IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Smart E-Commerce Recommendation System for Handling Limited Resource and Cold Start Problem

ThejaswiniN[1*], Aditya C R[2]

[1,2]Dept. of Computer Science and Engineering, VVCE, Mysuru

DOI: https://doi.org/10.26438/ijcse/v7i5.961964 | Available online at: www.ijcseonline.org

Accepted: 19/May/2019, Published: 31/May/2019

Abstract **–**With the development of E-commerce, Recommendation Systems are applied more widely to guide the customers to search for their interested products. A recommendation system includes a user model, a recommended model and a recommendation algorithm. Limited resource, data valid time and cold start problems are not well considered in existing E-commerce recommendation system. This paper proposes a limited resource based algorithm to provide an improvement to the existing product recommendation algorithm and also provides a solution to cold start problem.

Keywords **–**Limited resource, cold start, recommendation system.

I. INTRODUCTION

Data mining is a process of extracting the valuable information from the large amount of stored information which is related to various organizations that perform different functions and operations.

With the popularity of the usage of internet, the E-commerce system structure has become more complex in recent years. This has made harder for customers to find their interested products and the services effectively with minimum effort and time. E-commerce recommendation provides information and suggestions to customers to make proper purchasing decisions.

Even though the E-commerce recommendation system is successful, it faces lot of challenges with the development of the E-commerce. Nowadays, most of the E-commerce websites such as flipkart and amazon adopt memory based algorithms which are essentially heuristic, as in the user-based KNN system which uses statistical techniques to find the users with similar tastes and predicts a target item. Even though the existing recommendation systems are able to provide the recommendations for customer, there are some problems that need to be solved to improve customer satisfaction.

In this paper, we have addressed various factors that need to be taken into account for improving the user satisfaction.

1. *Limited Resource Situation*: For instance, personal recommendation for limited products with the special offer period. These kinds of products should be recommended with the priority, which provide the satisfaction to the consumer while buying the product.

2. *Data Valid Time*: The records which were recorded long time ago are not allowed to be used for recommendation process because those records are not accurate anymore.

3. *Cold start problem*: It is the problem faced by customers who visit the E-commerce website for the first time, which is not well solved in the existing E-commerce recommendation system.

An efficacious E-Commerce recommendation system can give out productive recommendations for customers which can be accepted by customers as far as possible.

II. LITERATURE SURVEY

In this section we briefly present some of the research literature related to recommendation systems, collaborative filtering and data mining.

Researchers have already classified many algorithms for collaborative recommendation including the memory based algorithms [2],which are essentially heuristic as in the user based KNN system which calculates the prediction of a target item based on statistical techniques in order to find users with similar tastes.

Zhimin Chen et. al. [3] have proposed an enhanced decision table classification approach for classification of online shopping dataset.Methods based on probability statistics have been used to assess clicking behavior data[4].

There are other technologies applied to recommender systems such as Horting, Clustering and Bayesian Network [5]. Bayesian Network creates a model based on a training set with a decision tree with nodes representing user information.

Shikhar Kesarwani[6] proposes a MSD-Apriori to discover the borderline rare elements which are very close and below to minimum support threshold and have strong correlation with frequent item. It formed by integrating MS Apriori with Dynamic Apriori algorithms.

Tapestry [10] is one of the earliest implementation of recommendation system of collaborative filtering. The explicit opinions of people from a close knit community are relied in this system, such as an office workgroup. After that several ratings-based automated recommender systems were developed.Video recommender [12] and Ringo are email and web based systems that generate recommendation music and movies respectively.

III. PROPOSED METHODOLOGY

With the development of E-Commerce, personalized recommendation has been paid more and more attention. Limited resource situation, data valid time and cold start problems have not been well considered in existing Ecommerce recommendation system. This paper proposes limited resource table method, an algorithm based on limited resource, and a solution to cold start problem which can enhance the effect of recommendation system. The solutions proposed here are meaningful for E-Commerce websites and recommendation system.

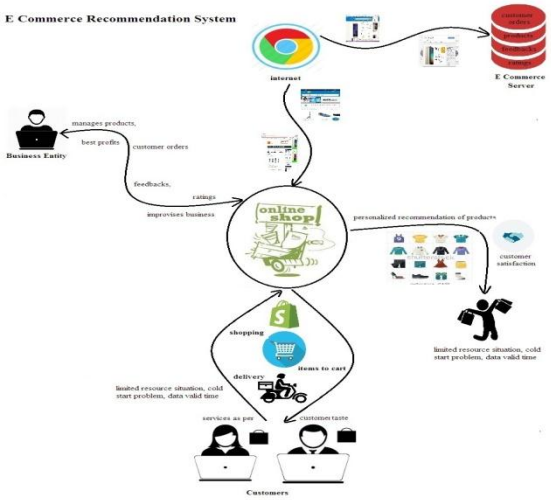

Figure 1: System Architecture

1. Hybrid Model for Limited Resource Situation with History(Existing user)

This model is used to handle limited resource situation for users with purchase history. Basically, it is a mixture of content and collaborative filtering algorithms.
It uses user previous orders and rating to handle the limited resource situation.Top 10 products will be recommended for the users. The hybrid model consists of content filtering and collaborative filtering methods [2]. The model proposed consists of methods for Recommendation process, Generation of neighbors and Generations of recommendations.

a) Recommendation Process

On the basis of collaborative filtering principle, the recommendation process of consumer's attractions can be divided into three steps;

1) *The representation of user information*: The purchasing history of attractions by consumer need to be analyzed and modeled.

