

Personalized Product Service Recommendation Based on User Portrait Mathematical Model

Xuesheng Lai ^a, Lili He ^b, Qingyan Zhou ^c

School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China.

^a15988816563@163.com, ^bllhe@zju.edu.cn, ^cswallowchou@126.com

Abstract. In the “Internet+” digital age, the use of customers’ information was not fully utilized when pushed the information. The traditional similarity exists some problems, such as incalculable, indistinguishable and high-level etc., it results in poor pertinence of personalized product services and a decline in the quality of recommendations. Concerning this shortcoming, this paper proposes a collaborative filtering recommendation algorithm based on the user’s portrait model customer rating, and use the customer’s rating results to make recommendations. To this end, a “user portrait” mathematical model was constructed, the similarity was improved by using a discrete-volume correlation theory and was weighted approaching to the users’ preference. It would be more accurate for “K” nearest neighbor set by similarity calculation. Furthermore, this paper recommends more suitable products to users. The experiment was conducted with the customer’s sales data of Le Bee net Cosmetics, it shows that the method proposed in the paper improves the accuracy of the recommendation effectively and improves the recommendation quality to some extent.

Keywords: big data, user portrait, customer rating, similarity, collaborative filtering.

1. Introduction

In the digital age, more and more data have been generated day by day, the user’s shopping behavior will produce trajectory. These tracks reflect the habits, personalities, interests and other information of consumer behavior to a certain extent. At the same time, these scattered data are collected, recorded and stored. After a series of processing and handling, the data can be transmitted to the user in the form of valuable information and reconstruct the consumer’s demand [1]. According to the user’s basic characteristics, interest, behavior etc which can abstract model. The tagged model is called “user portrait” [2-3]. The personalized recommendation system can be used to help companies recommend the most suitable product to users in a large amount of information, thereby improving the user’s awareness of the product.

Currently, the most widely used personalized recommendation system is the collaborative filtering recommendation algorithm [4-5]. Goldberg D [6] was the first to propose recommendation system based on collaborative filtering algorithm, and it also indicates that the target users are recommending other users with their own behavior types. Cai [7] studied model-based citation recommendation methods, through initiating random walks for individual query researchers in personalization recommendations to create personalization on subgraphs in three-tier interactive clustering methods. Pirasteh P [8] designed an adaptive weighted similarity calculation method that considers user non-common scoring items and some user’s behavioral habits. Park [9] proposed a fast collaborative filtering algorithm, RCF, which uses a K-nearest neighbor graph, uses the items which had been graded to find K-nearest neighbors. This paper improves the similarity calculation, propose a collaborative filtering algorithm that integrates discrete variables and approach to user preference. Through the customer score value obtained by the forecast, the paper finds the user sets with the most similar interest and recommend the products which the users are interested in. That is the Top-N recommendation.

2. The Construction of "User Portrait"

According to the demand characteristics of the market, user portrait model was constructed by consulting literature, research, interviews and suggestions of experts, as shown in Figure 1.

