Disparity-Aware Group Formation for Recommendation (Extended Abstract)

Lin Xiao
Institute of Interdisciplinary Information Sciences
Tsinghua University
Beijing, China
jackielinxiao@gmail.com

Zhang Min
Department of Computer Science and Technology
Tsinghua University
Beijing, China
z-m@tsinghua.edu.cn

Zhang Yongfeng
College of Information and Computer Science
University of Massachusetts
Amherst, MA 01003, USA
yongfeng@cs.umass.edu

Gu Zhaoquan
Department of Computer Science
HongKong University
HongKong, China
demin456@gmail.com

ABSTRACT

Group recommendation has attracted significant research efforts for its importance in benefiting a group of users, however, seldom investigation has been put into the essential problem of how the groups should be formed. This paper investigates the disparity-aware group formation problem in group recommendation. In this work, we present a formulation of the disparity-aware group formation problem, and further show its NP-Hardness. For the case when group satisfaction is maximized, we propose a cutting plane algorithm based on bilinear program that achieves a ε approximation to the optimum. For the general case, we design an efficient optimization algorithm based on Projected Gradient Descent and further propose a simplified swapping alike algorithm that accommodates to large datasets. We conduct extensive experiments on both simulated and real-world datasets. Experimental results verify that the performance of our algorithm is close to the optima. More importantly, our work reveals that proper group formation can lead to better performances of group recommendation in different scenarios. To our knowledge, we are the first to study the group formation problem with disparity awareness for recommendation, and more promising works are expected.

1. DISPARITY-AWARE GROUP FORMATION

In this section, we formulate the Disparity-Aware Group Formation (DAGF) problem. We first give some introductions about the semantics adopted in group recommendation problems from [1]. We assume the individual preference of an individual user i on item j is depicted as a number $R_{ij} \in [R_{min}, R_{max}]$.

DEFINITION 1. *Group Satisfaction: Given an item j and a group of users U, the satisfaction score $Sc(U,j)$ of the group given the item recommended to them is defined as a function in $[R_{min}, R_{max}]$: $Sc(U,j) = \sum_{i\in U} \frac{1}{|U|} R_{ij}$.*

DEFINITION 2. *Group Disparity: Given an item j and a group of users U, the disparity $D(U,j)$ of the group on item j is defined as a function in $[R_{min}, R_{max}]$: $D(U,j) = \frac{1}{|U|}\sum_{i\in U} |R_{ij} - \sum_{i\in U} \frac{1}{|U|} R_{ij}|$.*

Since most recommender systems follow the Top-K recommendation, the Top-K items with high satisfaction and low disparity are recommended to each group in our work. We set variables X_{ig} and Y_{jg} as indicator variables deciding whether user i is in group g and item j is recommended to group g respectively. Based on this, the Disparity-Aware Group Formation (DAGF) problem is rewritten into an integer programming:

$$\begin{aligned}
&\max .\ \omega \sum_{g=1}^{G}\sum_{i\in U}\sum_{j\in I} R_{ij}X_{ig}Y_{jg} + \\
&\quad (\omega-1)\sum_{g=1}^{G}\sum_{i\in U}\sum_{j\in I} |R_{ij} - \frac{\sum_{i\in U} X_{ig}R_{ij}}{\sum_{i\in U} X_{ig}}| X_{ig}Y_{jg} \\
&s.t. \\
&\quad \sum_{g=1}^{G} X_{ig} = 1, \forall i \in U;\ \sum_{j\in I} Y_{jg} = K, \forall g \in [1,G] \\
&\quad X_{ig} = \{0,1\}, \forall i,g;\ Y_{jg} = \{0,1\}, \forall j,g
\end{aligned} \tag{1}$$

Based on the intuition of Projected Gradient Descent, we propose a swapping alike algorithm. We introduce a simple yet effective projection method for the problem which acts like a swapping between groups.

