Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence

Personalized Recommendation Based on Co-Ranking and Query-Based Collaborative Diffusion

Xiao Yang[1,2], Zhaoxin Zhang[1], Qiang Wang[2]
[1]School of Computer Science and Technology, Harbin Institute of Technology, China
[2]Microsoft Research Asia, Beijing, China
shawcsn@gmail.org

Abstract

In this paper, we present an adaptive graph-based personalized recommendation method based on co-ranking and query-based collaborative diffusion. By utilizing the unique network structure of n-partite heterogeneous graph, we attempt to address the problem of personalized recommendation in a two-layer ranking process with the help of reasonable measure of high and low order relationships and analyzing the representation of user's preference in the graph. The experiments show that this algorithm can outperform the traditional CF methods and achieve competitive performance compared with many model-based and graph-based recommendation methods, and have better scalability and flexibility.

Introduction

Recommendation is a particular form of information filtering which exploits user historical behaviors and multidimensional similarities to generate a list of information items that is personally tailored to an end-user's preferences. There are many different methods have been presented to solve this problem from diverse perspectives. Traditional content-based models need to rely on sufficient information of user and item profiles which are usually difficult to obtain. Memory-based CF such as item-CF and user-CF can't guarantee the performance due to the data sparsity and lack of explicit optimization objective. Although model-based CF methods such as matrix factorization are known to provide accurate recommendation performances, many models have difficulties in interpretability and scalability. To improve the general performance of recommender systems, several graph-based methods have been introduced into recommendation to model the high-order relationship and interaction between users and items on a graph, and compute node similarity from a global perspective. Unified bipartite graph model incorporated with many implicit feedbacks has been explored to support multidimensional recommendation (Lee et al. 2011). Random walk model combining the trust network and collaborative filtering for recommendation can provide confidence in its predictions (Jamali et al. 2009). Query-centered random walk executed on clustered k-partite graph has high computational efficiency depending on the constructed subgraph (Cheng et al. 2007).

There is a wide range of other information such as user's demographic data and item's profiles that can be incorporated into the recommendation effort for better recommendation quality. All the additional heterogeneous information from attributive data to temporal data can reflect the user's latent preference in different ways. We want to use different connected bipartite graph to represent the diversity of objects and relationships, and could generate recommendations from multidimensional query. This paper presents a two-layer graph-based ranking methods with different initialization and surfing strategies for personalized recommendation. Experiments on typical dataset are conducted to compare its recommendation effectiveness with other graph-based method and model-based CF method.

Graph-based Model

This recommendation method contains two steps of different random walk procedures based on connected multiple bipartite graphs. First step is a co-ranking model in user-item graph to give objects with different propagation capacity scores for subsequent query-based recommendation. The second collaborative diffusion step is a targeted fusion of different ranks got by collective propagation for initial query vectors in many graphs.

Co-ranking for users and items

Essentially, a recommender system consists of users and items and additional data assigned to them, each user has selected or rated some items. Most of researches on recommendation algorithms treat every user equally. But there exist reliable users or experts that have much deeper and broader knowledge of a specific domain. These users' preference information is more credible and convincing. Moreover, prior research shows that user ratings are imperfect and noisy which limits the measurable predictive power of many recommender systems. To solve this latent problem in original data, we design a co-ranking method in user-item weighted bipartite graph. We denote the set of user's rank as $U = \{u_1, u_2 ... u_m\}^T$, and the set of items' rank as $I = \{i_1, i_2 ... i_n\}^T$ .The I_0 can be got by using the weighted average to measure the rating aggregation of items. The graph's adjacent matrix is denoted as $M = \{r_{ij}\}_{m\times n}$, where r_{ij} is the rating score from user i to item j . Then different ranking process of user and item is executed simultaneously as follows,

$$u_x = (1-\alpha)u_x^0 + \alpha\sum_{m\in I} w_{mx}^{iu} i_m, i_y = (1-\alpha)i_y^0 + \alpha\sum_{n\in U} w_{ny}^{ui} u_n \quad (1)$$

