Improved Post-Filtering Method Using Context Compensation

Be-Deu-Ro Kim and Jee-Hyong Lee

Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Korea

ljfis

Abstract

According to the expansion of smartphone penetration and development of wearable device, personal context information can be easily collected. To use this information, the context aware recommender system has been actively studied. The key issue in this field is how to deal with the context information, as users are influenced by different contexts while rating items. But measuring the similarity among contexts is not a trivial task. To solve this problem, we propose context aware post-filtering to apply the context compensation. To be specific, we calculate the compensation for different context information by measuring their average. After reflecting the compensation of the rating data, the mechanism recommends the items to the user. Based on the item recommendation list, we recover the rating score considering the context information. To verify the effectiveness of the proposed method, we use the real movie rating dataset. Experimental evaluation shows that our proposed method outperforms several state-of-the-art approaches.

Keywords: Context aware, Context compensation, Post-filtering, Relative average distance

1.   Introduction

Owing to the rapid pace of IT technological development recently, the number of smartphone users and wearable device users are sharply increased. As a result, context information like location of the user, time and weather can be easily collected through sensors which are attached to the devices [1]. Context aware recommender system which uses context information is under active study. But there is a problem, how this system deals with contextual attributes. Previous context approach studies are divided into three types [2]:

1) Contextual pre-filtering: It makes the matrix to unify the data, which consists of the user, item and context information, to a certain context before recommendation.

2) Contextual post-filtering: It makes the user and item matrix to recommend significant item. After this process, context information is additionally considered to find the recommendation item which applies to the context.

3) Contextual modeling: It makes the matrix to consider the user, item and context information simultaneously, while making recommendations.

In case of contextual pre-filtering, the recommender system performance gets lower when the quantity of information increases. This is due to the increase in sparsity problem. contextual post-filtering and contextual modeling have limitations to reflect all context information to the

increasing number of information, when high dimension matrix is generated [3].

In this paper, we propose context aware post-filtering approach to consider the context compensation. This approach maintains the high dimensional context information and reduces the data sparsity problem. Our method utilizes the relative average difference among the context and collaborative filtering algorithm. The main idea of this approach is to apply the context compensation for the movie rating while generating the recommendation list and adjusting the ranking list. To verify the effectiveness of the proposed method, the rating accuracy is measured by using the real movie rating dataset.

The rest of the paper is organized as follows. Section 2 provides the description of the previous research to relate the context compensation. In Section 3, the background methods of our study are introduced. Section 4 gives an overview of our recommender system and describes how to calculate the context compensation for the context attributes. Section 5 reports the experimental results and offers some discussions. In Section 6, conclusions are drawn and future work is presented.

2. Related Work

It is important to deal with the contextual information in context aware recommender system. The existing methods are classified three types: pre-filtering, post-filtering, contextual modeling.

Pre-filtering approach makes the 2-D matrix by using the most relevant contextual information when generate the recommendations. In general, previous researches apply the item splitting method. It reduces data dimension that items are splitting into two or more subsets following the value of a contextual factor [4]. But if the more context subsets are considered, the data sparsity problem also increased. To solve this problem, it uses machine learning techniques to generalize the contextual information such as principal component analysis [5], clustering [6]. However, these methods may lose the original data information and it does not consider all contextual information.

In case of the post-filtering approach, it is similar to pre-filtering approach. It also uses the user-item (2-D) matrix for generating recommendation lists. After generating the list, the contextual information is reflected to adjust the ranking of recommendations on the list [7]. Likewise, the contextual information generalization research is studied. Zhu et al. [8] applied the Latent Dirichlet Allocation (LDA) to convert the contextual information into vector type. It has advantage of the recommendation lists reordering according to calculated similar

contextual information. Ramirez-Garcia and Garcia-Valdez [9] used the post-filtering method in combination with the collaborative filtering method. They extract Top-N recommendation lists by using the collaborative filtering algorithm and adjust the lists to consider the user context. However, both methods focused on re-ranking the recommendation lists. In other words, these ignore the contextual information when they make the recommendation lists initially.

