service list is fair to the group users.

Definition Fairness. According to [3], Given a top-\( N \) recommended service list, \( S_{\text{red}} = \{s_{p_1}, s_{p_2}, ..., s_{p_{\text{red}}}\} \) \((p_i \text{ means the position of } s_i, S_{\text{red}} \in S)\), if a service \( s \) preferred by user \( \mu \) belongs to \( S_{\text{red}} \), i.e., \( s \in S_{\text{red}} \), we say \( S_{\text{red}} \) is fair to \( \mu \). For a group, if the position of each service of \( s_i \) is fair to its members as much as possible, \( S_{\text{red}} \) is fair to this group.

The goal of our model is to determine \( S_{\text{red}} \) in a specific sequence which is fair to group members as much as possible.

3.5. Observations

In this section, we generalize the following observations based on the real world, which provide support for the proposed model in theory.

- **Observation 1**: Each group is related with one or more topics. i.e., a sports club is more relevant to basketball or football games. The topics of this group may attract more users to join it. Besides, a group itself has some topic-based knowledge about services if they are related to certain topics, here is referred to as group preference [1].

- **Observation 2**: Besides user’s personal preferences, a user’s decision on services generally depends on other users. Several conditions should be considered when recommending a service \( s_i \) to a group user. 1) If a user \( \mu \) is expert in \( s_i \), his decision on \( s_i \) just depends on himself [1]. 2) If \( \mu \) knows little about \( s_i \), but his friends in this group are expert in it, \( \mu \)'s final decision on \( s_i \) relies on his friends’ decision. 3) If \( \mu \) has no friends or trusted members in this group, he may consult others who have a similar preference to him. Whether selecting \( s_i \) or not depends on those members with similar preference. Note that there exist some members with similar preference are also \( \mu \)'s friends. 4) \( \mu \) may tend to obey the group’s decision if he neither knows much about \( s_i \) nor has friends or members with similar preferences [1].
Four scenarios should be considered when \( \mu \) selects \( s \) (Observation 2). Here we use a switch \( d \) to decide which one may happen for \( \mu \)'s selection of \( s \), i.e.,

- if \( d = 0 \), \( \mu \) selects \( s \) based on his own expertise, which is a multinomial distribution over services \( \psi^s_\mu \).
- if \( d = 1 \), user \( \mu \) picks out \( s \) based on his social influence (e.g., friends), which is a uniform distribution on \( T_\mu \). Each member in \( T_\mu \) has his expertise on \( s \), which satisfies a multinomial distribution on \( \psi^s_\mu \).
- if \( d = 2 \), \( \mu \) selects \( s \) according to other members who have similar interests with \( \mu \), which is a uniform distribution on \( C_\mu \). Each member in \( C_\mu \) has his understanding about \( s \), which is a multinomial distribution on \( \psi^s_\mu \).
- if \( d = 3 \), that means \( \mu \) has neither expertise on \( s \) nor friends or users with common interests. Thus, \( \mu \) selects \( s \) according to group preference on \( s \), which is a multinomial distribution over \( \psi^s_\mu \).

Compared with other probabilistic graph-based works [9, 10], our approach has two improvements on group recommendation: 1) [9, 10] only consider two scenarios, i.e., the selection of services either depends on the user itself or group decision. However, their consideration can’t well reflect the practical situation in the real world. There exists explicit (e.g., friends or relatives) or implicit (e.g., common interest on sports) connection among users in a group. When a user \( \mu \) selects a service \( s \), he will consult other familiar users (e.g., friends or some people with common interests) if he is not clear about \( s \). Hence, the final decision of \( \mu \) on \( s \) generally depends on the opinions from these users instead of directly conforming to group influence. 2) From the perspective of services selection, [9, 10] apply a Bernoulli distribution on switch value to simulate the situation of user’s selection, i.e., if switch value is 0, user \( \mu \) select \( s \) depends on personal preference, otherwise on group preference. This simulation method can’t well reflect the real situations because each user can judge whether conforming to group preference, i.e., it should be that user decides the switch value (i.e., \( d \)) instead of random generation.

In this paper, we design a simple method to simulate users’ action on switch value: 

**Method:** each service \( s \) is related to certain topic \( z_s \), e.g., tent related with camping, restaurant related with a party, etc. In practice, each user \( \mu \) has experienced some services corresponding to several topics, denoted by \( Z_\mu = \{z_{\mu 1}, \ldots, z_{\mu l}\} \). If \( z_{\mu l} \in Z \), \( \mu \) has prior knowledge on \( s \), then \( d = 0 \); if \( z_{\mu l} \notin Z \), and \( \mu \)'s friends or other users who have common interests with \( \mu \) know \( z_{\mu l} \), then \( d = 1 \) or \( d = 2 \), otherwise \( d = 3 \). Algorithm 1 summarizes the complete generative process of MFPG.

To learn the parameters in MFPG, the estimation of...
the posterior likelihood function is defined by
\[ P(z, \mu, s | \alpha, \beta_1, \beta_2, \eta_1, \eta_2, \rho) = \int P(z | \theta_G) P(\theta_G | \alpha) d\theta_G \cdot \int P(\mu | z, \psi^s_\mu) P(\psi^s_\mu | \beta_1) d\psi^s_\mu \cdot A \] (1)
where A is defined as (2):
\[
A = \int \int \int P(s | \mu, z, d, \mu_c, \mu_k, \psi^s_\mu, \psi^s_k) P(\psi^s_\mu | \eta_2) \cdot P(\psi^s_k | \eta_1) P(\psi^s_\mu | \beta_2) d\psi^s_k d\psi^s_\mu d\psi^s_{\mu_k} d\psi^s_{\mu_c} d\psi^s_{\mu} d\psi^s_{\mu_2}
\] (2)

To infer the parameters \{\psi^s_\mu, \psi^s_k, \psi^s_c, \psi^s_0\}, we apply collapsed Gibbs sampling method to obtain samples from high-dimensional distribution. For a given latent topic variable \(z\), a Gibbs sampling method needs to calculate the full conditional probability for the assignment of the variable conditioned on all the assignment excluding \(z\).

