

A Study on Improvement of Serendipity in Item-based Collaborative Filtering Using Association Rule

Hiroaki Ito, Tomohiro Yoshikawa, Takeshi Furuhashi

Abstract—The number of available items in online shops are increasing by the spread of the Internet recently. Though users have a wide range of choices, they need to find their favorite items from a huge amount of information. Thus, a variety of recommendation systems are currently in use. “Accuracy” is the most important index in these recommendation systems. However, not only “Accuracy” but also “Serendipity” is said to be needed in terms of user satisfaction recent years. In this paper, we introduce a recommendation method of collaborative filtering based on association analysis which is one of the data mining techniques. We aim to improve Serendipity keeping Accuracy high by using the evaluation information that are rated differently from a target user. In addition, we show that Accuracy and Serendipity can be adaptable by a parameter in the proposed method. This paper compares the proposed method with a conventional method in terms of the performance of Accuracy and Serendipity.

I. INTRODUCTION

Recently, the number of e-commerce is increasing by the spread of the Internet, and the number of available items in online shops is also increasing. At the same time, it becomes difficult for users to find their favorite items from a huge amount of information. Thus, recommendation system is needed[1], and a variety of recommendation systems are currently in use. Amazon.com[4] is one of the most popular recommendation systems in the practical use. In this web cite, recommendation list of items which are related to each target user is shown.

Association analysis[5] is one of the data mining techniques which aim to extract valuable information from mass data. Some recommendation methods adapting the association analysis to user’s rating histories have been reported[6][7].

“Accuracy” is an important index in recommendation systems, which is the ratio of the number of user’s favorite items over that of recommended items. However, in terms of user satisfaction, it is said that to evaluate and improve “Serendipity” is also needed in addition to “Accuracy”[1][2][3][7][8][9].

Recommendation system is categorized into 2 types. One is based on collaborative filtering, which is a method to find users/items having similar tastes to a target user’s taste. In this method, uses’ ratings history is used as the

recommendation information. The other is based on content-based filtering, which is a method to find items having similar features of items which a target user likes. In this method, items’ features are used as the information. The collaborative filtering has the advantage of “Serendipity” because the items recommended by the content-based filtering become similar and it does not have to use much information but only use a rating history[2][9]. Thus we focus on the collaborative filtering.

In the conventional collaborative filtering based on the association analysis, the score for the recommendation items is calculated by the evaluation information that are rated same with a target user. In this paper, we aim to improve Serendipity keeping Accuracy high by using the evaluation information that are rated differently from a target user. In addition, we show that Accuracy and Serendipity can be adaptable by a parameter in the proposed method. This paper compares the proposed method with a conventional method in terms of the performance of Accuracy and Serendipity.

II. RECOMMENDATION SYSTEM

A. Association analysis

Association analysis is a method to find valuable combinations (association rules) from mass data. For example, when we find combinations of goods which are purchased continually, this knowledge could be useful for the display of them.

Association rules are expressed in the form of “ $A \Rightarrow B$,” in which A is the condition and B is the conclusion. This rule means that B will happen when A does. *confidence* is the typical evaluation index of association rules and calculated by the following equation (1).

$$confidence_{(A \Rightarrow B)} = \frac{N(A \wedge B)}{N(A)} \quad (1)$$

$N(A)$ represents the number of data which meet the condition A , and $N(A \wedge B)$ represents the number of data which meet the condition A and the conclusion B at the same time.

B. Item-based collaborative filtering

Memory based method in the collaborative filtering (CF) is categorized into item-base and user-base. In the item-based CF, a target user’s rating history serves as the association rules’ conditions, and “a candidate item = Like” is the conclusion. For example, when the target user evaluated an item A as “Like” and did not evaluate an item B , the value of $confidence_{(A=Like \Rightarrow B=Like)}$ is added to the score of a candidate item B . When the target

Hiroaki Ito is with the Department of Computer Science and Engineering, Nagoya University, Japan(email: itou@cmlpx.cse.nagoya-u.ac.jp).

Tomohiro Yoshikawa is with the Department of Computer Science and Engineering, Nagoya University, Japan(email: yoshikawa@cse.nagoya-u.ac.jp).

Takeshi Furuhashi is with the Department of Computer Science and Engineering, Nagoya University, Japan(email: furuhashi@cse.nagoya-u.ac.jp).

user evaluated the item A as “Don’t Like,” the value of $confidence_{(A=Don'tLike \Rightarrow B=Like)}$ is added. After all rating histories are used and the scores for all non-evaluated items are calculated, an item with the highest score is recommended to the target user.