Contextual modeling approach [10] is different previous from two methods. It considers the contextual information while generating recommendation lists. Contextual pre-filtering and post-filtering method use user-item recommendation function, while contextual modeling creates the preference predictive model based on user-item-contextual matrix. Typically, Tensor Factorization, which is matrix decomposition method, is used to deal with multidimensional data. Karatzoglou et al. [11] proposed recommendation method to utilize the context such as season, movie release date by modeling the n-dimensional tensor. Shi et al. [12] focused on the mood when users watch the movie. They combine the user-movie matrix and mood similar matrix between the two movies to recommend the movie for users. These kinds of methods have superior performance than other methods but it also has computational complexity.

As described above, previous study on contextual information did not consider that rating consists of item rating and contextual rating. To handle these problems, we propose the context compensation method when the recommendation lists are generated and adjust the item recommendation lists. Through the context compensation, we can improve the recommendation performance and prevent the contextual information loss. In addition, this proposed method increases accuracy of the recommendation lists by applying the improved post-filtering approach method.

3. Background

3.1 Context

Context aware recommender system is a kind of recommender system which recommends items to the user by considering various contextual information that can influence the user consumption or preference. It is important to define the context in that respect. Schilit and Theimer [13] introduced the term 'context aware' which means user's location, identities of nearby people, objects and change to those objects. Abowd et al. [14] defined that any information can be used which characterize the situation of entities such as user's emotional state, date and time,

location and people in the user's environment. Above this, some previous research has been done for the context concept. But it is not clear whether a type of contextual information affects the user's selection. And it still has the technical limitation to collect the varied contextual information. For such a reason, context recommender system has been studied in specific domain such as travel, movie, and music that collect the contextual information and understand relation between the item selection and contextual information relatively easy. In this paper, we assume that context is to influence the user's evaluation of the item and characterize the situation and users mutual relation.

3.2 Contextual Post-filtering

Adomavicius and Tuzhilin [15] introduced paradigms for context in recommender system. Contextual post-filtering is the one of them which ignores the contextual information initially. The rating for the items is predicted by using the $U \times I \times R$ ($user \times item \times rating$) recommender system on the data. After generating the ranked list of the recommendation items, Contextual information use to adjust the recommendation list. There is two type of methods exist for list adjustment. First thing is to filter out recommendation list that are irrelevant context. For example, if the item on recommendation lists has not a number of context. Another type is adjusting the ranking of recommendation based on the contextual information. As it mentioned in the related work, this method just considers the contextual information when adjust the recommendation lists. Figure 1 illustrated the contextual post-filtering approach.
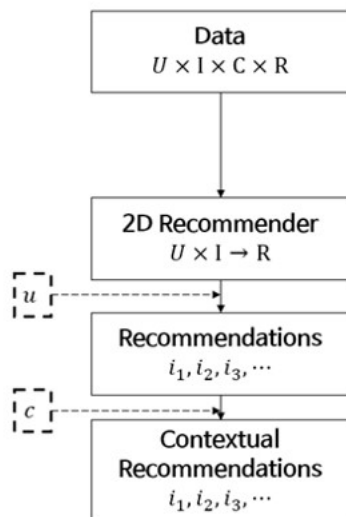

Figure 1. Contextual post-filtering.

3.3 Collaborative Filtering

Collaborative filtering is a technique that suggests new items to the user or predicts the users' preference for a certain item based on the users' purchase history record [16]. Figure 2 shows the schematic diagram of collaborative filtering process.
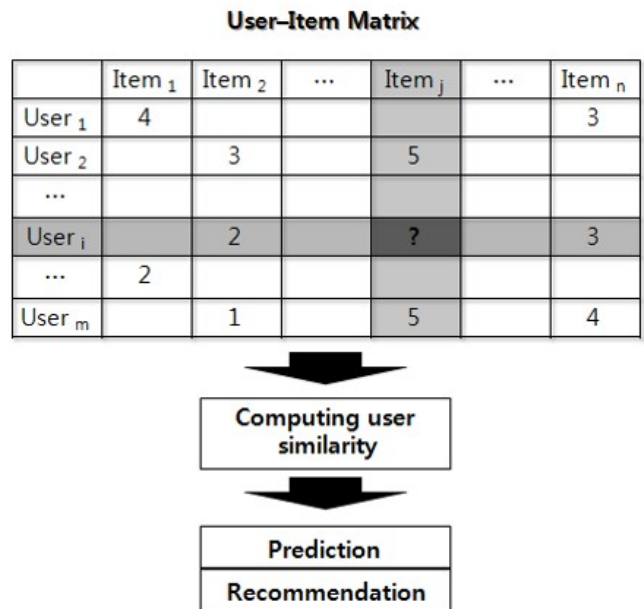

Figure 2. The collaborative filtering process.