However, this is intractable to get the full conditional probability because of complex inter-dependencies between user \(\mu\), service \(s\), topic \(z\) and switch value \(d\) i.e., the final decision of \(\mu\) on \(s\) depends on \(d\) which has 4 values in this paper.

To overcome this problem, we apply four-step Gibbs sampling method based on (1) by decomposing equation (2) as follows:
\[
A = \frac{\int \int \int P(s^0 | \mu, d, \psi^0_\mu) P(\psi^0_\mu | \beta_2) d\psi^0_\mu \cdot P(s^1 | \mu_c, d, \psi^1_\mu) P(\psi^1_\mu | \eta_2) d\psi^1_\mu \cdot P(s^2 | \mu_k, d, \psi^2_\mu) P(\psi^2_\mu | \beta_2) d\psi^2_\mu \cdot P(s^3 | \mu, d, \psi^3_\mu) P(\psi^3_\mu | \beta_2) d\psi^3_\mu}{A_0 A_1} A_2 A_3
\] (3)

where \(s^0\) means that user \(\mu\) chooses \(s\) according to his own expertise, \(s^1\) means that \(\mu\) chooses \(s\) according to his social links, \(s^2\) means that \(\mu\) selects \(s\) according to other users with common interests, \(s^3\) means that \(\mu\) selects \(s\) according to group influence.

Based on the new likelihood function shown in equation (1) and (3), we can determine the full conditional distribution of any topic \(z_j \in Z\) and switch \(d\) for \(\mu\) and \(s_j\). If \(s_j\) is selected by \(\mu\)'s personal expertise, i.e., \(d=0\), we sample \(z_j\) according to the following probability (20):
\[
P(z_j = k | z^{-j}, \mu, s^0) = \frac{\int P(Z | \theta_G) P(\theta_G | \alpha) d\theta_G \cdot \int P(\mu | Z, \psi^0_\mu) P(\psi^0_\mu | \beta_1) d\psi^0_\mu}{\int P(Z | \theta_G) P(\theta_G | \alpha) d\theta_G} \cdot \int P(\mu | Z, \psi^0_\mu) P(\psi^0_\mu | \beta_1) d\psi^0_\mu \propto \frac{\tau_{G,j} + \alpha_k}{\sum_{k \in Z} \tau_{G,j} + \alpha_k} \cdot \frac{\tau_{\mu,j} + \beta^\mu_1}{\sum_{\mu \in \mathcal{C}_\mu} \tau_{\mu,j} + \beta^\mu_1} + \beta^\mu_2 \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2} \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2}
\] (4)

where ‘\(-j\)’ means that we exclude the \(j\)th service for \(G\) when sampling. The similar derivation of collapsed Gibbs sampling equation for other \(d\)’s value is shown as:
\[
P(z_j = k | z^{-j}, \mu, s^1, s^2, s^3) \propto \frac{\tau_{G,j} + \alpha_k}{\sum_{k \in Z} \tau_{G,j} + \alpha_k} \cdot \frac{\sum_{\mu \in \mathcal{C}_\mu} \tau_{\mu,j} + \beta^\mu_1}{\sum_{\mu \in \mathcal{C}_\mu} \tau_{\mu,j} + \beta^\mu_1} + \beta^\mu_2 \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2} \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2}
\] (5)
\[
P(z_j = k | z^{-j}, \mu, s^3) \propto \frac{\tau_{G,j} + \alpha_k}{\sum_{k \in Z} \tau_{G,j} + \alpha_k} \cdot \frac{\sum_{\mu \in \mathcal{C}_\mu} \tau_{\mu,j} + \beta^\mu_1}{\sum_{\mu \in \mathcal{C}_\mu} \tau_{\mu,j} + \beta^\mu_1} + \beta^\mu_2 \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2} \frac{\sum_{c \in Z} \tau_{c,j} + \eta_1}{\sum_{c \in Z} \sum_{c^2} \tau_{c,j} + \eta_2}
\] (6)

After sampling a sufficient number of iterations, we
obtain the parameters $\psi_e^s$, $\psi_d^s$, $\psi_{1}^s$, $\psi_{2}^s$ and $\psi_{3}^s$ as follows:

\[
\hat{\psi}_e^s = \hat{P}(\mu|z) = \frac{\tau_{s,\mu} + \beta_{3}^e}{\sum_{\mu \in G} (\tau_{s,\mu} + \beta_{3}^e)} \quad (7)
\]

\[
\hat{\psi}_d^s = \hat{P}(s|\mu) = \frac{\tau_{p,s} + \beta_{3}^d}{\sum_{s \in S} (\tau_{p,s} + \beta_{3}^d)} \quad (8)
\]

\[
\hat{\psi}_{1}^s = \hat{P}(s|z) = \frac{\tau_{s,\mu} + \rho_{1}}{\sum_{s \in S} (\tau_{s,\mu} + \rho_{1})} \quad (9)
\]

\[
\hat{\psi}_{2}^s = \hat{P}(s|t) = \frac{\sum_{\mu \in C_{\mu}} \sum_{s \in S} (\tau_{s,\mu} + \eta_{1})}{\sum_{\mu \in C_{\mu}} \sum_{s \in S} (\tau_{s,\mu} + \eta_{1})} \quad (10)
\]

\[
\hat{\psi}_{3}^s = \hat{P}(s|t) = \frac{\sum_{\mu \in C_{\mu}} \sum_{s \in S} (\tau_{s,\mu} + \eta_{2})}{\sum_{\mu \in C_{\mu}} \sum_{s \in S} (\tau_{s,\mu} + \eta_{2})} \quad (11)
\]