2) *The generation of neighbor users (consumers):* The similarity of consumers can be computed according to the buying history data and the collaborative filtering algorithm. A neighbor consumer list can be calculated on the basis of known similarities.

3)*The generation of attraction recommendations:* Top-N attractions will recommended to the consumer according to the buying history of his neighbors.

According to above steps, user's basic information and past purchasing history can be used to calculate the user list of neighbors.

b)Generation of Neighbors

Neighbor users are generated mainly based on the similarity between each user. Suppose that the set of all consumers S={S1, S2... Sn}, for each consumer Si (i=1, 2... n), the system can calculate the neighbors list including the top N consumers with similarity higher than the given threshold. The similarity between customers is calculated using a similarity index given by equation(1).

$$\text{sim}(Ti, Tj) = \frac{|Si \cup Sj|}{|Si \cap Sj|} \qquad (1)$$

Table 1. Purchasing History of the customer

	Attractions				
Customers	**A1**	**A2**	**A3**	**A4**	**A5**
T1	True	False	True	True	True
T2	False	True	True	True	True
T3	True	False	True	True	False
T4	False	True	True	False	True
T5	False	True	False	True	False

Based on equation-1 and a sample purchasing history shown in Table I, we can calculate the similarity between T1 and

T2, T1 and T3, T1 and T4, T1 and T5 as follows. Si and Sj are the similarity between first user and second user.

$$\text{sim}(T1,T2) = \frac{|S1 \cup S2|}{|S1 \cap S2|} = \frac{3}{5} = 0.6 \quad (2)$$

$$\text{sim}(T1,T3) = \frac{|S1 \cup S3|}{|S1 \cap S3|} = \frac{3}{5} = 0.6 \quad (3)$$

$$\text{sim}(T1,T4) = \frac{|S1 \cup S4|}{|S1 \cap S4|} = \frac{2}{5} = 0.4 \quad (4)$$

$$\text{sim}(T1,T5) = \frac{|S1 \cup S5|}{|S1 \cap S5|} = \frac{1}{5} = 0.2 \quad (5)$$

Equation (2) shows the calculation of similarity between T1 with T2 with respect to Table I.
If the value of threshold Θ is set to be 0.5, then the neighbors of T1are T2 and T3.

b) Generation of Recommendations

Recommendations of attractions are computed by the purchasing times of neighbors. According to the calculation above, we know that the neighbors of customer T1 are T2 and T3, so we can list all the purchasing history of all the attractions by the neighbors so as to summarize the most popular ones. As listed in Table II, we can find that the maximal purchasing by neighbors are attraction A3 and attraction A4.

Table 2. Purchasing History of neighbor customer

Customers	A1	A2	A3	A4	A5
T1	True	False	True	True	True
T2	False	True	True	True	True
T3	True	False	True	True	False
Total	2	1	3	3	2

2. Random Algorithm Used for Cold Start Problem (New User)

When new customers enter the system, there is usually insufficient information to produce recommendation for them, because there is no purchasing history of the new consumers. The usual solution of the **cold start problem** is similarity calculation between each user by profile information, such as user area of interest, gender, Age, etc.

This algorithm uses user profile and matches with the existing users to identify the neighbours and based on the orders of neighbours, system will handle limited resource situation.

Step 1: Consider the user's area of interest, age, gender, location, marital status etc.
Step 2:Then we consider similar users of same interest and their transactions.
Step 3:Calculate the trading volume of each product

Trading volume is calculated using the following formulae;

Volume = transaction containing that product/total number of transactions

Step 4:Store it in array list, then sort and reverse
Step 5:Retrieve top 10 products as recommendation result.

IV. RESULTS AND DISCUSSION

The proposed methodology was implemented on an E-commerce application with self-populated data. The application was tested on 20 users, among which 15 were considered to be regular customers and 5 were cold start customers.

Table 3. Shows the purchase history of regular customers.

CUSTOMER SL.NO	ITEMS ID	TIME OF PURCHASING
1	2,8,12,5	10/4/2019,11:30AM
2	1,3,4,12	2/4/2019,5:30PM
3	6,8,21,7,1	8/4/2019,3:30AM
4	8,4,13,2	5/4/2019,10:00AM
5	2,3,17,20	9/4/2019,11:00AM
6	7,14,2,4	9/4/2019,1:30PM
7	1,4,12,3	10/4/2019,9:30PM
8	4,13,1	14/4/2019,2:45PM
9	2,8,12,20	15/4/2019,11:00PM
10	2,3,19,4	19/4/2019,2:00PM
11	6,14,11,9	20/4/2019,10:50AM
12	2,19,7,11	22/4/2019,12:05PM
13	2,5,7,8,9	26/4/2019,1:30PM
14	1,3,6,8,9	29/4/2019,7:40PM
15	12,5,8,10	29/4/2019,10:25PM

The feedback from these customers (Regular and cold start) was obtained to access the precision of the system.