This algorithm uses the user-item data as a ratings matrix. Each number in the square represents the user preference score of the ith user on the jth item. When the target user is selected, the process measures the similarity between the target user and other users. Then, it calculates the weighted average of the ratings by the neighbors who is similar to target user and recommend the items with the rating.

4. Proposed Method

In this section, we introduce our proposed method to apply context compensation that use relative distance among the rating average for each contextual information. This approach assumes that user's preference score (rating) not only consist of item rating but also context rating. Item rating means user's evaluation value of the content consumption. At this point of time, they are affected by the surrounding circumstances which are indicated context rating. In order to reflect this aspect, we conduct context compensation as two steps: pre-context compensation and post-context compensation.

Given the input data which has user-item-context matrix,

previous post-filtering approach make the recommendation lists by using only user-item matrix. But our method considers the contextual information in this stage which we call pre-context compensation. Figure 3 shows a simple structure of this method.
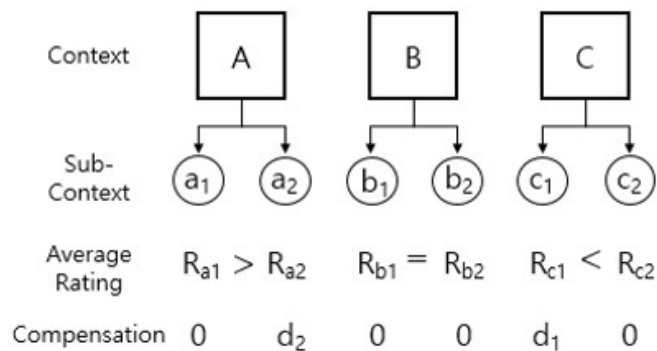

Figure 3. Context compensation according to average rating.

A, B, C mean context such as time, location, relation and each of them has context values (e.g., time is divided weekday or weekend and location has home or office). For input dataset, we calculate the average ratings of each context value and then compare the difference between the average ratings. If the average rating of a context value is bigger than the others, it has 0 compensation. On the other hand, the other context values have compensation values which are the difference from the maximum average rating. If the two average rating is the same, each of them has 0 compensation for the contextual information.

Table 1. An example of movie domain data

	Item	Time	Location	Relation	Rating
User1	T1	Weekday	Home	Family	3
User2	T2	Weekend	Home	Family	3

Table 1 shows the movie domain data example. Let us assume that average rating on the weekday is bigger than weekend and different between them is 0.5. Location and relation have same compensation. As you can see in the Table 1, the item T1 rating is 3. If we apply our proposed contextual compensation method to the rating, item T2 rating is converted to 3.5. T2 has already included content rating for the item and we summate the context compensation for each contextual information additionally. After input data transformed to reflect the context compensation into the rating, we generate the recommendation lists for user-item matrix by using the collaborative filtering algorithm. Then, we adjust the recommendation item rating to apply the context

compensation and we call post-context compensation. This process is similar to the pre-context compensation. On the contrary, we subtract the context compensation for the recommendation items rating in this processing.

5. Experiments

5.1 Dataset Description

In order to identify the performance of our proposed method, we use DePaul Movie Dataset [17]. DePaul Movie Dataset is collected from survey asking the students to rate the movies in different context situations. There are 5,029 ratings (scale 1-5) by 97 users on 79 movies within contexts "time, location, companion." Context "time, location" has two sub-contexts and "companion" has three sub-contexts.