After determining the above estimation of parameters, we will obtain the final decision of group $G$ on each candidate service $s$ via combining all of users’ decision according to (7)—(11), which is computed as follows:

\[
P(s|\mu, G) = \prod_{\mu \in G} \sum_{z \in Z} \theta_{G,z} \cdot \hat{\psi}_e^s (\lambda_0 \cdot \hat{\psi}_d^s \cdot \lambda_1 \cdot \hat{\psi}_{1}^s \cdot \lambda_2 \cdot \hat{\psi}_{2}^s \cdot \lambda_3 \cdot \hat{\psi}_{3}^s)
\]

where $\lambda_0, \lambda_1, \lambda_2$ and $\lambda_3$ can be computed as follows:

\[
\lambda_0 = \frac{\tau_{p,0}}{\tau_{p,0} + \tau_{p,1} + \tau_{p,2} + \tau_{p,3}} \quad \lambda_1 = \frac{\tau_{p,1}}{\tau_{p,0} + \tau_{p,1} + \tau_{p,2} + \tau_{p,3}} \\
\lambda_2 = \frac{\tau_{p,2}}{\tau_{p,0} + \tau_{p,1} + \tau_{p,2} + \tau_{p,3}} \quad \lambda_3 = \frac{\tau_{p,3}}{\tau_{p,0} + \tau_{p,1} + \tau_{p,2} + \tau_{p,3}}
\]

4.2. Activeness-based coalition ranking strategy

After obtaining the services preferred by a group and each user’s decision on services, i.e., $\psi_{C(\mu)}$, where $C(\mu) = \{u, c, t, z\}$, we should consider the fairness between users, i.e., determine the position of services, which guarantee fairness to each user as much as possible via coalition game theory. Here we consider users’ activeness. Based on the previous discussion, a group contains active users and inactive users, where their activeness contributes to the existence of group (Observation 3). According to social regularization principle [17], when sorting services, services preferred by active users should be in priority considered to rank at a top position.

First, we divide active users and inactive users according to activeness, we assume that the historical behavior of each user is shared with others (e.g., purchased items). To conveniently do experiments, activeness in our works consists of users’ historical services and his social links (e.g., friends in Douban data set). For a group $G = \{\mu_1, \mu_2, \ldots, \mu_m\}$, we use $S_{\mu}$ to denote the historical services purchased by a group $G$. For $\forall \mu \in G$, we get his historical services denoted by $S_{\mu}^t \subset S_{t}$, the proportion of $\mu$’s historical services is computed as follow:

\[
Pro_{\mu}^s = \frac{|S_{\mu}^t|}{|S_{t}|} \quad (13)
\]

After obtaining the services preferred by a group and users’ decision, the social-activeness of $\mu$ is computed as follow:

\[
Act_{\mu} = \frac{Pro_{\mu}^s + Act_{\mu}^l}{2} \quad (15)
\]

After computing the activeness of each user in $G$, we get a sorted order of users. Because each group contains two types of users, the active and the inactive, we divide $G$ into two subgroups by proportion $\varpi$ (defaulted by 20%), $G_{a}$ including $m_{1}$ active users and $G_{a}$ composed of $m_{2}$ inactive users (The effect of $\varpi$ will be discussed in Section 5).

For $\forall \mu \in G_{a}$, we computed the new estimate value for services based on activeness and user’s decision, denoted by $N_{ij}$.

\[
N_{ij} = e^{Act_{i} \cdot \psi_{C(\mu)}} \quad (16)
\]

450 we get total estimate value of $G_{a}$ on $s_{j}$ via calculating the mean value and a sorted services list is determined, denoted by $S_{a}^s = \{s_{1}, \ldots, s_{n}\}$, where $p_{i}$ means that the service is ranked at the $i$th position.

For $\forall \mu \in G_{a}$, we adopt a different strategy to get the ranked list of services because of their lower activeness. It is known that each user has his own decision on $S$ derived by section 4.1, a decision matrix $D_{a} = (\psi_{C(\mu)})_{m \times n}$ is obtained.

First, we convert decision matrix $D$ into ranking matrix $R_{-a}$ via sorting the decision value of each user on $S$, an example of the conversion is shown in Fig.3. Then let $P = \{p_{1}, p_{2}, \ldots, p_{n}\}$ be a sequence of position, we hope the $P$ is approximate to each row in a ranking matrix as soon as possible, which means that we must solve the following unconstrained optimization problem.

\[
\min_{P} F(P, R_{-a}) = \frac{1}{n \cdot m_{2}} \sum_{j=1}^{m_{2}} \sum_{j=1}^{m_{2}} (p_{j} - r_{ij})^{2} \quad (17)
\]

455 We apply stochastic gradient descent method to work out equation (17), and get the solution $P^* = \{p_{1}', p_{2}', \ldots, p_{n}'\}$. Another sequence of service list is obtained, i.e., $S_{a}^s = \{s_{p_{1}'}, s_{p_{2}'}, \ldots, s_{p_{n}'}\}$.
Table 2: Statistics of Data sets

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Epinions</th>
<th>Ciao</th>
<th>Douban</th>
</tr>
</thead>
<tbody>
<tr>
<td># Users</td>
<td>21926</td>
<td>7287</td>
<td>30438</td>
</tr>
<tr>
<td># Services</td>
<td>23863</td>
<td>12028</td>
<td>16277</td>
</tr>
<tr>
<td># Category</td>
<td>26</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td># Groups</td>
<td>8514</td>
<td>2175</td>
<td>6229</td>
</tr>
<tr>
<td># Ratings</td>
<td>498199</td>
<td>148093</td>
<td>359802</td>
</tr>
<tr>
<td># links</td>
<td>300053</td>
<td>57536</td>
<td>88759</td>
</tr>
<tr>
<td>Den.r (%)</td>
<td>0.095</td>
<td>0.16</td>
<td>0.07</td>
</tr>
<tr>
<td>Den.l (%)</td>
<td>0.12</td>
<td>0.21</td>
<td>0.019</td>
</tr>
<tr>
<td>Mem.Range</td>
<td>[2, 1304]</td>
<td>[2, 429]</td>
<td>[2, 326]</td>
</tr>
</tbody>
</table>

Note: 'Mem.Range' represents an interval which reflects the range of group size. 'Den.r' indicates the density on ratings, 'Den.l' indicates the density on trust or friend relationship.

