

# A Product Recommendation System Based on Hybrid Approach

Shivani Gore<sup>1</sup>, Jayashri Sonawane<sup>2</sup>, Laxmi Duggineni<sup>3</sup>

Department of Computer Engineering

<sup>1,2,3</sup> JSPM's Imperial College Of Engineering and Research, Wagholi, Pune.

**Abstract-**Product recommendation system are of great significance in e-commerce. Current recommendation approaches made great achievements, but do not consider different people's similar learning trajectories in the same major. This paper focuses on "book" as a product. In order to recommend books more accurately, this paper proposes a personalized book recommendation algorithm that is based on the time sequential collaborative filtering and content based collaborative filtering. In this algorithm we consider important factors: time sequence information of purchasing books of different users in database and content of the user profile. Here, results of both the algorithms are unified for obtaining the more accurate list of recommended books. This hybrid book recommendation system satisfies user by providing best and efficient books recommendations.

**Keywords-**Book Recommendation, Time Factor, Category of books, Content Based Algorithm, Time Sequence Based Collaborative Filtering .

## I. INTRODUCTION

Nowadays, vast amount of information is available and its difficult for user to find relevant information. Information related to books, songs, movies is available online in a scattered format. Recommender system are helpful to individuals who do not know the number of alternative for website for specific type of item or who lack sufficient experience. For example, A book recommender system suggest users to select the books to be read.

The e-commerce recommendation algorithm operate in a challenging environment as the scenarios change with respect to new customer. Why filtering?? The answer to this is can be the information overload existing in the current world. Too many movies, books, cameras, web pages etc where searching is difficult. For this purpose recommendation system came into existence[1]

There are various recommendation strategies which include content based recommendation, collaborative filtering - based recommendation, Association - based recommendation, Utility-based recommendation, Knowledge-based recommendation[2]. In this paper we propose a product

recommendation system based on hybrid approach. The term hybrid recommender system is used here to describe any recommender system that combines multiple recommendation techniques together to produce its output.

Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice versa); or by unifying the approaches into one model. Several studies empirically compare the performance of the hybrid with the pure collaborative and content-based methods and demonstrate that the hybrid methods can provide more accurate recommendations than pure approaches. These methods can also be used to overcome some of the common problems in recommender systems.

We have used two algorithms:

- A. Content based filtering algorithm
- B. Time Sequence Collaborative Filtering Algorithm

Time sequence-based recommendation algorithm adds time sequence information into existing information model. This algorithm helps us to learn changing data over time. Using recommendation system users identify one or more items as per their interest.

In content based algorithm, the recommenders suggest other items that are similar, based on comparison of item features or user features. Content-based filtering methods are based on a description of the item and a profile of the user's preference. In a content-based recommender system, keywords are used to describe the items and a user profile is built to indicate the type of item this user likes. In other words, these algorithms try to recommend items that are similar to those that a user liked in the past (or is examining in the present). In particular, various candidate items are compared with items previously rated by the user and the best-matching items are recommended. Description of items in terms of attributes for example, Directors, Actors. Description can also be stated via keywords which refers the user profile for example, in the restaurant the cuisines recommended to the customer would be dynamically generated based on his

choices like French, Italian, Indian according to his profile stated. If customer is Indian, Indian cuisines will be recommended to him.

**II. RELATED WORK**

Recommendation Systems have been widely used in e-commerce. Many domains such as book, electronic-products, music, movies etc. uses recommendations[2].In last few years, rapid development of information technology, an increasing number of books are available on e-commerce websites. Recommendation system helps user to find relevant books.

Recently many scholars have made significant progress in recommendation systems. Luet al.[3] proposed content based filtering and collaborative filtering recommendation methods. Antonio Hermando et al proposed a prediction method of collaborative filtering recommendation based on collaborative filtering for rating of users based on Bayesian probabilistic model[4].

Kouki et al.[5] designed hybrid probabilistic extensible hybrid recommendation method, which could automatically learn and make predictions. Typical model based collaborative filtering includes the clustering techniques based collaborative filtering, probabilistic method based collaborative filtering and matrix based decomposition collaborative filtering[6].