C. User-based collaborative filtering

A target user’s rating (Like/Don’t Like) serves as the association rules’ conditions, and “the target user = Like” is the conclusion. For example, when user 1 evaluated an item A as “Like” and the target user did not evaluate the item A , the value of $confidence_{(user1=Like \Rightarrow target\ user=Like)}$ is calculated and added to the score of item A . If the value of $confidence_{(user1=Like \Rightarrow target\ user=Like)}$ is high, the target user is likely to evaluate the item A as “Like.” After calculating the scores for all non-evaluated items of the target user based on all users’ rating histories, an item with the highest score is recommended to the target user.

III. RELATED WORK

One of the item-based CF is Weighted Sum[10]. This method calculates $P(u)_i$ which is the prediction of user u on item i .

$$P(u)_i = \frac{\sum_{j \in I} (s_{i,j} * R_{u,j})}{\sum_{j \in I} |s_{i,j}|} \quad (2)$$

Where, $R_{u,j}$ is the rating value of user u on item j , $s_{i,j}$ is the adjusted cosine similarity between item i and j , and I is the set of similar items to item i .

Adjusted cosine similarity is calculated by the following equation (3). U is the set of users who rated both item i and item j . \bar{R}_u is the average of the u -th user’s rating.

$$s_{i,j} = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}} \quad (3)$$

IV. PROPOSED METHOD

In the subsection II-B, a target user’s rating history only serves as the association rules’ conditions. In the proposed method, regarding to the item-base CF, not only a target user’s rating history but also the contrary rating history serves as the association rules’ conditions, and “a candidate item = Like” is the conclusion for both conditions.

We define R_t as a target user’s rating and $\neg R_t$ as the contrary rating. R_t and $\neg R_t$ are “Like” or “Don’t Like.” If the value of $confidence_{(A=R_t \Rightarrow B=Like)}$ is high, the target user who evaluated the item A as R_t is likely to evaluate the item B of the association rule’s conclusion as “Like.” So high “Accuracy” is expected when the item with high value of $confidence_{(A=R_t \Rightarrow B=Like)}$ is recommended.

Next, we defined d in eq. (4) which is the subtraction between $confidence_{(A=R_t \Rightarrow B=Like)}$ and $confidence_{(A=\neg R_t \Rightarrow B=Like)}$.

$$d = confidence_{(A=R_t \Rightarrow B=Like)} - confidence_{(A=\neg R_t \Rightarrow B=Like)} \quad (4)$$

When $|d|$ is high, the users who evaluated the item A as R_t have a different taste from those as $\neg R_t$ in terms of “item B = Like.” It means that the target user’s rating history “ $A = R_t$ ” has more information than the information of the items with low $|d|$.

For example, suppose that item B_1 and item B_2 are candidate items, and both the value of $confidence_{(A=R_t \Rightarrow B_1=Like)}$ and $confidence_{(A=R_t \Rightarrow B_2=Like)}$ are 0.8. The item-based CF in the subsection II-B gives the same score to the item B_1 and the item B_2 . However, when the value of $confidence_{(A=\neg R_t \Rightarrow B_1=Like)}$ is 0.2 and that of $confidence_{(A=\neg R_t \Rightarrow B_2=Like)}$ is 0.6, the meaning of “item $A = R_t$ ” between the item B_1 and the item B_2 is much different. It shows that the target user is likely to evaluate the item B_1 as “Like” with high probability different from those who evaluated the item A as $\neg R_t$, and the target user is also likely to evaluate the item B_2 as “Like” but less surprise because it is not specified to the target user. So $|d|$ intends the difference of the taste between a target user and others.