Table 3: Parameters Setting

<table>
<thead>
<tr>
<th>Parameters</th>
<th>α</th>
<th>β₁</th>
<th>β₂</th>
<th>ρ</th>
<th>η₁</th>
<th>η₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.2</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

After getting the two service lists, $S_{order}^u$, $S_{order}^v$, we design another ranking strategy to get the final order of $S$: 1) If $p_k = p_i$, we put $s_p_k$ at the position $p_k$ in $S$. 2) If $p_k \neq p_i$, there must exist $p_j \neq k$, s.t., $s_p_k = s_p_j$. We apply activeness to get the new position $\overline{p_k}$ as follows:

$$\overline{p_k} = \frac{1}{Act_{a} + Act_{a} (Act_{a} \cdot p_k + Act_{a} \cdot p_j)}$$

where $Act_{a}$ and $Act_{a}$ is the minimum activeness in $G_a$ and $G_{-a}$. We put $s_p_k$ at the position $\overline{p_k}$ in $S$. Repeat the above step until a ranked services list $S_{red}$ is finally obtained.

5. Experiments

5.1. Data sets and statistics

To validate the performance, we apply our scheme to three real-world data sets. Table 2 shows the statistics of data sets (items in this section are identical to the services mentioned above).

- **Epinions**: Tang [42] crawled it from a well-known online consumer review site Epinions. On this site, a user writes not only critical reviews for various products but also adds other members to his trusted list if he feels that their reviews are useful to the choice of items (the items are classified into 26 categories).

- **Ciao**: Tang [42] also provides the second data set crawled from Ciao, another famous review site which is similar to Epinions. Items in the Ciao data set are divided into 28 categories.

- **Douban**: The last data set is Douban dataset crawled by Ma [43] from a popular Chinese social networking service website, Douban. It allows registered users to record information and create content related to entities such as film, books, music, and recent events. This dataset contains movie items.

For Epinions and Ciao data sets, we filter out some terms that are rated less than five times and get 23863 items with 489700 ratings, 12028 items with 148093 ratings respectively. For the Douban data set, we sample a subset of Douban dataset which contains 31240 users and 16277 movies.

**How to form group.** Each data set includes social relationships matrix denoted by $T_{mat}$ which is a 0-1 matrix, i.e., if user $u_j$ is socially connected with user $u_i$, $T_{ij}$ is 1, otherwise 0. For $\forall u_i$ recorded in $T_{mat}$, we select users directly connected with $u_i$ and put them into a group $G$, note that there may exists social links between these users excluding $u_i$. Finally, we get 8514, 2175 and 6229 groups corresponding to these data sets respectively. Each group is assumed to be independent during experiments.

5.2. Evaluation methodology

In our experiments, we apply a five-time Leave-One-Out Cross Validation (LOOCV) to evaluate the performance of various schemes. In each run, each data set is split into two subsets, i.e., a training set and a testing set by randomly selecting one of the rated terms for each user and putting it into the testing set. Since it is quite time-consuming to rank all items for each user during evaluation, we followed the common strategy [45, 46] that randomly samples 100 items that are not interacted by the user, ranking the test item among the 100 items, i.e., the testing set of this user contains 101 items. For a given group including $K$ users, the testing set is the union of its inside $K$ users’ testing set, which at most contains $K$ testing items+100*K sampling items. The training set is used to train a model, then for each group, a size-$N$ recommendation list in a descendent sequence is generated via our scheme. In the majority of the results presented in Section V-D, we set $N$ as 5, 10, 15, 20, and 25 to compare the result difference.

The recommendation accuracy and fairness is measured via Hit Rate (HR) and Average Reciprocal Hit Rank (ARHR). HR is computed by

$$HR = \frac{\# \text{ hits}}{\#G}$$

where $\#G$ is the size of group $|G|$, $\# \text{ hits}$ is the number of users who have items in the testing set recommended in the Top-$N$ recommendation list. The second measure for

---

1http://www.cse.msu.edu/~tangjili/trust.html
2http://www.ciao.co.uk

3https://drive.google.com/file/d/1jnRwejx0oenpWkQHsmGLASSq9zlZh8o/view?usp=sharing
evaluation is ARHR, which is defined as follows:

\[
ARHR = \frac{1}{\#[G]} \sum_{i=1}^{\#hits} \frac{1}{p_i}
\]  

(20)

where \( p \) is the position of the item in the ranked recommendation list when an item of a group is hit. ARHR measures the inverse of the position of the recommended item in the recommendation list. In our work, the fairness is also converted to a ranking problem, i.e., the higher the ARHR value is, the more fair the recommended service list will be. Table 3 shows the parameters in PFGR.

5.3. Comparison schemes

To demonstrate the effectiveness, we compare the proposed approach with the following baselines and state-of-the-art schemes.