The ratings of related items usually calculated in process of recommendation, but related time sequence behavior information easily recommendation, but related time sequence behavior information is easily ignored. Many authors and proposed the recommendation algorithm based on time sequence information for this problem.

Time sequence based recommendation algorithm adds time sequence into the existing recommendation model. This algorithm enables the model to learn the data changing over time. Hence the accuracy of recommendation results would be improved. Time sequence algorithm is newly introduced and hence if it is combined to the existing algorithm, result achieved will be accurate and efficient. Hence, current scholars uses hybrid approaches like same.

Gao.et.al[7] developed an improved collaborative filtering recommendation also with time adjusting. A real time stream based recommendation algorithm was proposed based on collaborative filtering[8].Some authors also add time sequence information into feature vectors of product[9].

There are also some scholars assigning products to different clusters dynamically by evolutionary combined clustering and make final recommendation[10].

**III. PROBLEM STATEMENT**

PROBLEM STATEMENT: To develop a recommendation system for e-commerce as a tool that emphasizes on:

- A) Accuracy of recommendation
- B) User’s ease of product purchase

**IV. METHODOLOGY**

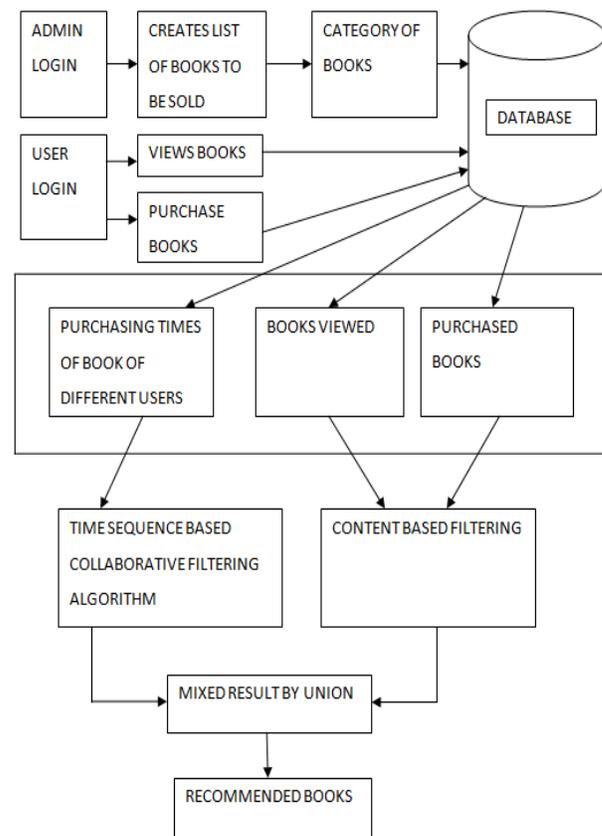


Fig: System Architecture

**4.1: Algorithms:**

1) Time sequence collaborative filtering algorithm: When user purchase a book or search a book on the website, the system will recommend related books by time sequence based algorithm. Algorithm uses distance to denote the many relationship between books. The shorter distance means the closer relationships.

Table 1: User’s book purchasing record.

| Average purchasing time | User 1 | User 2 | User 3 |
|-------------------------|--------|--------|--------|
| T1                      | b1     | b2     | b4     |
| T2                      | b3     | b5     | b2     |
| T3                      | b2     | b3     | b1     |
| T4                      | b6     | b1     | b2     |

In the above table purchasing records of the books of different users is shown, where b1,b2..bN are books purchased.T1,T2....TN are average time stamp of book purchase. For example:

The new user who wants to purchase book b2,then average time distance of the books purchased after b2 is calculated for each user having purchase record of b2. Finally, the books having minimum average timestamp will be recommended to the user from minimum to maximum timestamp order.

$b_j$  = The book that users purchase or view.

$b_k$  = A book appeared after  $b_j$  in the purchase set of all the users who purchased  $b_j$

$T_k$ = Time sequence information of  $b_k$ .

$N_k$ = No. of times  $b_k$  appears after  $b_j$  in all sets.