As described above, the value of $confidence_{(A=R_t \Rightarrow B=Like)}$ is expected to lead high “Accuracy,” and the value of $|d|$ is expected to lead high “Serendipity” because the recommendation of the item with high $confidence$ but low $|d|$ means non-personalized recommendation[15][16]. Thus, the score for recommendation in the proposed method is calculated by the following equation (5).

$$s_B = \begin{cases} confidence_{(A=R_t \Rightarrow B=Like)}^\alpha * d & \text{if } d \geq 0 \\ confidence_{(A=R_t \Rightarrow B=Don'tLike)}^\alpha * d & \text{otherwise} \end{cases} \quad (5)$$

If d is only used for the recommendation score, the value of $confidence_{(A=R_t \Rightarrow B=Like)}$ is not considered. So d multiplied by $confidence_{(A=R_t \Rightarrow B=Like)}$ is applied to s_B . When $d \geq 0$, the higher both d and $confidence_{(A=R_t \Rightarrow B=Like)}$ are, the greater chance to be recommended the item B has. When $d < 0$, $confidence_{(A=R_t \Rightarrow B=Don'tLike)}$ is used instead of $confidence_{(A=R_t \Rightarrow B=Like)}$, then the recommendation score for the item B will be reduced as much as the value of $confidence_{(A=R_t \Rightarrow B=Don'tLike)} \cdot (confidence_{(A=R_t \Rightarrow B=Like)} + confidence_{(A=R_t \Rightarrow B=Don'tLike)} = 1)$.

α is the parameter for the weight of $confidence$. It is expected the balance between Accuracy and Serendipity can be adaptable by tuning the parameter α , large α leads high Accuracy and small α leads high Serendipity.

V. EXPERIMENT

A. Sample data

In this experiment, MovieLens[11] published by GroupLens[12] was employed. The dataset was linked to Internet Movie Database (IMDb)[13] and Rottan Tomatoes movie review system[14]. The range of rating

scores was between 0.5 to 5.0 with the step size 0.5. The range from 0.5 to 3.5 was regarded as “Don’t Like,” and that from 4 to 5 was regarded as “Like.”

The number of users who rated items as “Like” at least 51 times and as “Don’t Like” at least 50 times was 1118. The number of items which were evaluated by at least 300 users was 611. These users and items were employed for the experiment of recommendation.

B. Methods

Evaluation indexes were calculated by the use of 10-fold cross-validation, one tenth of the dataset was test users and others were training users. And a set of 10-fold cross-validation was conducted 10 times.

In this experiment, test users were treated as target users. An item which a target user rated as “Like” was randomly chosen and used for the first rating history. The other items were treated as non-evaluated items. The number of recommendation was 50 for each target user.

C. Evaluation index

Three evaluation indexes employed in this experiment are shown below. The total number of recommendation was N . The set $\{i_1, i_2, \dots, i_N\}$ was the recommended items. The rating history of i_k was defined as $R(i_k)=1/-1$ (Like/Don’t Like).

$$\frac{1}{N} \sum_{k=1}^N t_k \quad (6)$$

a) Accuracy

“Accuracy” means like ratio of the number of target user’s favorite ($R(i_k) = 1$) items over that of recommended items (N).

$$t_k = \begin{cases} 1 & \text{if } R(i_k) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

b) Novelty[15][16]

In eq. (8), I_{NP} is the set of recommended items provided by Non-Personalized method[15][16]. “Novelty” is the ratio of the number of items which were rated as “Like” and did not appear in the Non-Personalized recommendation set over that of recommended items.

$$t_k = \begin{cases} 1 & \text{if } R(i_k) = 1 \text{ and } i_k \notin I_{NP} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

c) Personalizability[7]

In eq. (9), $P(R(i_k) = 1)$ is the like ratio of item i_k in training users. “Personalizability” is the information quantity based on the lowness of the like ratio. Personalizability becomes higher when recommended items were rated as “Like” and its like ratio is lower.

$$t_k = \begin{cases} \log_2 \frac{1}{P(R(i_k)=1)} & \text{if } R(i_k) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

b) Novelty and c) Personalizability have been proposed as one of the factors of “Serendipity.”

D. Experiment1

Figure 1 shows the comparison between item-based CF in the subsection II-B and the user-based CF in the subsection II-C using association rules. Though Accuracy of the item-based CF was higher than that of the user-based CF, Novelty and Personalizability, which are the evaluation indexes for “Serendipity,” were lower. In the item-based CF, a target user’s rating history served as the association rules’ conditions directly. So Accuracy was high but Novelty and Personalizability tended to be lower.