- **Ave/LM Ranking CF Algorithm** [7]: This algorithm ranks items based on Average/Least Misery relevance and recommends the top-\( k \) items.
- **EFGreedy Algorithm** [6]: This method proposes a fairness metric called proportionality and greedily selects items to maximize fairness.
- **Greedy-LM/Var** [13]: Lin et al. propose this approach using a greedy algorithm for Least Misery/Variance Fairness-aware group recommendation.
- **USRG** [15]: This work proposes an approach based on non-cooperative games to maximize the preference of user in group via determining Nash equilibrium state.
- **COM** [1]: A probabilistic model based on LDA is proposed to model the generative process of group recommendation
- **CrowdRec** [9]: This model is an extension of COM, which is applied in crowd-funding domains.
- **NIGR** [14]: This work aims to find Narch equilibrium during group recommendation with social influence between users consideration.
- **CoaGR** [2]: CoaGR, based on coalition game theory, divides users into several exhaustive and disjoint coalitions to maximize the social welfare function (defined in [2]) of group.
- **GTBT** [28]: GTBT is a game theory-based scheme which is applied to trust evaluation during recommendation.
- **Simple_PFGR (our scheme)**: This scheme neglect the social relationships in PFGR, i.e., \( d \)'s value is only set as 0,1 or 3 during service selection, and \( Act_{R}^{p} \) is set as 0 for fairness evaluation.
- **PFGR (our scheme)**: PFGR with social relationships account combines probabilistic graph and coalition game to maximize the satisfaction when making recommendations to a group.

The comparison schemes are divided into two parts: Schemes without social links account: Ave/LM Ranking algorithm, EFGreedy, Greedy-LM/Var, COM and CrowdRec; Schemes with social links consideration: USRG, CoaGR, NIGR and GTBT. To be fair, we compare with these two kinds of schemes, respectively.

5.4. Results and analysis

In this section, we analyze Top-N recommendation performance of PFGR with other compared schemes on different data sets to answer the following questions:

- How does PFGR compare with state-of-the-art methods (Section 5.4.1)?
- How does PFGR compare with other approaches in different sizes of groups (Section 5.4.2)?
- How does our approach tackle the conflicted preferences case (Section 5.4.3)?
- How does the users' activeness of a group affect the fairness (Section 5.4.4)?
- What's the advantage of our coalition strategy over other game theory-based schemes (Section 5.4.5)?
- How do the parameters applied in our work affect the recommendation performance (Section 5.4.7)?

5.4.1. Overall performance comparisons

Tables 4 and 5 summarize the performance of the state-of-the-art schemes and ours (i.e., Simple_PFGR and PFGR).

In Table 4, all of the approaches don’t consider social links, therefore we assume that no social links exist in the formed group and input information is just rating information of group members. As shown in Table 4, Simple_PFGR significantly improves the HR and ARHR compared with EFGreedy. For other six schemes such as Ave ranking CF, LM ranking CF, Greedy-LM, Greedy-Var, COM, CrowdRec, our scheme attains a maximum increase of 43.01% in HR and 54.75% in ARHR. Compared with the current best scheme COM, PFGR attains higher HR and ARHR with 7.45% and 5.64% increase on average.

In Table 5, our input information includes group members’ rating information and their mutual social relationship. As indicated in Table 5, PFGR outperforms USRG and GTBT because there is more than 80% increase in HR and ARHR value. Compared with NIGR, the best approach based on game theory, PFGR hit higher HR and
Table 4: Overall Comparison on Three Real-world Datasets (without social links)

<table>
<thead>
<tr>
<th>Methods</th>
<th>HR@05</th>
<th>HR@10</th>
<th>HR@15</th>
<th>HR@20</th>
<th>HR@25</th>
<th>ARHR@05</th>
<th>ARHR@10</th>
<th>ARHR@15</th>
<th>ARHR@20</th>
<th>ARHR@25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ave Ranking CF</td>
<td>0.0675</td>
<td>0.1063</td>
<td>0.1709</td>
<td>0.2348</td>
<td>0.3051</td>
<td>0.3056</td>
<td>0.0574</td>
<td>0.0618</td>
<td>0.0653</td>
<td>0.0692</td>
</tr>
<tr>
<td>(129.48%</td>
<td>(128.70%)</td>
<td>(116.93%)</td>
<td>(117.16%)</td>
<td>(119.33%)</td>
<td>(121.46%)</td>
<td>(121.47%)</td>
<td>(122.47%)</td>
<td>(123.04%)</td>
<td>(123.96%)</td>
<td>(125.00%)</td>
</tr>
<tr>
<td>LM Ranking CF</td>
<td>0.0612</td>
<td>0.0917</td>
<td>0.1578</td>
<td>0.2213</td>
<td>0.2931</td>
<td>0.3017</td>
<td>0.0512</td>
<td>0.0557</td>
<td>0.0592</td>
<td>0.0627</td>
</tr>
<tr>
<td>(136.14%)</td>
<td>(130.51%)</td>
<td>(119.07%)</td>
<td>(118.99%)</td>
<td>(114.75%)</td>
<td>(125.43%)</td>
<td>(124.87%)</td>
<td>(124.25%)</td>
<td>(123.03%)</td>
<td>(125.73%)</td>
<td></td>
</tr>
<tr>
<td>EFGreedy (++)</td>
<td>0.0262</td>
<td>0.0415</td>
<td>0.0803</td>
<td>0.1471</td>
<td>0.2057</td>
<td>0.0122</td>
<td>0.0138</td>
<td>0.0151</td>
<td>0.0169</td>
<td>0.0204</td>
</tr>
<tr>
<td>Greedy-LM</td>
<td>0.0736</td>
<td>0.1155</td>
<td>0.1771</td>
<td>0.2243</td>
<td>0.2804</td>
<td>0.0081</td>
<td>0.0102</td>
<td>0.0117</td>
<td>0.0128</td>
<td>0.0140</td>
</tr>
<tr>
<td>(123.80%)</td>
<td>(118.22%)</td>
<td>(112.82%)</td>
<td>(111.74%)</td>
<td>(115.30%)</td>
<td>(111.28%)</td>
<td>(119.50%)</td>
<td>(119.25%)</td>
<td>(117.42%)</td>
<td>(110.61%)</td>
<td></td>
</tr>
<tr>
<td>Greedy-Var</td>
<td>0.0728</td>
<td>0.1214</td>
<td>0.1843</td>
<td>0.2571</td>
<td>0.3290</td>
<td>0.0596</td>
<td>0.0672</td>
<td>0.0701</td>
<td>0.0754</td>
<td>0.0798</td>
</tr>
<tr>
<td>(120.05%)</td>
<td>(112.37%)</td>
<td>(98.41%)</td>
<td>(13.80%)</td>
<td>(19.23%)</td>
<td>(14.61%)</td>
<td>(16.70%)</td>
<td>(15.77%)</td>
<td>(18.40%)</td>
<td>(15.76%)</td>
<td></td>
</tr>
<tr>
<td>COM</td>
<td>0.0702</td>
<td>0.1238</td>
<td>0.1926</td>
<td>0.2596</td>
<td>0.3211</td>
<td>0.0612</td>
<td>0.0669</td>
<td>0.0715</td>
<td>0.0753</td>
<td>0.0798</td>
</tr>
<tr>
<td>(110.35%)</td>
<td>(78.82%)</td>
<td>(13.74%)</td>
<td>(75.97%)</td>
<td>(63.15%)</td>
<td>(67.68%)</td>
<td>(43.62%)</td>
<td>(34.15%)</td>
<td>(19.62%)</td>
<td>(23.37%)</td>
<td></td>
</tr>
<tr>
<td>CrowdRec</td>
<td>0.0815</td>
<td>0.1283</td>
<td>0.1937</td>
<td>0.2641</td>
<td>0.3351</td>
<td>0.0631</td>
<td>0.0673</td>
<td>0.0738</td>
<td>0.0788</td>
<td>0.0842</td>
</tr>
<tr>
<td>(57.24%)</td>
<td>(76.70%)</td>
<td>(11.15%)</td>
<td>(14.17%)</td>
<td>(21.91%)</td>
<td>(21.17%)</td>
<td>(21.36%)</td>
<td>(21.02%)</td>
<td>(21.27%)</td>
<td>(21.37%)</td>
<td></td>
</tr>
<tr>
<td>Simple_PFGR</td>
<td>0.0874</td>
<td>0.1369</td>
<td>0.1998</td>
<td>0.2751</td>
<td>0.3415</td>
<td>0.0651</td>
<td>0.0703</td>
<td>0.0748</td>
<td>0.0796</td>
<td>0.0865</td>
</tr>
</tbody>
</table>