$N'$ =  $\max(N_1,N_2,N_3,....)$ .i.e. biggest time interval of book purchase.

$t_{ik}$  = Purchasing time of  $b_k$  for single user.

$t_{ij}$  = Purchasing time of  $b_j$  for single user.

**Algorithm is described as follows:**

- Step1: Search users in database with same book purchase history.
- Step2: Calculate the time distance between current book and other books purchased by each user.
- Step3: Average time distance of each purchased book with respect to each user from current book.
- Step4: Recommend books whose average time distance is minimum.
- Step5: Stop.

$$dis(b_j,b_k) = \sqrt{T_k^2 + (N' - N_k)^2} \dots\dots (1)$$

$$T_k = \sum_{i=1}^m (t_{ik} - t_{ij}) / N_k \dots\dots (2)$$

Above equations are used to calculate the time-distance between( $b_j,b_k$ ) and results will be displayed from minimum to maximum order of time-distance.

**Content based collaborative Filtering algorithm:**

**Algorithm is described as follows:**

- Step1: Take input from user.
- Step2: Search features of book name entered by user.
- Step3: If features available  
Then recommend books having matched feature
- Step4: Stop

**4.2 Advantages:**

1. In content based algorithm the recommendation of the users is provided based on their set of interest and in Time Sequence algorithm recommendation is provided based on the minimum time stamp of all users where varying interest of other users is considered.
2. Thus, hybrid system considers similar as well as different interest of users making it more efficient.
3. User rating is optional feature in our recommendation system which gives customers the freedom of choice irrespective of which the recommendation will be done.
4. A better understanding of user needs is fulfilled through the hybrid approach.
5. In the comparison to the number of items sold without recommendation system, recommender systems are able to sell additional set of items.

**V. EXPERIMENTS AND ANALYSIS**

Experimental data sets are sample books searched and purchased records.

**EXPERIMENT 1: CALCULATE THE ACCURACY AND RECALL RATE OF THE ALGORITHM**

We select five users at random and use the recommendation algorithm to give readers Top 5 recommended books (five books). The books on connecting lines between users and recommendation results are recommended for different users successfully. We analyze the number of successfully recommended books and calculate the precision rate and the recall rate (the calculation results are

shown in Table 2). Precision and recall rate are two metrics that are used for estimating the qualities of recommended results. The equations used in this paper are shown as follows:

Precision Rate = a number of books successfully recommended / a number of books actually recommended.

Recall Rate = a number of books successfully recommended / a number of books purchased

Table 2: Results of recommendations for different users.

| User | Apurch-<br>asig<br>amount | Actual recom-<br>mendation<br>amount | Success-<br>ful recomme<br>ndation | Precisio<br>n rate | Recall<br>rate |
|------|---------------------------|--------------------------------------|------------------------------------|--------------------|----------------|
| 1    | 2                         | 5                                    | 3                                  | 0.6                | 1.5            |
| 2    | 7                         | 5                                    | 0                                  | 0.00               | 0.00           |
| 3    | 10                        | 5                                    | 2                                  | 0.4                | 0.2            |
| 4    | 24                        | 5                                    | 3                                  | 0.6                | 0.12           |

The precision rate of Top5 is 0.4(the average of precision rates in Table 2) while recall rate is 0.45 (average recall rates in table 2). Compared with other recommendation algorithms, both of them are lower due to influence by the small number of book duplicates.

If we exclude these books, the precision rate and recall rate could be enhanced. The recommendation algorithm in this paper is built on the time sequence information which is closely related to the process of professional knowledge acquisition and recommend a book set related to our future professional learning. If we exclude this part of books, we cannot get a concrete result. Hence, recommendation results are not filtered.

## EXPERIMENT 2: DATA PROCESSING

We found an interesting pattern in view of recommendation results. When recommendations are for arts users, the types of recommended books coincide with professional fields of the users. When recommendations are for science users, the types of recommended books do not coincide with professional fields of the users and the literature books are in the majority. There are some books of other major in recommendation results, i.e arts books are in the majority among recommendation books.

