Journal of Web Development and Web Designing
Volume 3 Issue 2

Analysis of Hybrid Recommendation System for E-commerce Application

Mrs Jayashree Salunke, Ms Anagha Chaudhari
Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India.
Email: jayu.salunke@gmail.com, anagha.chaudhari@gmail.com

Abstract
E-commerce sites are the major developing patterns in the present situation, which encourages online item product selection, purchase and sales. These days E-commerce sites have better popularity and coming nature, so various check of clients wish to share their opinion about their involvement through making reviews, ratings and blogs. Great deals of Recommender System (RS) have taken after the previously mentioned factors for finest item recommendation to the clients. In spite of the fact that, the outcomes are best and reliable, the e-commerce framework should take additional considerations on the related/comparative item analysis. The personalization can't be resolved with just item closeness, this additionally should be recognized by their customize features and interest. So, the Hybrid recommendation system performs effective product recommendation and increases the customer satisfaction.The major ones of these techniques are combining collaborative filtering with sequential pattern analysis, Hybrid model of collaborative filtering, combining knowledge based with user profile and most frequent item technique, combining collaborative filtering with behaviour prediction model, combining content based filtering, collaborative filtering and association rule algorithms. In this paper we explained Hybrid Recommendation System approaches ,which algorithms have highest accuracy, which algorithm solve the cold start problem,gray sheep problem, sparsity problem, Types of Hybrid Recommendation System, Comparison of various types of Hybrid recommendation System & issues of recommendation system.

Keywords: *Association Rule (AR), Behaviour prediction model , ,Collaborative-based filtering (CF) with Sequential pattern analysis (SPA), Content Based Filtering, Knowledge based filtering with user Profile, Recommendation System(RS), Types of hybrid recommendation system.*

INTRODUCTION

In an age of information overload, the importance of personalized recommendation systems for online products and services is rapidly growing. Such systems allow buyers to find what they want without wasting their time and also enable sellers to provide buyers with the items they are likely to purchase, thereby furnishing benefits to both parties. As a result of this growing importance, fundamental knowledge and techniques for developing recommendation systems have been studied, including content-based filtering (CBF), collaborative filtering (CF), association rule or sequential pattern analysis and hybrid approaches [1–3].

A number of studies have attempted to resolve several typical problems of each recommendation technique such as the new user (or cold start) problem, the new item (or the first rater) problem, and the sparsity problem .However, there are still issues for how online shopping malls can make better recommendations for their users [1, 4].

E-commerce sites are the major rising patterns in the present situation, which encourages online item determination, buy and deals. Now a days E-commerce sites have better prominence and coming nature, so various tally of clients wish to impart their insights about their involvement through making reviews, rating and blogs. A lot of Recommender System (RS) have taken after the previously mentioned factors for finest item recommendation to the clients. Despite the fact that, the outcomes are best and reliable, the web based business framework should take additional contemplations on the related/comparative product analysis. The personalization can't be resolved with just item similarity, this likewise should be distinguished by their customized features and interest. In this way, the Hybrid recommendation system algorithms perform effective product recommendation and increase the customer satisfaction.

Most of the current recommendation systems recommend products that have a high probability of being purchased. They employ content-based filtering (CBF), collaborative filtering (CF) and data mining techniques. We analyze the shortcomings of each algorithm and find that the existing algorithms have less direct interaction with customers. The comparative results are shown in Table 1. In order to overcome the problems of each recommendation algorithm and exploit their respective advantages, many scholars studied the combined recommendation algorithms that fuse multiple recommendation techniques. However, most of the research was based on two algorithms: CBF and CF. Few researchers studied the dynamic composition of multiple recommendation techniques with interactive designs [5–8].

Table 1. *Comparative analysis of main recommendation algorithms.*

	Algorithms	Accuracy	Automaticity	Real-Time	Diversity	Scalability	Cold-Start Problem	Sparsity Problem
	CBF	Inferior	Good	Good	Bad	Bad	New users	Not
CF	User-based CF	Better	Bad	Bad	Better	Bad	Serious	Serious
	Item-based CF	Better	Inferior	Inferior	Better	Bad	Serious	Serious
	Association rules	General	Good	Good	Good	General	New projects	General

To overcome these issues, In this paper we explained Hybrid Recommendation System approaches ,which algorithms have highest accuracy, which algorithm solve the cold start problem, gray sheep problem, sparsity problem, Types of Hybrid Recommendation System, Comparison of various types of Hybrid recommendation System.

LITERATURE SURVEY

Table 2. *Literature Survey Table*

Sr No.	Paper Title	Author Name	Publication Details	Contribution
1.	A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis.	Keunho Choi , Donghee Yoo , Gunwoo Kim, Yongmoo Suh	ScienceDirect 2012	Combining implicit rating-based collaborative filtering and sequential pattern analysis
2.	Hybrid Collaborative Filtering Model for improved Recommendation	Hao Ji, Jinfeng Li, