A Collaborative Recommender System: Lexicographic Consensus and Web Usage Mining Approach

Seok Kee Lee[1] , Young-Gab Kim[2], and Do-Gil Lee[3,1]

[1] Department of Industrial & Management Engineering, Hansung University, Seoul, Korea
seelee@hansung.ac.kr
[2] College of Information & Communication, Korea University, Seoul, 136-701, Korea
always@korea.ac.kr
[3] Research Institute of Korean Studies, Korea University, Seoul, 136-701, Korea
motdg@korea.ac.kr

Abstract. Collaborative filtering (CF) is one of the most successful and widely used methods of automated product recommendation domain [1, 2, 3]. However, cardinal scale generally used for representing the preference intensity is also ineffective owing to its increasing estimation errors. In this paper, we propose a new CF-based recommendation methodology that constructs an ordinal scale-based customer profile under the implicit ratings condition. An experiment with the Web transaction data from a real online shopping mall shows that the proposed method performs better than existing CF methodologies.

Keywords: collaborative filtering, recommender system, lexicographic consensus, web usage mining

1 Introduction

In this paper, we propose a CF recommendation methodology, called Lex-CF, which implicitly constructs an ordinal scale-based customer profile by using Web usage mining (WUM) and lexicographic consensus. In order to capture implicit preference information, we employ a WUM technique [4] and a specific consensus method, called lexicographic consensus [5]. It generates a compromised preference from various partial pieces of ordinal preference information and then generates an ordinal scale-based customer profile.

[1] Do-Gil Lee is a corresponding author.

2 Proposed Methodology

2.1 Phase1: Web Usage Mining (WUM)

In this phase, we apply the WUM technique to extract customer's preference from information on his/her previous behaviors shown on the Web sites. The raw web log files are pre-processed and customer's transaction data including the information such as <time, customer-id, product-id, shopping behavior> is abstracted [6]. The transaction data collected are converted into some specific matrix, called the customer behavior set (CBS). CBS $V^* = (v_{sj})$ is a matrix containing the target customer's shopping behaviors across products.

2.2 Phase 2: Ordinal Scale-based Customer Profile Creation

Step 1. Preference sequence elicitation

The preference sequence, $E_b = \{r_{bj} | m_j \in M_b\}$, is a sequence of rankings of a customer over all products he/she experienced during the navigation. The process is repeated until all the criteria have been used.

Step 2. Ordinal scale-based customer profile creation

A complete customer profile, $U = \{E_1,...,E_Y\}$ is created by iterating step 1 repeatedly for all Y customers (Y is the total number of customers). Given U, the remaining Phase of Lex-CF is performed in the next phase 3.

2.3 Phase 3: Neighborhood Formation / Recommendation Generation

Step 1. Neighborhood formation

The similarity between customers and forms a neighborhood between a target customer and a group of like-minded customers is calculated [7]. Using this similarity measure, this step determines which previous customers will be used in the recommendation for the target customer. In our methodology, the best-*n*-neighbor technique is adopted because of its superior performance.

Step 2. Recommendation generation

This step derives the top-*N* recommendation from the neighborhood of customers. For each customer, we produce a recommendation list of *N* products that the target customer is most likely to purchase.

3 Experimental Evaluation

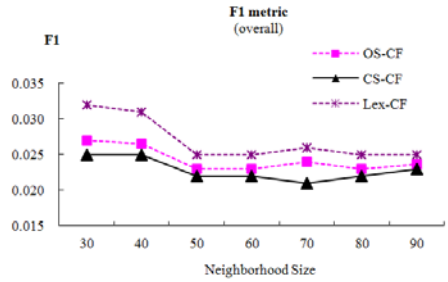

Fig. 1. Overall F1 metrics for the three CF methodologies

4 Conclusion

This paper has presented a new CF-based recommendation methodology called Lex-CF with the aim of addressing the product overload problem in large e-commerce sites. The experiments to verify the performance of Lex-CF by using real Web transaction data confirmed the superiority of Lex-CF over existing CF-based recommender systems. It proved that Lex-CF may be a viable CF solution for the large e-commerce sites.

Acknowledgements. This research was financially supported by Hansung University.

References

1. Martin, L-N., Yolanda, B-F., Jose, J.P., Aberto, G-S.: Property-based Collaborative Filtering for Health-Aware Recommender System. Expert Systems with Applications. 39, 7451--7457 (2012)
2. Tsai, C-F., Hung, C.: Cluster Ensembles in Collaborative Filtering Recommendation. Applied Soft Computing. 12, 1417--1425 (2012)
3. Lee, S. K., Cho, Y. H., Kim, S.H.: Collaborative Filtering with Ordinal Scale-based Implicit Ratings for Mobile Music Recommendations. Information Sciences. 180, 2142--2155 (2010)
4. Cook, W.D.: Optimal Allocation of Proposals to Reviewers to Facilitate Effective Ranking. Management Science. 51, 655--61 (2005)
5. Dombi, J., Imreh, C., Vincze, N.: Learning Lexicographic Orders. European Journal of Operational Research. 183, 748--756 (2007)
6. Cho, Y.H., Kim, J.K.: Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendations in e-Commerce. Expert Systems with Applications. 26, 236--246 (2004)
7. Kamishima, T.: Nantonac Collaborative Filtering: Recommendation Based on Order Response. In: 9th International Conference on Knowledge Discovery and Data Mining, pp. 583--588 (2003)