

Personalized Recommendation Systems: An Application in Case-based Reasoning

Jieli Sun, Qingwei Zhao

Hebei University of Business & Economics, China

Solomon Antony, Si Chen

Murray State University, USA

Abstract—This paper explores the application of case-based reasoning (CBR) to personal recommender system. The proposed approach incorporates users' characteristics as well as the object's attributes. Important aspects of the approach are similarity determination and case retrieval techniques are described. It is proposed that such a focus can lead to more effective personal recommendation system. This paper focused on the case similarity determining tactics and case retrieval tactics.

Keywords—case based reasoning; personalized recommendation system; case similarity; case retrieval.

I. INTRODUCTION

Recommender systems make recommendations to a collection of users for products that might interest the users. The recommendations are based on users' characteristics and on transactional data [15]. Two main approaches to Recommender system are (i) Collaborative filtering and (ii) Content based recommendations. While collaborative filtering uses historical ratings from all users, content based filtering is primarily based on product profile attributes. Collaborative filtering relies on collecting user ratings of products in a specific domain and determining similarities among user ratings to make the recommendation [15].

The CBR based personal recommendation approach involves (i) personalized recommendation expression (ii) construction of the related case library, (iii) determination of case similarity, and (iv) retrieving qualified cases. This study focuses on the Case expression and case library organization. Multiple case libraries form the data source for the recommendation reasoning system. Similarity determination and case retrieval are also important parts to effective recommender system [1][2][3].

II. CASE SIMILARITY DETERMINATION

A. Relevant research

According to the principles of Case-based reasoning (CBR), to solve new problems, we need to look up cases that are most similar to the current case on issues of properties and characteristics. Once we find the similar case, we can use its results to make the decision to the new problems. Until now, most of the CBR retrieval models have used the nearest neighbor method for determining similarity, emphasizing on the one-to-one attribute match of the case.[4][5][6].

For recommender systems, collaborative filtering based on neighbors is one of the most successful technologies and

one of the most widely used technologies. This recommendation method needs to find similar users (neighbors) and seek their recommendations. Clustering theory and algorithms play a role in collaborative filtering. User-based collaborative filtering is built on the assumption that if two or more users' ratings are quite similar on some projects, then their ratings on the other projects are also will quite similar. The basic idea of the algorithm is that target users can use the nearest neighbor to approach the project rating, to rate on the none-rated projects. Until now there are three methods to determinate similarities between personalized recommendation users. They are cosine similarity, the modified cosine similarity and correlation similarity [7][8].

All three methods base their computations on the similarity of the vector, and strictly match objects attributes. Cosine similarity measure method is that seeing user ratings as a vector use cosine angle between vectors to measure the similarity between users. However, no statistical characteristics of user ratings are included. The modified cosine similarity method is to lessen users' average rating, basing on the cosine similarity method. This method reflects more about the correlations between users, rather than the similarities. No matter which of the above methods is chosen, we need to calculate the similarity between users and among all the other users. When the number of users increases gradually, the required computation will increase sharply. While doing the case retrieval, the regular CBR systems need to determine the case similarity, and to look up the similar cases in the case library in the system. There also exists the case similarity calculation problems [9][10].

B. Similarity determining methods

Vectors of characteristic variables of the cases are stored in a searchable library. When determining case similarity, we compare characteristic vectors of the case to be judged to vectors of cases in the user case library and compute similarity measures. The formula for similarity computation can be chosen by the system administrator and it can be specified as configuration parameters of C-PRS recommendation system. This feature would make the recommendation system more flexible.

In the personal recommender system, the characteristics words of user's profile reflect his/her interests. Assigning each characteristic word to a category, we can determine the user's case category. We do this by adding the word class vectors of these characteristic words. We compute similarity measures for each element in the category vector. Next, we

will use the category with the highest similarity as the representative category of the user's case.

There may be a lot of registered users, and the needs of users' interests are in different categories, so there is no need to calculate the similarity between all the users' cases. The system only calculates the similarity between users cases representing the same categories. At first, according to the representative category of the user case, look up in the similar users' case library and calculate the similarity. Based on the similarity threshold setting in the recommender system, the cases that exceed the threshold are the similar ones, and they will be recorded to the similar user library. Since the similar user cases are the set of a group of user cases that meet the similarity requirements, the similar user cases in this paper can also be called similar user set.

III. PERSONALIZED RECOMMENDATION CASE RETRIEVAL

A. *Relevant research*

CBR uses similarity determination and case characteristics to identify one or more best cases from the case library that are associated with the current issue. Case retrieval is to find the similar case from the system case library. Case retrieval can be completed by three steps: characteristics identification, preliminary matches and the best selection. Case retrieval requires that similarity between the relevant cases be compared from all angles. In the view of the case retrieval technology, there are three methods of CBR case retrieval. They are (a) the nearest neighbor method, (b) inductive indexing and (c) knowledge guidance. The nearest neighbor method and the inductive indexing are more general case retrieval methods[6][9][11].

The nearest neighbor retrieval method aims to identify the cases from the library that are most similar to the current issue. If there is a good measure of similarity, this method will be a good retrieval method. But its retrieval time is directly proportional to the size of the case library. With the increase of case library, retrieval time will increase accordingly [6][9][11].

The inductive indexing retrieval is to continuously extract the components, which exhibit most difference between the case and the other cases, based on the case's components. The disadvantage of inductive indexing is that when there are new cases, it needs to go over the inductive indexing process again, requiring more time for indexing. This method requires a lot of cases to make induction. The accidental events can be inducted, and produce the wrong index [6][9][11].