Table 4. Shows the Item recommended item based on Area of Interest (AOI) for Existing user.

CUSTOMER SL.NO	Recommended Item(ID)	Customer satisfaction (0-10)
1	2,3,12,4,8,5,1	7
2	6,14,19,4,2,3,11	8
3	2,8,12,5,9,7,20	8
4	13,4,2,1,9,3,6,8	8
5	2,7,12,3,20,	7
6	1,4,12,3,13	7
7	20,12,2,8	8
8	12,17,8,4	9
9	1,3,6,5,8,9,10,2	9
10	6,14,11,9,1,3,6,8,9,12,5,8,10	10
11	2,3,19,4,8,1,	6
12	12,8,5,2,4,11,3	8

13	2,3,1,6,5,8,9,10	7
14	2,7,8,9,12,5,10	7
15	12,17,8,4,2,7,3	8

Table V. shows the items that are displayed to the customers that were of their interest for new user.

CUSTOMER SL.NO	Recommended Item(ID)	Customer satisfaction (0-10)
16	1,4,12,3,13	8
17	2,5,7,8,9,4,13,1	9
18	1,3,6,8,9,4,13	8
19	1,2,3,4,17,6,8	5
20	2,6,5,14	8

From the results obtained it was evident that both regular and cold start customers were satisfied with the application. Compared to other existing systems, even cold start customers were able to find their items of interest effectively using the proposed methodology.

All the customers were satisfied with the personal recommendation given for limited product. Also the data valid time problem was effectively handled.

V. CONCLUSION

In this work, a new recommendation method is proposed to improve customer satisfaction in buying their interested product. Some of the major factors like limited resource situation, data valid time and cold start problem are solved in the proposed system by using the hybrid algorithm for existing user and random algorithm for the new user.

REFERENCES

[1] B.Sarwar, G.Karypis, J.Konstan, J.Riedl, "Analysis of recommendation algorithms for E-commerce", In: ACM Conference on Electronic Commerce, pp.158-167, 2000.

[2] Billsus D., and Pazzani M. J., "Learning Collaborative Information Filters", In Proceedings of ICML '98, pp. 46-53, 1998.

[3]ZhiminChen, Yi Jiang, Yao Zhao, "A Collaborative Filtering Recommendation Algorithm Based on User Interest Change and Trust Evaluation",International Journal of Digital Content Technology and its Application,vol.4,November 9, December2010.

[4] G.Karypis, "Evaluation of item-based top-N recommendation algorithms," in Proc. of CIKM 2001, pp. 247–254, 2001.

[5] ShikharKesarwani,AsthaGoel and Dr.NeetuSardana,"MSD-Apriori:Discovering Borderline-rare items using Association Mining",In proceedings of 2017 Tenth International Conference on Contemporary Computing (IC3),10-12 August 2017,Noida,India.

[6] Herlocker L. J, KonstanA.J, Terveen G. L, et a1, "Evaluating collaborative filtering recommender systems", ACM Transactionon Information Systems, vol. 22, no.1, pp. 5-53, 2004.

[7]Basu, C., Hirsh, H., and Cohen, W. (1998). Recommendation as Classification: Using Social and Content-based Information in Recommendation. In Recommender System Workshop '98. pp. 11-15.

[8] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992). Using Collaborative Filtering to Weave an Information Tapestry.Communications of the ACM.December.

[9]Xue, W., B. Xiao, et al. (2015). Intelligent mining on purchase information and recommendation system for e-commerce.2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).

[10] Goldberg D, Nichols D, Oki B M, et a1, "Using collaborative filtering to weave fin information Tapestry", Communications of the ACM, vol. 35, no.12, pp. 61-70, 1992.

[11] Li, D., G. Zhao, et al. (2015). A Method of Purchase Prediction Based on User Behavior Log. 2015 IEEE International Conference on Data Mining Workshop (ICDMW).

[12] Ramezani, M. and F. Yaghmaee (2016). "A novel video recommendation system based on efficient retrieval of human actions."Physica a-Statistical Mechanics and Its Applications 457: 607-623.

[13] Yi, Z., D. Wang, et al. (2015). Purchase Behavior Prediction in MCommercewith an Optimized Sampling Methods. 2015 IEEE International Conference on Data Mining Workshop (ICDMW).

[14] Schafer, J. B., Konstan, J., and Riedl, J. (1999). Recommender Systems in E-Commerce. In Proceedings of ACM E-Commerce 1999 conference.

[15] Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). Recommending and Evaluating Choices in a Virtual Community of Use. In Proceedings of CHI '95