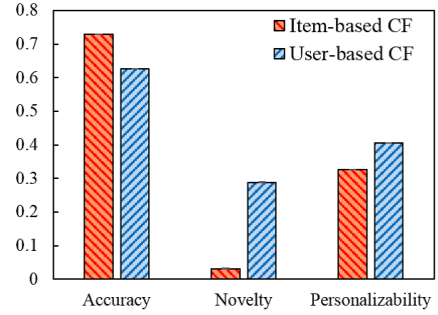


Fig. 1. Comparison between item-based CF and user-based CF

E. Experiment2

Figure 2 shows the comparison of the proposed method with the item-based CF using association rules. When α in eq. (5) increases, Accuracy reaches approximately same level with the item-based CF. In high α , around $\alpha = 1$, by using the rating history of the users who rated different from a target user, Novelty and Personalizability of the proposed method were higher than those of the item-based CF while it kept high Accuracy.

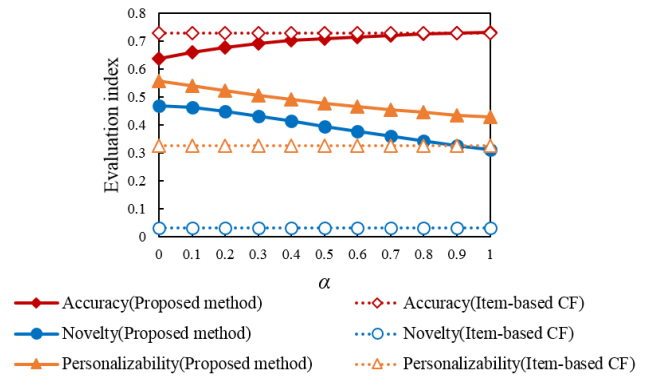


Fig. 2. Comparison of proposed method with item-based CF

F. Experiment3

Figure 3 shows the comparison of the proposed method with the user-based CF using association rules. In the subsection V-D, Novelty and Personalizability of the item-based CF were lower than those of the user-based CF. However, the proposed method was superior to the user-base CF in

Accuracy, while it kept approximately same level in Novelty and Personalizability around $\alpha = 1$. It is shown that the proposed method achieved high Accuracy same with the item-based CF and high Novelty and Serendipity same with the user-based CF.

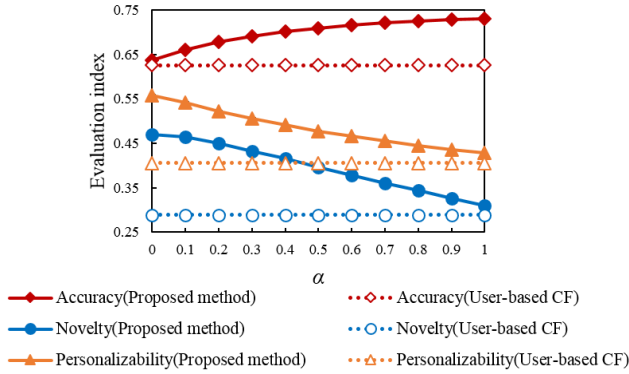


Fig. 3. Comparison of proposed method with user-based CF

G. Experiment4

In this subsection, we compared the proposed method with the conventional method, Weighted Sum shown in the section III. When user u rated item j as “Like,” $R_{u,j}$ becomes 1, and $R_{u,j}$ becomes -1 in “Don’t Like” in the conventional method. Figure 4 and 5 show the comparison of the proposed method with the conventional method ($|I|=300, 610$).

When the number of $|I|$ was 300, Novelty and Personalizability were the highest in the conventional method. In the case that α was lower than 0.3, Novelty and Personalizability of the proposed method were higher than those of the conventional method while it kept that Accuracy was higher.

When the number of $|I|$ was 610, Accuracy was the highest in the conventional method. In the case α was higher than 0.2, Accuracy was higher than the conventional method. In addition, not only Accuracy but also Novelty and Personalizability were higher from 0.2 to 0.4 in α .

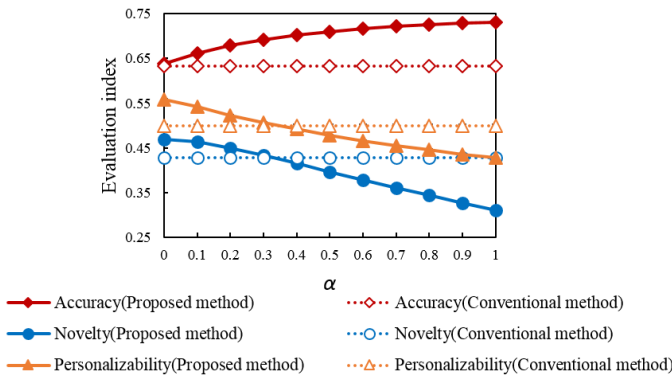


Fig. 4. Comparison of proposed method with conventional method ($|I|=300$)