Note: ‘+’ means that the performance of Simple_PFGR exceeds more than 80% compared with other approaches. ‘↑’ means the improvement in accuracy compared with other approaches.

ARHR value, respectively. The higher HR and ARHR values indicate that PFGR can efficiently rank the services for a group in the top position.

Besides, the results also show that: 1) For methods based on greedy algorithm, the recommendation performance of Greedy-Var is better than that of Greedy-LM in total. 2) Methods based on probabilistic graph, i.e., COM and CrowdRec is superior to those methods based on greedy algorithms or ranking in recommendation performance. 3) PFGR is better than Simple_PFGR in HR and ARHR, which indicates that the social relationship can improve recommendation performance (More details are in Section 5.4.4).

In total, our scheme PFGR accomplishes more accuracy recommendation and determines a comparatively satisfied ranked list for groups, which efficiently tackle the fairness issue between users compared with current state-of-the-art.

5.4.2. Recommendation on different size of group

In this section, we discuss the schemes on different group sizes shown in Figs. 3-6. We divide data sets into five categories according to the number of users in a group as shown in Table 6. Here, we set N as 10 and 20. Besides Tables 4 and 5, which shows the remarkable comparison results with EFGreedy, USRG and GTBT, we additionally compare the following schemes in this part, namely Ave Ranking CF, LM Ranking CF, Greedy-LM, Greedy-Var, COM, CrowdRec, CoaGR and NIGR.

From Figs. 3-6 we can acknowledge that 1) Simple_PFGR and PFGR outperforms all the compared schemes for groups with different sizes on the three data sets whether considering social links or not. Our scheme attains the highest values in both HR and ARHR, indicating our schemes has
better recommendation accuracy than other schemes. 2) Compared with Ave Ranking CF and LM Ranking CF, Greedy-LM and Greedy-Var, COM and CrowdRec, CoaGR and NIGR, the maximum increase in HR and ARHR attains 26.51% and 17.84%, respectively. Besides, NIGR, COM and CrowdRec also achieve good recommendation performance on group sizes of two to ten because the values of HR and ARHR hit by these four schemes are quite similar to ours. However, the recommendation performance of them would decrease when group size becomes more substantial.

In summary, our scheme PFGR consistently achieves more accurate results when compared with the state-of-the-art approaches. The results prove that PFGR can produce satisfactory recommendations via effective optimizing the fairness within the groups of users and integrating social trust simultaneously. Our empirical studies also demonstrate that our proposed model has good scalability and suitability when recommending to a larger size of groups.