Where $w_{mx}^{iu} = 1/\mathrm{d}_m$ and $w_{ny}^{ui} = \mathrm{r}_{ny}/(\mathrm{r}_{\max}\mathrm{d}_n)$, w^{ui} is the weight of directed edges from users to items and w^{iu} denotes the similar way. This iterative refinement process can converge and the ranking of users and items are mutually enhanced. By the iteration equation (1), we can compute the final rank values directly without iterative computing as follows,

$$u_x = (1-\alpha)u_x^0 + \alpha(1-\alpha)\sum_{m\in I} w_{mx}^{iu} i_m^0 + \alpha^2 \sum_{n\in U} w_{nx}^{uu} u_n \quad (2)$$

The i_y can be got in the similar form, and it's provable that this iterative framework has an equivalent optimization framework (Zhou et al. 2004) as min(O), and the object function can be written as,

$$O = \frac{1}{2}\sum_{i,j=1}^{m+n} w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^{m+n} \left\| f_i - f_i^0 \right\|^2 \quad (3)$$

After this co-ranking for users and items, we can get the user's rank as his expertise score and item's rank as its estimated quality score.

Query-based collaborative diffusion

Integrating and mediating heterogeneous sources of information can be the efficient way to improve the performance of recommender systems. Based on different types of explicit and implicit personal preferences and item profiles, such as ratings, tags, categories, locations, timestamps and so on, multidimensional relationships can be built in n-partite heterogeneous graph. We treat the recommendation for target user as an input query vector Q, and then implement the collaborative diffusion of initial query from different type of direct and indirect links in this multi-graph to get the ranking scores for target items. The collaborative iterative diffusion of arbitrary adjacent nodes or bipartite graph can be formulated as follows,

$$\vec{N} = \theta P^T \vec{N} + (1-\theta)\vec{Q} \quad (4)$$

Q is the query vector which can be any type of node in the diffusion process. N is the ranking vector of Q's neighbors, and we can compute the transition possibility matrix P by normalizing the adjacent matrix M. This iterative query process from nodes to nodes in multiple paths can get the final fused ranks for items which user hasn't explored as the recommendations.

Experimental Results and Conclusion

We verify our approach on the *Movielens-100k* which is a typical movie recommendation dataset, which contains personal ratings and movie profiles such as genres etc. We compare the top-k recommendation performance with several existing methods: item-based and user-based CF, Item-based random walk (RW) and matrix factorization (MF). The performance of these different recommendation algorithms is shown in Table 1.

Table 1: Performance comparison

	prec@5	prec@10	MAP	AUC
ItemKNN	0.691	0.677	0.503	0.919
UserKNN	0.759	0.714	0.525	0.921
RW	0.772	0.721	0.528	0.926
MF	0.814	0.752	0.541	0.947
Our Model	**0.836**	**0.781**	**0.564**	0.953

From the performance comparison results we can find that the two-layer collaborative ranking method can achieve better performance due to its graph-based iterative optimization and incorporating multidimensional data. As the initial query vector can be multiple and target ranking can be arbitrary nodes, this method is more scalable and flexible for various recommendation types in practice. As the future work we will explore integrating into temporal information and high-order complex relationships.

References

Lee, S., Song, S., Kahng, M., Lee, D. and Lee, S. 2011. Random walk based entity ranking on graph for multidimensional recommendation. Proceedings of the fifth ACM conference on Recommender systems , 93–100.

Jamali, M. and Ester, M. 2009. TrustWalker: a random walk model for combining trust-based and item-based recommendation. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 397–406.

Zhou, D., Bousquet, O., Lal, T.N., Weston, J. and Schölkopf, B. 2004. Learning with local and global consistency. Advances in neural information processing systems, NIPS 2004, 321–328.

Cheng, H., Tan, P.N., Sticklen, J. and Punch, W.F. 2007. Recommendation via query centered random walk on k-partite graph. Seventh IEEE International Conference on Data Mining, ICDM 2007, 457–462.