VI. CONCLUSIONS

In this paper, we proposed the recommendation method using association rules which improved “Serendipity” keep-

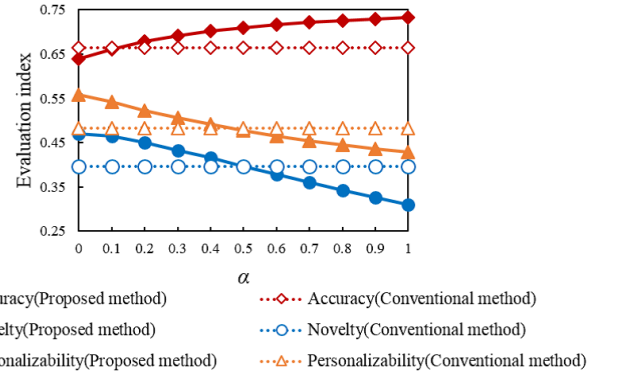


Fig. 5. Comparison of proposed method with conventional method ($|I|=610$)

ing “Accuracy” high by using the evaluation information that were rated differently from a target user. We also showed that “Accuracy” and “Novelty,” “Personalizability” could be adaptable by tuning the parameter α in the proposed method. In the particular values of the parameter, the proposed method was superior to the conventional method in “Accuracy,” “Novelty” and “Personalizability.”

REFERENCES

- [1] T. Kamishima, “Algorithms for Recommender Systems (1) (in Japanese),” *Journal of Japanese Society for Artificial Intelligence*, Vol. 22, No. 6, pp. 826-837, 2007.
- [2] T. Kamishima, “Algorithms for Recommender Systems (2) (in Japanese),” *Journal of Japanese Society for Artificial Intelligence*, Vol. 23, No. 1, pp. 89-103, 2008.
- [3] T. Kamishima, “Algorithms for Recommender Systems (3) (in Japanese),” *Journal of Japanese Society for Artificial Intelligence*, Vol. 23, No. 2, pp. 248-263, 2007.
- [4] Amazon, <http://www.amazon.com/>
- [5] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules in large databases,” *Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487-499, 1994.
- [6] W. Lin, S. A. Alvarez and C. Ruiz, “Efficient adaptive-support association rule mining for recommender systems,” *Data Mining and Knowledge Discovery*, Vol. 6, No. 1, pp. 83-105, 2002.
- [7] T. Yoshikawa, T. Mori and T. Furuhashi, “Proposal of a New Recommendation System that Addresses “Personalizability”,” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 17, No. 2, pp. 167-175, 2013.
- [8] S. M. McNee, J. Riedl and J. A. Konstan, “Being accurate is not enough: how accuracy metrics have hurt recommender systems,” *CHI '06 Extended Abstracts on Human Factors in Computing*, pp. 1097-1101, 2006.
- [9] J. L. Herlocker, J. A. Konstan, L. G. Terveen and J. T. Riedl, “Evaluating collaborative filtering recommender systems,” *ACM Transactions on Information Systems (TOIS)*, Vol. 22, No.1, pp. 5-53, 2004.
- [10] B. Sarwar, G. Karypis, J. Konstan and J. Riedl, “Item-based collaborative filtering recommendation algorithms,” *Proceedings of the 10th international conference on World Wide Web*, pp. 285-295, 2001.
- [11] I. Cantador, P. Brusilovsky and T. Kuflik, “2nd workshop on information heterogeneity and fusion in recommender systems (HetRec 2011),” *5th ACM RecSys*, pp. 387-388, 2011.
- [12] the original Movielens dataset from GroupLens research group, <http://www.grouplens.org>
- [13] IMDb website, <http://www.imdb.com>
- [14] Rotten Tomatoes website, <http://www.rottentomatoes.com>
- [15] H. Chandrashekhar and B. Bhasker “Personalized recommender system using entropy based collaborative filtering technique,” *Journal of Electronic Commerce Research*, Vol. 12, No. 3, pp. 214-237, 2011.
- [16] V. Schickel-Zuber and B. Faltings “Inferring user’s preferences using ontologies,” *Proceedings of the 21st national conference on Artificial Intelligence*, pp. 1413-1418, 2006.