### Table 5: Overall Comparison on Three Real-world Datasets

<table>
<thead>
<tr>
<th>Methods</th>
<th>Metrics</th>
<th>HR@5</th>
<th>HR@10</th>
<th>HR@15</th>
<th>HR@20</th>
<th>HR@25</th>
<th>ARHR@5</th>
<th>ARHR@10</th>
<th>ARHR@15</th>
<th>ARHR@20</th>
<th>ARHR@25</th>
</tr>
</thead>
<tbody>
<tr>
<td>USRG (+)</td>
<td></td>
<td>0.0428</td>
<td>0.0764</td>
<td>0.1216</td>
<td>0.1622</td>
<td>0.2066</td>
<td>0.2036</td>
<td>0.2025</td>
<td>0.1988</td>
<td>0.1945</td>
<td>0.1939</td>
</tr>
<tr>
<td>NIGR</td>
<td></td>
<td>0.0803</td>
<td>0.1352</td>
<td>0.2006</td>
<td>0.2710</td>
<td>0.3435</td>
<td>0.6443</td>
<td>0.6932</td>
<td>0.7048</td>
<td>0.8900</td>
<td>0.9060</td>
</tr>
<tr>
<td>CoaGR</td>
<td></td>
<td>0.0602</td>
<td>0.1089</td>
<td>0.1609</td>
<td>0.2417</td>
<td>0.3264</td>
<td>0.6772</td>
<td>0.6962</td>
<td>0.7067</td>
<td>0.9769</td>
<td>0.9769</td>
</tr>
<tr>
<td>GTFR (+)</td>
<td></td>
<td>0.0901</td>
<td>0.1402</td>
<td>0.2042</td>
<td>0.2758</td>
<td>0.3526</td>
<td>0.6079</td>
<td>0.7027</td>
<td>0.7083</td>
<td>0.8822</td>
<td>0.8879</td>
</tr>
<tr>
<td>PFGR</td>
<td></td>
<td>0.1204</td>
<td>0.1485</td>
<td>0.2310</td>
<td>0.3573</td>
<td>0.4403</td>
<td>0.6025</td>
<td>0.7013</td>
<td>0.7852</td>
<td>0.8841</td>
<td>0.9035</td>
</tr>
</tbody>
</table>

Note: ‘+’ means the performance of PFGR exceeds more than 80% compared with other approaches. ‘↑’ means the improvement in accuracy compared with other approaches.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Metrics</th>
<th>HR@5</th>
<th>HR@10</th>
<th>HR@15</th>
<th>HR@20</th>
<th>HR@25</th>
<th>ARHR@5</th>
<th>ARHR@10</th>
<th>ARHR@15</th>
<th>ARHR@20</th>
<th>ARHR@25</th>
</tr>
</thead>
<tbody>
<tr>
<td>USRG (+)</td>
<td></td>
<td>0.0447</td>
<td>0.0711</td>
<td>0.1302</td>
<td>0.2066</td>
<td>0.2880</td>
<td>0.5494</td>
<td>0.6077</td>
<td>0.6090</td>
<td>0.6026</td>
<td>0.6051</td>
</tr>
<tr>
<td>NIGR</td>
<td></td>
<td>0.1166</td>
<td>0.1448</td>
<td>0.2124</td>
<td>0.3434</td>
<td>0.4215</td>
<td>0.6988</td>
<td>0.6960</td>
<td>0.7029</td>
<td>0.7062</td>
<td>0.7082</td>
</tr>
<tr>
<td>CoaGR</td>
<td></td>
<td>0.0938</td>
<td>0.1205</td>
<td>0.1795</td>
<td>0.2634</td>
<td>0.3289</td>
<td>0.5785</td>
<td>0.5785</td>
<td>0.5785</td>
<td>0.5785</td>
<td>0.5785</td>
</tr>
<tr>
<td>GTFR (+)</td>
<td></td>
<td>0.0513</td>
<td>0.0852</td>
<td>0.1442</td>
<td>0.2038</td>
<td>0.2766</td>
<td>0.3772</td>
<td>0.4117</td>
<td>0.4406</td>
<td>0.4488</td>
<td>0.4521</td>
</tr>
<tr>
<td>PFGR</td>
<td></td>
<td>0.1201</td>
<td>0.1492</td>
<td>0.2106</td>
<td>0.3306</td>
<td>0.4087</td>
<td>0.5473</td>
<td>0.5957</td>
<td>0.6662</td>
<td>0.7951</td>
<td>0.8885</td>
</tr>
</tbody>
</table>

### Table 6: The Statistics of Group Size

<table>
<thead>
<tr>
<th>Data</th>
<th>Cat1</th>
<th>Cat2</th>
<th>Cat3</th>
<th>Cat4</th>
<th>Cat5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epinions</td>
<td>1</td>
<td>2-9</td>
<td>10-30</td>
<td>30-50</td>
<td>50-100</td>
</tr>
<tr>
<td>Douban</td>
<td>1</td>
<td>2-9</td>
<td>10-30</td>
<td>30-50</td>
<td>50-100</td>
</tr>
</tbody>
</table>

Note: ‘Cat’ is short for category. Cat1 contains groups whose total members are 2-10. The total members of groups in Cat2 are 10-30. For Cat3 and Cat4, the total members are 30-50, and 50-100, respectively. Cat5 contains groups whose total member is larger than 100.

5.4.3. Conflicted preference cases study

In this section, we specially discuss the proposed PFGR in conflicted preferences cases which can’t be well solved in current schemes. Here we first reconstruct the group according to users’ preferences. Compared with Douban, Epinions and Ciao data sets contain the categories about services, i.e., games, books, movies and so on. Actually, in
Because of the space limitation, here we firstly randomly simulate the conflicted preferences scenario, we randomly divide users’ preferences into two parts, denoted by $pre_A$ and $pre_B$, where no overlapped users are both between $pre_A$ and $pre_B$. More details about these 5 groups are shown in Table 7. All parameters setting in this section are default the same as section 5.4.2. As shown in Fig.7, we know that 1) both PFGR and Simple_PFGR hit the maximum value in ARHR value, which indicates that our scheme can efficiently solve the fairness when confronted with conflicted preferences. 2) Compared with schemes based greedy algorithm such as Greedy_LM/Ave, LDA-based approaches, i.e., COM and CrowdRec can solve fairness better when confronted with conflicted preferences.
but with a slight improvement. 3) Social links can also help to solve fairness issue because of PFGR hitting larger ARHR than SimplePFGR. To conclude, our schemes have advantage in solving fairness issue under the conflicted cases.