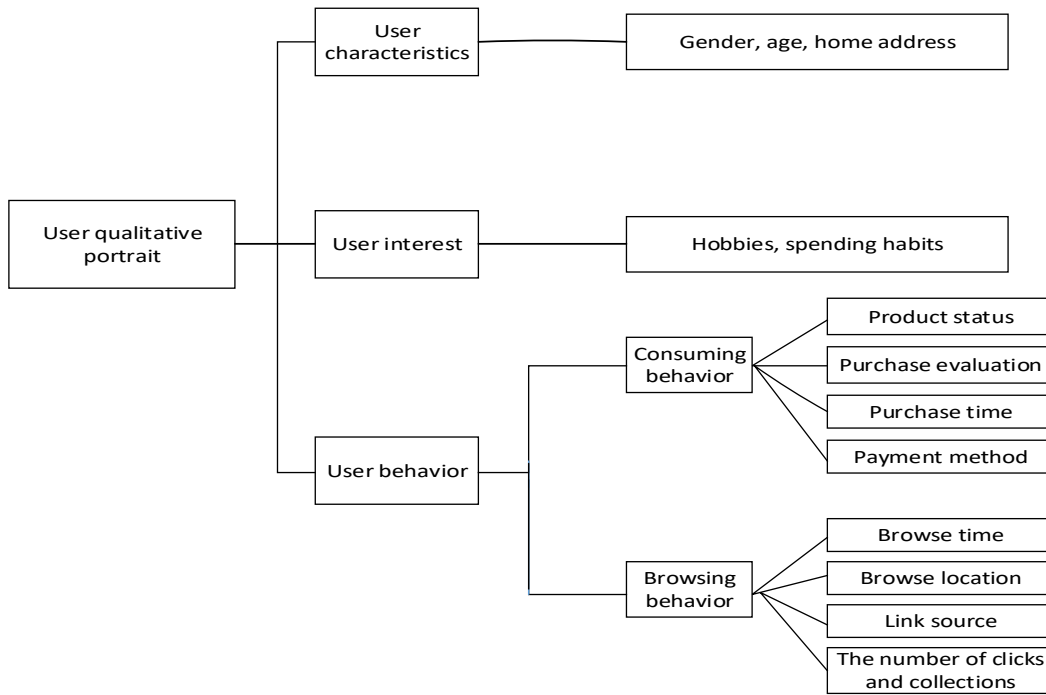


Fig. 1 User portrait

3. Research on Recommendation Technology Based on User Portrait Customer Ratings

In this paper, by improving the traditional collaborative filtering algorithm, as shown in Figure 2.

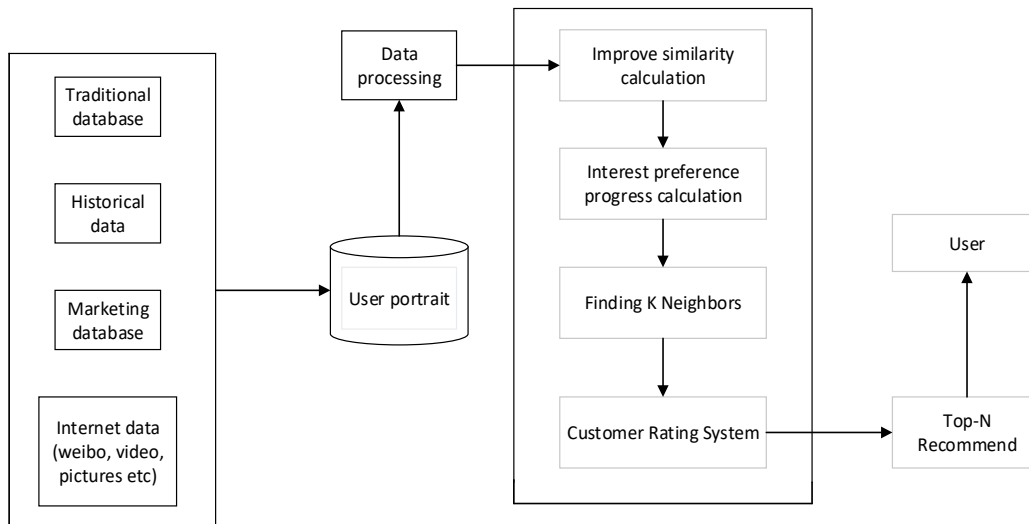


Fig. 2 Mathematical model of user portrait customer rating system

3.1 Calculation of Traditional Similarity

The traditional common similarity calculation methods include modified Pearson (CPCC), Pearson similarity (PCC), Co-sine similarity (COS), modified Pearson (CPCC), adjusted Cosine similarity (ACOS) and so on. Suppose the user rating matrix is represented by the matrix $A(m,n)$, the m rows represent m users, the n columns represent n items, and the element of the row J of the I row represents the rating table 1 of the user I to the J of the project, which is a Person-Project scoring matrix, and uses 6 users to score the 4 items, where "-" indicates missing user ratings.

Table 1. User Rating Matrix

	Project 1	Project 2	Project 3	Project4
Person1	4	3	4	5
Person2	4	3	3	4
Person3	5	2	-	-
Person4	2	3	-	-
Person5	-	-	4	-
Person6	5	4	2	-

3.2 Similarity Calculation Method for Fusion Discreteness and User Interest Post Progress

The similarity calculation method proposed in the literature [10]. In the calculation process, the reliability factors $\alpha \neq 1$ and $\alpha = 1$ between users are divided into two cases. The similarity calculation method proposed in this paper is expressed as formula (1).

$$sim(u, v) = \begin{cases} I_1 \sigma_{u,v}(U, V) & \alpha \neq 1 \\ I_1 \cdot \sigma_{u,v}(U, V) & \alpha = 1 \end{cases} \quad (1)$$

Table 2. Results of User Interest Preference and Progress Correction

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6
Person1	0.229	0.803	0.311	0.232	0.4170	0.220
Person2		0.242	0.700	0.437	0.399	0.5799
Person3				0.246	0.501	0.293
Person4				0.470	0.423	0.595
Person5					0.541	0.191
Person6						0.245

As can be seen from the table, similarity calculation results calculated by the algorithm can reflect the degree of similarity accurately and reasonably between users, and improve the similarity between users. Thus, the core similarity calculation problem is solved for the recommendation system, which has benefit of scoring prediction and recommendation subsequently.

