

A cross-domain group recommender system with a generalized aggregation strategy

Ruxia Liang^{1,2}, Qian Zhang^{1*}, Jie Lu¹, Guangquan Zhang¹

¹*Decision Systems and e-Service Intelligence Laboratory*

¹*Centre for Artificial Intelligence, University of Technology Sydney, Australia*

²*School of Business, Central South University, Changsha 410083*

**E-mail: Ruxia.Liang@uts.edu.au,*

Qian.Zhang-1@uts.edu.au,

Jie.Lu@uts.edu.au, Guangquan.Zhang@uts.edu.au

Developing group recommender systems has been a vital requirement due to the prevalence of group activities. However, existing group recommender systems still suffer from data sparsity problem because they rely on individual recommendation methods with a predefined aggregation strategy. To solve this problem, we propose a cross-domain group recommender system with a generalized aggregation strategy in this paper. A generalized aggregation strategy is developed to build group profile in the target domain with the help of individual preferences extracted from a source domain with sufficient data. By adding the constraints between the individual preference and the group profile, knowledge is transferred to assist in the group recommendation task in the target domain. Experiments on a real-world dataset justify the effectiveness and rationality of our proposed cross-domain recommender systems. The results show that we increase the accuracy of group recommendation on different sparse ratios with the help of individual data from the source domain.

Keywords: recommender system, group recommender system, cross-domain recommender system, collaborative filtering

1. Introduction

Recommender system has attracted great attention from academics and industries due to its ability of handling information overload problem¹. Many online service systems, such as Amazon, Google, and Yelp are increasingly adopting recommender systems. The increasing group activities in various online service systems, such as trips, parties and family activities, have triggered the development of group recommender systems². Although extensive research have been conducted to improve performance of individual recommender systems, investigation in group recommender systems is still lacking. The recommendation task is different when the users change from

individuals to a group. The preference of the group is modeled to generate a group profile, which depends on the preference of individuals and the composition of the group. In this way, the strategies of group profile generation are very important in group recommender systems.

Group recommender systems usually contain two necessary phases: group profile generation phase and recommendation phase³. In the group profile generation phase, several efficient aggregation strategies to fuse preferences of group members, such as average, least misery, maximum satisfaction are developed. In the recommendation phase, individual recommendation techniques are utilized, such as memory-based collaborative filtering (CF)³ and latent factor model-based CF⁴, or hybrid methods⁵. Similar to recommender systems for individuals, data sparsity is also a challenging issue in group recommender systems. Since the history records of groups are even less than those of individuals, the data sparsity problem in group recommender systems is more severe.

Previous studies attempted to tackle the group rating sparsity problem by introducing additional information, such as social information⁶, trust^{7,8}, and content information^{9,10} to depict member interaction or personality. Unfortunately, such additional information is not always available. Another solution is to transfer knowledge from a domain where relatively sufficient data is collected, which is known as cross-domain recommender systems. There is still a gap between cross-domain recommender systems and group recommender systems. How can the group recommender systems benefit from a source domain with sufficient data remains a problem to be solved.

In this paper, we propose a cross-domain group recommender system (CDGRS) with a generalized aggregation strategy. For each group member, their individual preferences are modeled from a source domain with sufficient data. We integrate the group profile generation and the recommendation into one unified matrix factorization (MF) method with constraints between group profile and individual preferences. Different aggregation strategies to build group profile are integrated to provide decision-makers with more flexibility.

The contributions of this paper are as follows:

- A generalized aggregation strategy is proposed to build group profile with individual preferences. Several aggregation strategies are integrated, therefore offering flexibility for decision-making in different group recommendation scenarios.
- A CDGRS with a generalized aggregation strategy is developed

which first transfers knowledge from a source domain to assist group recommendation in a target domain, thus alleviating the sparsity problem in group recommender systems.

- Experiments and comparison analysis are conducted on a real-world dataset. The results show that our proposed CDGRSs can increase the accuracy of recommendation compared with group recommender systems built on one domain.

The rest of this paper is as follows. Section 2 presents some notations and preliminaries related to our proposed CDGRS, and constructs a CDGRS. Experiments and comparison analysis are presented in Section 3. Finally, Section 4 concludes this paper.

2. A Cross-domain Group Recommender System

In this section, we first present several frequently used notations and preliminaries of MF. Afterwards, we present the details of our proposed CDGRS.