Denote L as the objective function, $L_i = [\frac{\partial L}{\partial X_{i1}}, ..., \frac{\partial L}{\partial X_{ig}}, ...], \forall i \in U$; $s_i(g)$ as the projection of $\frac{\partial L}{\partial X_{ig}}$ and $s_j(g)$ as the projection of $\frac{\partial L}{\partial Y_{jg}}$. We relax the requirement of objective function so that the computed gradient is a descent direction for the objective, i.e. $L_i \cdot s_i \geq 0$ and we have the following constraint set (without objective functions):

$$s.t. \sum_{g} s_i(g) = 0,\ L_i \cdot s_i \geq 0,\ \text{and} \begin{cases} s_i(g^p) \leq 0, \forall X_{ig^p} = 1 \\ s_i(g^n) \geq 0, \forall X_{ig^n} = 0 \end{cases} \tag{2}$$

Appears in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), Das, Durfee, Larson, Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*
Copyright © 2017, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

Algorithm 1 DISPARITY AWARE GROUP OPTIMIZATION

Input: Rating matrix R, the set of users U and items I

Output: Formed groups: $X_{ig}, \forall i \in U, g$; $Y_{jg}, \forall j \in I, g$

```
Initialize the group indicators of users and items: X_ig, ∀i ∈ U, Y_jg, ∀j ∈ I;
while |F^t − F^{t+1}| ≤ ε OR iter<MaxIter do
  for each group g do
    Calculate the Top-K items of group g: S(g,K) = {j|Y_jg = 1, ∀j ∈ I};
  end for;
  for each user i do
    for each group g do
      Calculate the gradient L_i(g) as Eq. 3
    end for;
    Assign the user to g = max_{g∈[1,G]}{L_i(g), ∀g};
  end for;
end while
```

where

$$L_i(g) \approx \omega \sum_{j \in I} R_{ij} Y_{jg} + (\omega - 1) \sum_{j \in I} \sum_{i \in U} |R_{ij} - \frac{\sum_{i \in U} R_{ij} X_{ig}}{\sum_{i \in U} X_{ig}}| Y_{jg} \tag{3}$$

This new sub-problem has a simple solution. When $X_{ig^p} = 1$ and $L_i(g^p) \neq \max\{L_i(g)\}$:

$$s_i(g) = \begin{cases} 1, L_i(g) = \max\{L_i(g)\} \\ -1, X_{ig} = 1 \\ 0, \text{otherwise} \end{cases} \tag{4}$$

Otherwise, we have $s_i = 0$.

Judging from the derivation, the main idea of our swapping procedure is to swap users between groups. For a given group formation, we first calculate the Top-K recommended items in each group. Suppose that the items are fixed, we find those users who can obtain higher ratings of Top-K items if swapped into other groups. For those users, we finally swap them into the group where they can get the highest increase of objective function. We repeat the swapping procedure until no user can get higher increase on objective function by swapping. The detailed specification of the algorithm is presented in Alg. 1.

2. EXPERIMENT

2.1 Experiment Settings

The real-world datasets are chosen from MovieLens, Filmtrust and Ciao. For ML-10M (MovieLens-10M, released by MovieLens) and Ciao, We choose 10000/5000 (from ML-10M and Ciao respectively) users who rated most items and 2000 items that get rated by most users. The ratings of these datasets take values from 1 to 5 and the missing entries are estimated with state-of-the-art Collaborative Filtering method (similar processing has been conducted in [1] and [4]) **PMF** (Probabilistic Matrix Factorization) [3]. Moreover, we generate a Randomly Generated DataSet (RD) with the same size of Filmtrust, and the ratings follow a uniform distribution in the interval $[0, 5]$.

Algorithms for Comparison: We compare our approaches with some state-of-art approaches, including **GRD[4]**, **SCC[2]**, **BCC[5]** and **KTD[4]**.