3.3 Description of Collaborative Filtering Algorithms Based on Discrete and User Interests

Algorithm 2-1. Collaborative Filtering Algorithm Based on Discreteness and User Interest Preference Progress

Input: scoring matrix R, target user u and nearest neighbor K

Output: Target user u's customer rating forecast for item i

1. for all users in test sets do
2. the discrete values of user u and v are calculated.
3. the discrete values of the user u and v are calculated,
4. the discrete sums of user u and v scores are calculated.
5. calculated information coefficient measure.
- 6.calculate the final similar result.
7. finding K-nearest neighbors of active users u.
8. calculate user's score prediction
- 9.end for
10. calculate the MAE value of the test score in the test set.

4. Experimental Results and Analysis

4.1 Datasets

The similarity the experimental data was mainly collected on the data from 2017-12-07 to 2018-01-08 at <http://www.lefeng.com/>. After statistics, 1103 cosmetics were scored by 983 users. The data

set for the available information for the user is 110034, and the score is an integer of 1~5. The data sparseness level calculation method is 2.

$$1 - \frac{110034}{1103 \times 983} = 0.898 \tag{2}$$

4.2 Experimental Metrics

The experimental results in this paper are evaluated using MAE, which is one of the criteria for evaluating the quality of recommended algorithms. The accuracy of the recommendation results is measured by calculating the difference between the predicted score and the actual evaluation data. The smaller the MAE value, the higher the accuracy of the recommendation. Suppose the user's predicted score set is denoted by $M = \{m_i | i = 1, 2, \dots, n\}$ and the corresponding user's actual score set is $N = \{n_i | i = 1, 2, \dots, n\}$, n is the number of articles, then the MAE calculation formula is as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |m_i - n_i| \tag{3}$$

4.3 Determining the Number of Optimal User Neighbors

In this experiment, the data set was divided into 20 sets, 17 of which were used as training set and 3 were used as test set. Experiments were performed several times to obtain the average value. This method can prevent data from being unevenly distributed among data sets. At the same time, verifying the mean value can reduce the error caused by random sampling effectively.

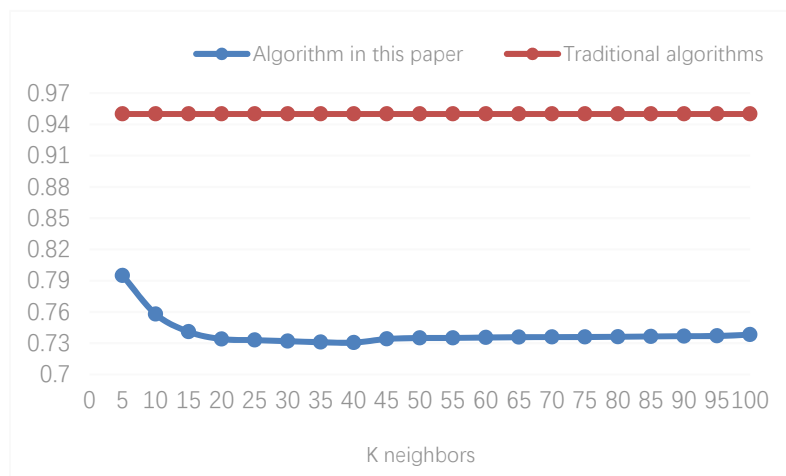


Fig. 3 Best Neighbor K

The purpose of this experiment is to introduce the mathematical model of the user portrait customer rating system and use the improved similarity to find the optimal neighbor number K . Compared with the traditional algorithm, it shows that it is superior to the traditional algorithm. In Fig. 3, the initial value is set to 0, progressively increased by 5, and the experiments are performed successively to find the average value. The results are shown in the figure. When the algorithm of this paper is less than 20, the MAE value of the algorithm decreases sharply. As the K number increases, it can be obtained that the calculated MAE value is the smallest, when the number of neighbors of the user reaches around 40. When the K value is greater than 40, the MAE value does not change very much and tends to be stable, so the optimal user neighbor number is selected to be 40.