2.1. Notations and Preliminaries

Suppose we have S groups and H items in one domain, we denote the interactions between groups and items as $\mathbf{G} \in \mathbb{R}^{S \times H}$ (bold uppercase letter represents a matrix). If each group expresses their preferences through explicit ratings, such as $\{1, 2, 3, 4, 5, ?\}$ (“?” means a missing value), the group-item rating matrix \mathbf{G} is denoted as $[g_{l,k}]_{S \times H}$. Based on the biased MF model¹¹, an unknown group rating $\hat{g}_{l,k}$ can be estimated by

$$\hat{g}_{l,k} = \eta + b_l^g + b_k^g + (\mathbf{p}_l^g)^T \mathbf{q}_k^g \quad (1)$$

where η indicates the overall average rating of groups, and the parameters b_l^g and b_k^g indicate the observed deviations of group g_l and item v_k , respectively. The vectors $\mathbf{p}_l^g \in \mathbb{R}^K$, and $\mathbf{q}_k^g \in \mathbb{R}^K$ are the group and item latent factors, which are in a common latent factor feature space of a lower dimensionality K .

The recommendation task is to predict the missing values in the rating matrix based on the groups’ historical rating records. Since there are many missing values in the rating matrix \mathbf{G} , directly factorizing the highly sparse rating matrix is proven to overfitting. To tackle this problem, regularization is introduced to the model. Generally, the optimization problem is

$$J = \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k})^2 + \lambda_1 (\|\mathbf{p}_l^g\|_F^2 + \|\mathbf{q}_k^g\|_F^2 + (b_l^g)^2 + (b_k^g)^2) \quad (2)$$

The parameter $\lambda_1 \geq 0$, and $g_{l,k}$ indicates the real group rating on item v_k . The optimization problem can be solved by gradient descent.

2.2. Cross-domain Group Recommender System

To our best knowledge, there is still no research that combines cross domain recommender systems with group recommendation. Based on the unified MF model in Eq. (2), we construct a CDGRS by extracting knowledge from individual user ratings. The cost function of our proposed CDGRS is as follows:

$$J_{CDGRS}(\theta) = \sum_{(i,j) \in K} (r_{i,j} - \hat{r}_{i,j})^2 + \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k})^2 + \lambda_1 \sum_{l=1}^s \|f(\mathbf{p}_i) - \mathbf{p}_l^g\|_F^2 + \lambda_2 (\|\mathbf{p}_i\|_F^2 + \|\mathbf{q}_j\|_F^2 + b_i^2 + b_j^2) + \lambda_3 (\|\mathbf{p}_l^g\|_F^2 + \|\mathbf{q}_k^g\|_F^2 + (b_l^g)^2 + (b_k^g)^2) \quad (3)$$

where $\hat{r}_{i,j} = \mu + b_i + b_j + \mathbf{p}_i^T \mathbf{q}_j$, which represents the predicted rating of individual user μ_1 on item i_j . Similar to Eq. (2), μ indicates the overall average ratings of individual users, and b_i and b_j are the user and item biases. The vectors $\mathbf{p}_i \in \mathbb{R}^K$ and $\mathbf{q}_j \in \mathbb{R}^K$ are the user and item latent factors, respectively. Particularly, $f(\cdot)$ is an aggregation function, such as average (f_{avg}), minimum (f_{min}) and power average (f_{pa}) operator. $\|\cdot\|_F$ denotes the Frobenius norm.

The parameter set $\theta = \{\eta, \mathbf{p}_i^g, \mathbf{q}_k^g, b_l^g, b_k^g\}$ can be learned for group recommendation from the optimization problem $\min \frac{1}{2} J_{CDGRS}(\theta)$. Gradient descent can be used to solve the proposed system, and the cost function is non-increasing under the following update rules:

$$\mathbf{p}_l^g: = \mathbf{p}_l^g + \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k}) \times \mathbf{q}_k^g + \frac{\partial J_3(\theta)}{\partial \mathbf{p}_l^g} - \lambda_3 \mathbf{p}_l^g \quad (4)$$

$$\mathbf{q}_k^g: = \mathbf{q}_k^g + \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k}) \times (\mathbf{p}_l^g)^T - \lambda_3 \mathbf{q}_k^g \quad (5)$$

$$b_l^g: = b_l^g + \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k}) - \lambda_3 b_l^g \quad (6)$$

$$b_k^g: = b_k^g + \sum_{(l,k) \in H} (g_{l,k} - \hat{g}_{l,k}) - \lambda_3 b_k^g \quad (7)$$

where $J_3(\theta) = \frac{1}{2} \lambda_1 \sum_{l=1}^s \|f(\mathbf{p}_i) - \mathbf{p}_l^g\|_F^2$.