Evaluation Metrics: The first metric is the Average Fulfilment (AF) of users:

$$AF = \frac{\sum_i \sum_{j \in I_g} R_{ij}}{\sum_i \sum_{j \in I(i,K)} R_{ij}} \tag{5}$$

Table 1: AF and AD when $w = 0.8$, $G = 10$, $K = 10$

Metrics	Average Fulfilment				Average Disparity			
Dataset	RD	F.T.	ML	Ciao	RD	F.T.	ML	Ciao
GRD	0.525	0.818	0.841	0.848	1.238	0.625	0.544	0.359
SCC	0.582	0.929	0.850	0.903	1.187	0.491	0.515	0.369
KTD	0.590	0.912	/	/	1.175	0.490	/	/
BCC	0.575	0.894	0.853	0.887	1.184	0.521	0.459	0.346
DAGO	**0.653***	**0.942***	**0.893***	**0.921***	**1.043***	0.501	**0.457***	0.350
PGD	**0.662***	**0.951***	/	/	**1.001***	0.498	/	/

Table 2: AF and AD when $w = 0.2$, $G = 10$, $K = 10$

Metrics	Average Fulfilment				Average Disparity			
Dataset	RD	F.T.	ML	Ciao	RD	F.T.	ML	Ciao
GRD	0.514	0.689	0.744	0.787	1.193	0.269	0.098	0.181
SCC	0.558	0.844	0.761	0.849	1.101	0.327	0.098	0.208
KTD	0.568	0.814	/	/	1.096	0.266	/	/
BCC	0.548	0.805	0.779	0.849	1.096	0.288	0.076	0.211
DAGO	**0.606***	**0.850***	0.773	**0.870***	**0.900***	**0.198***	**0.036***	**0.179***
PGD	**0.611***	**0.853***	/	/	**0.886***	**0.192***	/	/

I_g denotes the set of items recommended to the group g; $I(i, K)$ denotes the set of K items with highest ratings from user i.

The second metric is the Average Disparity (AD) of the users:

$$AD = \frac{\sum_g \sum_{j \in I_g} |U_g| D(U_g, j)}{K \times \sum_g |U_g|} \tag{6}$$

which evaluates the disparity between users inside same groups on the recommendation.

Intuitively, AF evaluates the ratio of user ratings on the recommended items in their group against the ratings of their favourite items. Note that the optimal solution can never gain higher ratings than the sum of all the ratings of each user's favourite items, as a result, $AF \leq 1$; Therefore, higher AF and lower AD are expected.

2.2 Quality Analysis under AF&AD Metric

The performances of the algorithms under the metrics of AF and AD are summarized in Table 1 and Table 2, where the settings we choose are $\omega = 0.8$ and $\omega = 0.2$ (for different levels of trade-off between satisfaction and disparity), $G = 10$, and $K = 10$. In the tables, those values with stars have passed the significance test on the level of $p < 0.01$. Notice that PGD and KTD does not fit for large datasets, we list their performances on RGDS and FilmTrust.

From the results in the tables, we see that our algorithm has a remarkable better performance than the other benchmark algorithms on almost all datasets. Besides, our algorithm achieves not only better overall satisfaction, but also relatively lower disparity. Notice that when $\omega = 0.8$, the objective leans towards maximizing the satisfaction rather than minimizing the disparity, DAGO achieves **highest** AF on all datasets and also induces **low** disparity; when $\omega = 0.2$, the objective leans towards minimizing the disparity rather than maximizing the satisfaction, DAGO induces **lowest** disparity on all datasets and also achieves **high** satisfaction. This indicates that DAGO has a good flexibility in accordance with the value of ω and outperforms other approaches given different specified objectives (determined by the value of ω).

The average fulfilment is the ratio between users' satisfaction in the formed groups and the satisfaction achieved from their ideal Top-K items, which is by definition lower than that of the optimal solution (AF=1), because users have to make compromises when they are in a group with other users.

REFERENCES

[1] S. Amer-Yahia, S. B. Roy, A. Chawlat, G. Das, and C. Yu. Group recommendation: Semantics and efficiency. *Proc. VLDB Endow.*, 2(1), Aug. 2009.

[2] I. S. Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. In *KDD '01*, New York, NY, USA, 2001. ACM.

[3] R. Salakhutdinov and A. Mnih. Bayesian probabilistic matrix factorization using markov chain monte carlo. In *ICML '08*, New York, NY, USA, 2008. ACM.

[4] L. V. L. Senjuti Basu Roy and R. Liu. From group recommendations to group formation. In *SIGMOD '15*, New York, NY, USA, 2015. ACM.

[5] H. Shan and A. Banerjee. Bayesian co-clustering. In *ICDM '08*, Washington, DC, USA, 2008. IEEE Computer Society.