4.4 Comparison with Other Algorithms

This experiment is mainly to verify the improved algorithm compared with Cosine, Pearson and Spearman and other traditional algorithms to compare the MAE index. The unrated items of the user are predicted on the test set, and the MAE value of the collaborative filtering algorithm that does not

use similarity metrics is used, and the evaluated index is the MAE value under different K values as shown in Figure 4:

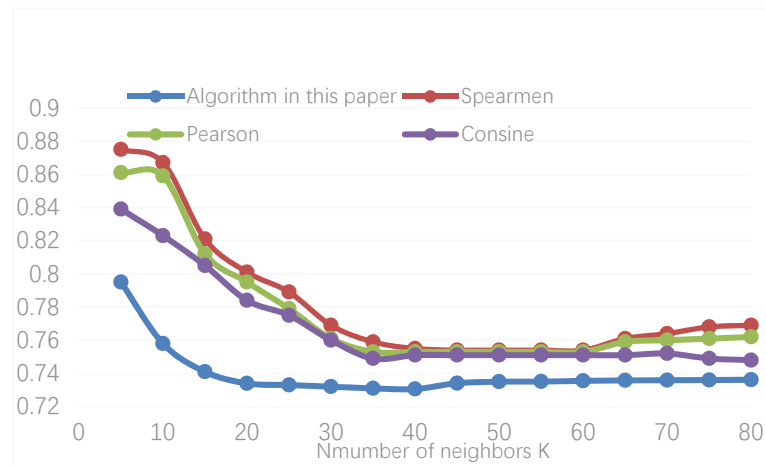


Fig. 4 Comparison of nearest neighbor K

Figure 4 shows the Pearson correlation, Spearman correlation, and cosine similarity on the same data set, and the MAE value of different neighbors for this algorithm. As can be seen from Figure 4, after introducing the model of the user portrait customer rating system, it is more accurate to calculate the similarity between the users' common favorite products. The MAE value of the improved algorithm is always less than the MAE value of the other three similarity measurement methods, and the best one is obtained when the nearest neighbor number is 40. By comparison, it can be shown that the algorithm of this paper is superior to other algorithms in recommendation quality, and can significantly improve the service of recommending personalized products.

5. Conclusion

This article aims to improve personalized product services. Therefore, this paper proposes a method of calculating the similarity between the discrete magnitude and the user's preferences to find the nearest neighbor number. The improved algorithm is applied to the personalized product recommendation function. It verifies that this method makes the target user's nearest neighbor number more accurate. It can solve the drawbacks of the traditional similarity measurement methods effectively and improve the quality of products and services of the recommendation system significantly. It provides a reference for corporate decision makers to implement personalized marketing strategies, and makes scientific and rapid decisions.

References

- [1]. Hai Liu, Hui Lu. Research on Accurate Marketing Segmentation Model Based on User Portrait Mining[J]. *Journal of Silk*, 2015, 12(52):37-42.
- [2]. Xie Ran. Seeing the application logic from the five sectors of large data case[J]. *China Internet Weekly*, 2015(2):30-35.
- [3]. Zhang Peng, Liu Yijing. Consumer map[J]. *China Market: Channel Edition*, 2013(9):30-32.
- [4]. Pu P, Chen L, Hu R. A user-centric evaluation framework for recommender system[C] // *Proceedings of the Fifth ACM Conference on Recommender Systems*, 2011.
- [5]. Knijnenburg B P, Willemsen M C, Gantner Z, etc.. Explaining the user experience of recommender systems[J]. *User Modeling and User-Adapted Interaction*, 2012, 22(4):441-504.
- [6]. Goldberg D, Nichols D, Oki BM, Terry D. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 1992, 35(12):61~70.

- [7]. Tsai C F ,Hung C.Cluster ensembles in collaborative filtering recommendation.Applied Soft Computing,2012.12(4):1417-1425.
- [8]. Pirasteh P,Hwang D,Jung J E.Weighted Similarity Schemes for High Scalability in User-Based Collaborative Filtering.Mobile Networks and Application,2015,542:236-242.
- [9]. Park Y,Park S,Jung W,Lee S G.Reversed CF:A fast collaborative filtering algorithm using a k-neighbor graph.Expert Systems with Application,2015,42(8):4022-4028.
- [10]. Shaocheng Ding, Donghong Ji,Lulu Wang.Collaborative filtering recommendation algorithm based on user attributes and scores[J].Computer Engineering And Design,2015.36(2):487-491.