If the case library is updated frequently, the inductive indexing retrieval method may not be most efficient method to use. Thus, based on the corresponding improvements of the nearest neighbor retrieval method, this paper use case retrieval method that are the integration of the similar case and the model case.

As a preparation of the multiple case library of the case-based reasoning personal recommender system, we need to employ appropriate case indexing mechanisms and create different case categories by classification. This is a general

process of creating the personal recommendation case library. As we need to select the most similar case with the users' case. The nearest neighbor retrieval method is among the most intuitive and most easy-to-use case retrieval method. The case-based reasoning personal recommendation case retrieval method in this paper, is based on the nearest neighbor method, making the corresponding improvements according to the characteristics of the personal recommender system [6][9][12].

The basic idea of the nearest neighbor retrieval system is to retrieve the case space in the recommender system, identifying the nearest neighbors to the users' requests. As to the case-based reasoning personal recommender system, the case space may gradually increase, subsequently affecting the system's performance. If the system's performance decline is noticeable, it may affect the satisfaction of users in using the system. When using the integration of the similar case and the model case method to do the case retrieval, the system can generate similar case and store into the similar case library offline. Users can find the similar neighbors in the similar case library when they do the retrieval. In this process, the model case helps the system to determine the retrieval category, and determines whether considering the recommended data by the model case according to users' recommendation choices [13][14].

B. *Integration of model and similar case*

In the case-based reasoning personal recommender system, case retrieval is divided into two stages namely user case library retrieval and similar user case library retrieval. The steps involved in this process are outlined below:

Step1: According to the mode case data and category information in the system knowledge library, determine the type of users' requirements, i.e. forming category index of the retrieval.

Step2: According to the category index of the retrieval and user's case data, do the case retrieval in the similar user case library, to find the similar user cases.

Step3: Judge whether the similar user case library retrieval is completed. If so, it will return to the most similar case, and turn to Step 4. Otherwise, go back to step 2.

Step4: According to the returned similar case, get the most similar case to meet users' needs. It will be regarded as the solution of the personal recommender, providing basis for the personal recommendation reasoning.

Step5: Algorithm ends.

IV. CONCLUSIONS

The key issues of case-based reasoning personal recommender system is the quality of personal recommendation results. It depends on the case similarity determining strategies, case retrieval strategies and case-based personal recommendation strategy. Case similarity determining strategy and case retrieval strategy are important parts of the design process of CBR system. Their quality can greatly affect the efficiency of the CBR system. Case similarity determining strategy and case retrieval strategy ensures the successful application of personal

recommender system. They are the guarantee of the success of the case-based reasoning personal recommendation.

ACKNOWLEDGEMENTS

This research work was supported by Social Science Foundation of HeBei province under Grant No. HB2011QR47.

REFERENCES

- [1] Kolodner, Janet L. "An introduction to case-based reasoning." *Artificial Intelligence Review* 6.1 (1992): 3-34.
- [2] Aamodt, Agnar, and Enric Plaza. "Case-based reasoning: Foundational issues, methodological variations, and system approaches." *AI communications* 7.1 (1994): 39-59.
- [3] Gilboa, Itzhak, and David Schmeidler. "Case-based decision theory." *The Quarterly Journal of Economics* 110.3 (1995): 605-639.
- [4] Juell, Paul, and Patrick Paulson. "Using reinforcement learning for similarity assessment in case-based systems." *IEEE Intelligent Systems* 18.4 (2003): 60-67.
- [5] Pearce, Michael, et al. "Case-based design support: A case study in architectural design." *IEEE Expert* 7.5 (1992): 14-20.
- [6] Zhang, Rongmei . A Study of the smart decision support system based on CBR and MAS and its applications [D]. Univeristy of Science and Technology Beijing, 2001.
- [7] Chuangfeng Wei, Huo Pang, Yunze Li. The application of an improved nearest neighbor case-based reasoning. *Journal of System and Simulation*, 2005,17(5):1045-1047.
- [8] Schafer, J. Ben, Joseph Konstan, and John Riedl. "Recommender systems in e-commerce." *Proceedings of the 1st ACM conference on Electronic commerce*. ACM, 1999.
- [9] Ting Liu, Bing Qing, Yu Zhang, Wanxiang Che. *Introduction to Information Search System* [M]. Mechanical Engineering Press, 2008,12:150.
- [10] Zhirui Xu, Lanrong Yang. A discussion of similarity measures in case-based recommendation system. *Journal of Xiamen University (Natural Science)* , 2000,39(4):441-445.
- [11] Liao, T. Warren, Zhiming Zhang, and Claude R. Mount. "Similarity measures for retrieval in case-based reasoning systems." *Applied Artificial Intelligence* 12.4 (1998): 267-288.
- [12] Shuguang, Zhang, et al. "A MODEL OF CASE MATCHING METHOD IN CASE-BASED REASONING [J]." *Pattern Recognition and Artificial Intelligence* 3 (2002).
- [13] Jian Ying, Shixian Li. A case-based search model [J]. *Acta Scientiarum Naturalium Universitatis Sunyatseni*. summary, 1999, 38(2):1-5.
- [14] Shishen Zhong, Xingui He. A case-based hybrid search model. *Journal of Software*, 1999, 10(5):521-526.
- [15] Melville, Prem, and Vikas Sindhwani. "Recommender systems." *Encyclopedia of machine learning*. Springer US, 2010. 829-838.