5.2 Experimental Design

We compare performances of three baseline methods with our method to verify the validity. Baseline methods used are: traditional collaborative filtering (CF) method, and Slope One [18] and Item-Splitting [19]. Slope One method compare to the user preference different between new item and other item. It dramatically reduces the overfitting. Item-splitting is the novel pre-filtering technique. According to the value of an item-dependent contextual condition, the rating of certain items are split. In the experiment, we apply the 10-fold cross-validation, 90% of ratings are used as training set and remaining 10% are used as test set. In order to handle the data sparsity, we fill the prediction rating value into the unvalued rating by using the user similarity and other users rating. Mean absolute error (MAE) and root mean squared error (RMSE) are used as evaluation criteria [20]. MAE predicts the rating between the target user and other users and measures the degree of accuracy of the recommendations. MAE is defined as follows:

$$MAE = \frac{\sum_{i=1}^{n} |P_i - P_j|}{n}. \tag{1}$$

In Eq. (1), the symbol P_i is the corresponding real ratings, and P_j is the prediction of user's ratings. n is the total number of rating in the data. More the evaluation value is close to zero; lower will be the predicted score error. And RMSE is shown in Eq. (2).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (p_i - p_j)^2}. \tag{2}$$

P_i and P_j have the same meaning as MAE. But It is the RMSE that mean of the square all of the error and has large numerical errors as compared with MAE.

5.3 Experiment Results

This section gives an evaluation of the proposed method for context compensation in recommender system. We compare our proposed method with three baselines which include CF, Slope One and item-splitting by using MAE and RMSE. As shown Figure 4, our proposed method outperformance than other baseline recommender system. Especially, our method is better performance than existing CF approach about 0.08. The result is presented in Figure 4 in this paper.
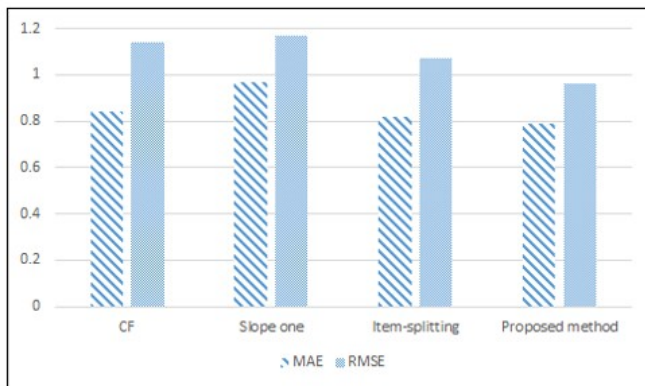

Figure 4. Performance of context compensation method. MAE, mean absolute error; RMSE, root mean square error.

6. Conclusion

This study is focused on reflecting the contextual information in the context aware recommender system. The main idea of this approach is to calculate the relative distance of each context value by using the contextual average rating and compensate the rating to add the context evaluation value. Our experiment result shows that proposed method is outperformance than traditional CF method and contextual filtering method. Through this result, we can prove that users rating is not only consist the content rating but also context rating. As a future work, We will consider more contextual information and structure for the context aware recommender system.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

[1] D. H. Kim, J. M. Shin, and S. W. Park, "Decision tree based application recommendation system," *Proceedings of Korea Computer Congress*, vol. 39, no. 1, pp. 140-142, 2012.

[2] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin, "Incorporating contextual information in recommender systems using a multidimensional approach," *ACM Transactions on Information Systems*, vol. 23, no. 1, pp. 103-145, 2005. http://dx.doi.org/10.1145/1055709.1055714

[3] K. C. Choi, J. H. Yang, and H. J. Lee, "Context-aware recommender system using purchase history data," *Proceedings of Korean Institute of Information Scientists and Engineers*, Pyeongchang, Korea, 2014, pp. 595-597.

[4] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734-749, 2005. http://dx.doi.org/10.1109/TKDE.2005.99

[5] M. Unger, A. Bar, B. Shapira, and L. Rokach, "Towards latent context-aware recommendation systems," *Knowledge-Based Systems*, vol. 104, pp. 165-178, 2016. http://dx.doi.org/10.1016/j.knosys.2016.04.020

[6] V. Codina, F. Ricci, and L. Ceccaroni, "Distributional semantic pre-filtering in context-aware recommender systems," *User Modeling and User-Adapted Interaction*, vol. 26, no. 1, pp. 1-32, 2016. http://dx.doi.org/10.1007/s11257-015-9158-2