### 5.4.4. Fairness evaluation

In this section, we mainly discuss the game mechanism which is applied to guarantee fairness, i.e., determining a sequence of services that can satisfy the preference of users as much as possible. In other words, if more users are content with the ranking service, the value of HR and ARHR will become larger. Fairness in our work is related with users’ activeness. In our model, a group is composed of active users and inactive users based on Observation

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 7: The Statistics of Group

<table>
<thead>
<tr>
<th>Epinions</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td></td>
</tr>
<tr>
<td>Group2</td>
<td></td>
</tr>
<tr>
<td>Group3</td>
<td></td>
</tr>
<tr>
<td>Group4</td>
<td></td>
</tr>
<tr>
<td>Group5</td>
<td></td>
</tr>
</tbody>
</table>

Note: there is no overlapped preference between preA and preB, while users' preferences are overlapped in preA or preB.

![Figure 8](image)

(a) HR on Epinions  
(b) ARHR on Epinions

![Figure 9](image)

(c) HR on Ciao  
(d) ARHR on Ciao  
(e) HR on Douban  
(f) ARHR on Douban

Figure 8: The proportion $\omega$ of active users in group

Figure 9: Comparison between Different Game Strategies

3. The proportion of active users depends on $\omega$ which is defaulted as 20%, e.g., If a group contains 100 users, the number of active users is 20. Here we vary the value of $\omega$ from 10% to 50% to observe the effect on fairness brought by activeness (Here we set $N$ as 10, 15, 20).

As shown in Fig.8, there is an increase in HR and ARHR with the variation of $\omega$, which means that if the activeness of users is considered, the recommendation performance will be enhanced, i.e., more users are satisfied with the ranking services. In addition, we find that when the value of $\omega$ is larger than 30%, the tendency of increase in HR and ARHR becomes gentler because the slope from 30% to 50% is smaller than that between 10% and 30%. This result shows that the recommendation performance will remain stable with the increase in the number of active users. Moreover, compared with other approach relevant with fairness issue, e.g., [13], PFGR achieves better recommendation performance in HR and ARHR shown in Table 4 and Figs 4, 5 and 6.

### 5.4.5. The analysis of our coalition strategy

In this section, we mainly discuss the efficiency of our proposed activeness-based coalition strategy (ACG) when tackling fairness. To validate the advantage of our coalition strategy, we select another two ubiquitous game theory-based strategies for comparison, i.e., Non-Cooperative game strategy (NonCG) [15] and preference-based coalition strategy (PCG) [2]. Note that these two game strategies are the part of USRG [15] and CoaGR [2].

Our PFGR contains two parts: the first part is probabilistic graph-based model which is designed to select the services preferred the groups, while the second part is
In Fig.10, the comparison is significant because the height difference between Simple_PFGR and PFGR is quite evident. On average, PFGR has more than 15% increase in HR and ARHR compared with Simple_PFGR. The same conclusion can also be drawn on the Ciao and Douban data sets; 2) MAE values shown in Tables 8-13 are slight, where the maximum is less than 0.4%. The MAE Values indicate that the impact of the variation of parameters on RSs is slight, which validates the robustness of PFGR.

### 5.4.7. Parameters effect

In this section, we investigate the effect of parameters recorded in Table 3. We conduct experiments on parameters using the control variable method. The control variable method is a scientific method that keeps one parameter changeable while other parameters hold unchanged during experiments. Here, we set $N$ as 10.

Tables 8-13 show the effect of parameters by varying $\alpha$, $\beta_1$, $\beta_2$, $\rho$, $\eta_1$ and $\eta_2$ from 0.2 to 0.8. We can summarize the following from the results: 1) Different parameter values lead to different HR and ARHR values. For example, Epinions attains the highest HR and ARHR values (i.e., 0.1290 and 0.0760 shown in Table 6) when $\alpha=0.6$, while get different highest HR and ARHR values (i.e., 0.1307 and 0.0740) at $\beta_1=0.8$ and $\beta_2=0.2$. The same conclusion can also be drawn on the Ciao and Douban data sets; 2) MAE values shown in tables are slight, where the maximum is less than 0.4%. The MAE Values indicate that the impact of the variation of parameters on RSs is slight, which validates the robustness of PFGR.

### 6. Conclusions

In this paper, we mainly study the fairness problem in group recommendation based on probabilistic graph model and coalition game and propose a novel approach called PFGR which can achieve higher recommendation performance with fairness account. The proposed approach first selects the services satisfied the preferences of a group via modelling the selection behavior of users according to several observations existing in the real world. After determining the services, PFGR further considers users’ activeness and designs a sorted strategy based on coalition game to suggest a ranked recommendation list which can maximize all the members’ preference (i.e., fairness). Our experimental results show that PFGR outperforms...
### Table 11: The effect of $\rho$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Epinions</th>
<th>Ciao</th>
<th>Douban</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.1314</td>
<td>0.0706</td>
<td>0.1432</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1252</td>
<td>0.0658</td>
<td>0.1402</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1246</td>
<td>0.0649</td>
<td>0.1429</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1269</td>
<td>0.0640</td>
<td>0.1457</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1253</td>
<td>0.0648</td>
<td>0.1418</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1221</td>
<td>0.0632</td>
<td>0.1436</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1207</td>
<td>0.0674</td>
<td>0.1411</td>
</tr>
<tr>
<td>MAE</td>
<td>0.23%</td>
<td>0.18%</td>
<td>0.20%</td>
</tr>
</tbody>
</table>

### Table 12: The effect of $\eta_1$

<table>
<thead>
<tr>
<th>$\eta_1$</th>
<th>Epinions</th>
<th>Ciao</th>
<th>Douban</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.1249</td>
<td>0.0683</td>
<td>0.1460</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1218</td>
<td>0.0667</td>
<td>0.1411</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1206</td>
<td>0.0659</td>
<td>0.1448</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1207</td>
<td>0.0634</td>
<td>0.1401</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1255</td>
<td>0.0621</td>
<td>0.1411</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1253</td>
<td>0.0645</td>
<td>0.1438</td>
</tr>
<tr>
<td>MAE</td>
<td>0.40%</td>
<td>0.18%</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

### References