After filtering out irrelevant books with professions, we find that the types of recommended books are basically consistent with the professional knowledge learning sequence. It is very common that some science students who like literature spend free time in reading literary works. These reading behaviors affect the rate of accuracy of the recommendation results. Therefore, our algorithm filters out the irrelevant professional books when doing recommendation to improve the books recommendation accuracy rate.

## VI. FUTURE WORK

Recommendation systems will be used in the e-commerce websites for providing better recommendation of items to be purchased and keeps on updating records. It can also be used in educational institutes for providing improved learning by recording their previous performance. This system can also be used for providing relevant records from huge amount of information overload existing in the world.

## VII. CONCLUSION

We have proposed a novel method for book recommendation. The greatest advantage of this method is that it combines time sequence information and contents features for the accurate results. The accuracy of recommendation results is affected due to our experimental dataset has certain limitation. For example, small copies of books but when appropriate dataset will be used, result accuracy will be improve.

## ACKNOWLEDGMENT

We thank our colleagues who provided insight and expertise that greatly assisted the research. We thank our Professor Sharmila Shinde for inspiring us to do this project and her comments and suggestions.

## REFERENCES

- [1] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowl.-Based Syst.*, vol. 46, pp. 109–132, Jul. 2013.
- [2] Q. Liu, E. Chen, H. Xiong, C. H. Q. Ding, and J. Chen, "Enhancing collaborative filtering by user interest expansion via personalized ranking," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 1, pp. 218–233, Feb. 2012.
- [3] Z. Lu, Z. Dou, J. Lian, X. Xie, and Q. Yang, "Content-based collaborative filtering for news topic recommendation," in *Proc. AAI*, 2015, pp. 217–223.

- [4] A. Hernando, J. Bobadilla, and F. Ortega, “A non negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model,” *Knowl.-Based Syst.*, vol.97, pp.188-202, Apr.2016.
- [5] P. Kouki, S. Fakhraei, J. Foulds, M. Eirinaki, and L. Getoor, “HyPER: A flexible and extensible probabilistic framework for hybrid recommender systems,” in *Proc. 9th ACM Conf. Recommender Syst.*, Sep.2015, pp. 99–106.
- [6] D. K. Bokde, S. Girase, and D. Mukhopadhyay, (2015). “Role of matrix factorization model in collaborative filtering algorithm: A survey.” [Online].Available:<http://arxiv.org/abs/1503.07475>.
- [7] L. Gao and M. Huang, “A collaborative filtering recommendation algorithm with time adjusting based on attribute center of gravity model,” in *Proc. 12th Web Inf. Syst. Appl. Conf. (WISA)*, Sep. 2015, pp. 197–200.
- [8] Y. Huang, B. Cui, W. Zhang, J. Jiang, and Y. Xu, “Tencent Rec: Real-time stream recommendation in practice,” in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, May 2015, pp. 227–238.
- [9] L. Xiong, X. Chen, T. K. Huang, J. Schneider, and J. G. Carbonell, “Temporal collaborative filtering with Bayesian probabilistic tensor factorization,” in *Proc. SIAM Int. Conf. Data Mining*, 2010, pp. 211–222.
- [10] B. Li, X. Zhu, R. Li, C. Zhang, X. Xue, and X. Wu, “Cross-domain collaborative filtering over time,” in *Proc. 22nd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2011, pp. 2292–2298.
- [11] H. Xia, B. Fang, M. Gao, H. Ma, Y. Tang, and J. Wen, “A novel item anomaly detection approach against shilling attacks in collaborative recommendation systems using the dynamic time interval segmentation technique,” *Inf. Sci.*, vol. 306, pp. 150–165, Jun. 2015.
- [12] Y. Wang and W. Shang, “Personalized news recommendation based on consumers’ click behavior,” in *Proc. 12th Int. Conf. Fuzzy Syst. Knowl. Discovery (FSKD)*, Aug. 2015. pp. 634–638.
- [13] S. Feil, M. Kretzer, K. Werder, and A. Maedche, “Using gamification to tackle the cold-start problem in recommender systems,” in *Proc.19th ACM Conf. Comput. Supported Cooperative Work Social Comput. Companion*, Feb. 2016, pp. 253–256.