[7] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic, and E. Duval, "Context-aware recommender systems for learning: a survey and future challenges," *IEEE Transactions on Learning Technologies*, vol. 5, no. 4, pp. 318-335, 2012. http://dx.doi.org/10.1109/TLT.2012.11

[8] H. Zhu, E. Chen, H. Xiong, K. Yu, H. Cao, and J. Tian, "Mining mobile user preferences for personalized context-aware recommendation," *ACM Transactions on Intelligent Systems and Technology*, vol. 5, no. 4, article no. 58, 2015. http://dx.doi.org/10.1145/2532515

[9] X. Ramirez-Garcia and M. Garcia-Valdez, "Post-filtering for a restaurant context-aware recommender system," in *Recent Advances on Hybrid Approaches for Designing Intelligent Systems*, O. Castillo, P. Melin, W. Pedrycz, and J. Kacprzyk, Eds. Cham: Springer International Publishing, 2014, pp. 695-707. http://dx.doi.org/10.1007/978-3-319-05170-3_49

[10] L. Baltrunas, "Context-aware collaborative filtering recommender systems," Ph.D. dissertation, Free University of Bozen-Bolzano, Bolzano, Italy, 2011.

[11] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, "Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering," in *Proceedings of the 4th ACM Conference on Recommender Systems*, Barcelona, Spain, 2010, pp. 79-86. http://dx.doi.org/10.1145/1864708.1864727

[12] Y. Shi, M. Larson, and A. Hanjalic, "Mining mood-specific movie similarity with matrix factorization for context-aware recommendation," in *Proceedings of the Workshop on Context-Aware Movie Recommendation*, Barcelona, Spain, 2010, pp. 34-40. http://dx.doi.org/10.1145/1869652.1869658

[13] B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *IEEE Network*, vol. 8, no. 5, pp. 22-32, 1994. http://dx.doi.org/10.1109/65.313011

[14] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Proceedings of 1st International Symposium on Handheld and Ubiquitous Computing*, Karlsruhe, Germany, 1999, pp. 304-307. http://dx.doi.org/10.1007/3-540-48157-5_29

[15] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Proceedings of the 2008 ACM Conference on Recommender Systems*, Lousanne, Switzerland, 2008, pp. 335-336. http://dx.doi.org/10.1145/1454008.1454068

[16] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th International Conference on World Wide Web*, Hong Kong, 2001, pp. 285-295. http://dx.doi.org/10.1145/371920.372071

[17] Y. Zheng, B. Mobasher, and R. Burke, "CARSKit: a Java-based context-aware recommendation engine," in *Proceedings of 2015 IEEE International Conference on Data Mining Workshop*, Atlantic City, NJ, 2015, pp. 1668-1671. http://dx.doi.org/10.1109/ICDMW.2015.222

[18] D. Lemire and A. Maclachlan, "Slope one predictors for online rating-based collaborative filtering," in *Proceedings of the 2005 SIAM International Conference on Data Mining*, Newport Beach, CA, 2005, pp. 471-475. http://dx.doi.org/10.1137/1.9781611972757.43

[19] L. Baltrunas and F. Ricci, "Context-based splitting of item ratings in collaborative filtering," in *Proceedings of the 3rd ACM Conference on Recommender Systems*, New York, NY, 2009, pp. 245-248. http://dx.doi.org/10.1145/1639714.1639759

[20] G. Shani and A. Gunawardana, "Evaluating recommendation systems," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. New York, NY: Springer, 2011, pp. 257-297. http://dx.doi.org/10.1007/978-0-387-85820-3_8

Be-Deu-Ro Kim received his B.S in mechatronic engineering from Korean Polytechnic University, Korea in 2015. He is currently pursuing his M.S in computer engineering at Sungkyunkwan University. His research interests recommender system and machine learning.
E-mail: kimbdr@skku.edu

Jee-Hyong Lee received his B.S., M.S., and Ph.D. in computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1993, 1995 and 1999, respectively. From 2000 to 2002, he was an international fellow at SRI International, USA. He joined Sungkyunkwan University, Suwon, Korea, as a faculty member in 2002. His research interests include fuzzy theory and application, intelligent system and machine learning.
E-mail: john@skku.edu