Generally, there are many types of predefined aggregation strategies to model group profile, including the most widely used strategies such as average, the least misery, the maximum satisfaction, and the expertise strategy. Since the function $f(\cdot)$ is a generalized form, our proposed system is a generalized CDGRS which can be reduced to various forms with different aggregation strategies. For example, $f_{avg}(\mathbf{p}_i) = \sum_{i \in g_l} \frac{1}{|g_l|} \mathbf{p}_i$ and $f_{min}(\mathbf{p}_i) = Min(\mathbf{p}_i)$, and $f_{max}(\mathbf{p}_i) = Max(\mathbf{p}_i)$.

Finally, a group rating predicted can be determined using Eq. (2), and a group recommendation list can be generated according to the *top*-K highest scores of $\hat{g}_{l,k}$.

3. Experiments and Analysis

This section conducts experiments on a real world data set **CAMRa2011**¹ using our proposed CDGRS and some baselines. Then, the experimental results and comparison analysis are presented to conclude this section.

3.1. Dataset Description and Metrics

CAMRa2011 is a real-world movie rating dataset containing both individual user and group's rating information. We aim to recommend movies for the groups in this dataset. However, the group rating domain suffer from serious data sparsity problem. To deal with this issue, We try to extract knowledge from individual user ratings to assist group recommendation, and the modified user ratings are treated as the source domain. To testify the impact of sparse ratio of group rating on the group recommendation performance, we randomly extracted three subsets of groups and items to form three group rating datasets with three different sparse ratios, *target1*, *target2*, and *target3*. Finally, we normalized all the ratings in both domains to a rating scale from 1 to 5. The details of the final datasets are summarized in Table 1.

To evaluate the accuracy of our proposed CDGRSs and baselines, we use mean absolute error (MAE) and root mean square error (RMSE)¹² as the evaluation metrics.

3.2. Baselines

To justify the effectiveness and superiority of our proposed CDGRS, we compare it with three existing baselines as below.

¹<http://2011.camrachallenge.com/2011>

Table 1. Statistical information on the **CAMRa2011** dataset.

	source domain	<i>target1</i>	<i>target2</i>	<i>target3</i>
# User/Group	1644	289	289	289
# Item	7792	1701	1701	1701
# Rating	146408	5445	3925	1965
Sparse ratio	98.87%	99.6%	99.2%	99%
Average	3.6838	3.8124	2.1803	2.1306

- **SVD**¹¹. This method is the basic model applied by our proposal.
- **Pearson user-based CF**¹³. This method is a classical memory-based CF approach which utilizes the Pearson correlation coefficient to calculate user similarities
- **After factorization (AF)**¹⁴. This method computes the group profile by merging the latent factors of the users that belong to the group. Then, group recommendation is generated by calculating the inner product between group profile and every item vector.

3.3. Experiments Results

Tables 2 and 3 show the comparison results of our CDGRS with other baselines. To save space, we only compare our CDGRS with three aggregation strategies with the baselines. As shown in the last three rows of Tables 2 and 3, the three proposed CDGRS systems are denoted as *CDGRS_avg*, *CDGRS_max*, and *CDGRS_min*, respectively. We can see that our proposed system performs better than the baselines. When the sparse ratio varies in the target group domain, the performance of the CDGRS with different aggregation strategies is also different. It is reasonable because different strategies cater for different decision scenarios. Therefore, our proposed CDGRS provides decision-makers with more flexibility through integrating different aggregation strategies into one system. In addition, the Pearson user-based CF failed or can only generate a very small portion of group rating predictions due to the particularly sparsity problem of our group domain. Overall, our proposed CDGRS system demonstrates superior accuracy and flexibility even when the group domain is very sparse.

4. Conclusions and Future Work

This paper presents a CDGRS with a generalized aggregation strategy. It has two advantages compared with extant research. First, previous group recommender systems can only consider one aggregation strategy at a time,

Table 2. MAE prediction performance for group recommendation.

	<i>target1</i>	<i>target2</i>	<i>target3</i>
SVD	0.6767 \pm 0.0136	0.4949 \pm 0.0096	0.3298 \pm 0.0073
AF	1.6892 \pm 0.0141	0.4286 \pm 0.0087	1.7574 \pm 0.0042
CDGRS_ave	0.6073 \pm 0.0154	0.3177 \pm 0.0075	0.3204 \pm 0.0055
CDGRS_max	0.6222 \pm 0.0150	0.3158 \pm 0.0079	0.3169 \pm 0.0054
CDGRS_min	0.6211 \pm 0.0104	0.3160 \pm 0.0070	0.3175 \pm 0.0050

Table 3. RMSE prediction performance for group recommendation.

	<i>target1</i>	<i>target2</i>	<i>target3</i>
SVD	0.8290 \pm 0.0174	0.6315 \pm 0.0085	0.4402 \pm 0.0153
AF	1.7401 \pm 0.0138	1.4897 \pm 0.0084	1.8074 \pm 0.0035
CDGRS_ave	0.7316 \pm 0.0166	0.4220 \pm 0.0097	0.4166 \pm 0.0063
CDGRS_max	0.7459 \pm 0.0162	0.4204 \pm 0.0099	0.4137 \pm 0.0067
CDGRS_min	0.7451 \pm 0.0170	0.4195 \pm 0.0090	0.4145 \pm 0.0056

this study integrates a generalized aggregation strategy into a unified MF method. Therefore, our proposed CDGRS provides decision-makers with more flexibility to choose the appropriate aggregation strategy according to practical decision scenarios. Second, this proposed CDGRS leverages the auxiliary information from individual user preferences to assist group recommendation in the target domain with insufficient data. Experiments are conducted on a real-world movie rating dataset. The results show that our proposed system performs better than three other baselines. In the future, we will try to extract knowledge from other domains, such as books and musics, to facilitate the group recommendation for movies.

Acknowledgments

This work was supported by the Australian Research Council (ARC) under Discovery Grant [DP170101632].

References

1. C. Zhang, H. Zhang and J. Wang, Personalized restaurant recommendation method combining group correlations and customer preferences, *Information Sciences* **454**, 128 (2018).
2. D. Cao, X. He, L. Miao, Y. An, C. Yang and R. Hong, Attentive group recommendation, *41st International ACM SIGIR Conference on*

- Research and Development in Information Retrieval, SIGIR 2018*, 645 (2018).
3. W. Wang, G. Zhang and J. Lu, Member contribution-based group recommender system, *Decision Support Systems* **87**, 80 (2016).
 4. C. Villavicencio, S. Schiaffino, J. Andres Diaz-Pace and A. Monteserin, Group recommender systems: A multi-agent solution, *Knowledge-Based Systems* **164**, 436 (2019).
 5. Y. Yang, D. Hooshyar and H. S. Lim, Gps: Factorized group preference-based similarity models for sparse sequential recommendation, *Information Sciences* **481**, 394 (2019).
 6. M. Mao, J. Lu, G. Zhang and J. Zhang, Multirelational social recommendations via multigraph ranking, *IEEE Transactions on Cybernetics* **47**, 4049 (2016).
 7. X. Wang, Y. Liu, J. Lu, F. Xiong and G. Zhang, Trugrc: Trust-aware group recommendation with virtual coordinators, *Future Generation Computer Systems* **94**, 224 (2019).
 8. T. Yu, J. Guo, W. Li, H. J. Wang and L. Fan, Recommendation with diversity: An adaptive trust-aware model, *Decision Support Systems* **123**, p. 113073 (2019).
 9. H. J. Jeong and M. H. Kim, Hggc: A hybrid group recommendation model considering group cohesion, *Expert Systems with Applications* (2019).
 10. P. Skowron, P. Faliszewski and J. Lang, Finding a collective set of items: From proportional multirepresentation to group recommendation, *Artificial Intelligence* **241**, 191 (2016).
 11. Y. Koren, Factorization meets the neighborhood: A multifaceted collaborative filtering model, in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2008.
 12. Q. Zhang, D. Wu, J. Lu, F. Liu and G. Zhang, A cross-domain recommender system with consistent information transfer, *Decision Support Systems* **104**, 49 (2017).
 13. T. A. Adomavicius, G., Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE Transaction on Knowledge and Data Engineering* **17**, 734 (2005).
 14. J. B. J. H. K. Fernando Ortega, Antonio Hernando, Recommending items to group of users using matrix factorization based collaborative filtering, *Information Sciences* **345**, p. 313–324 (2016).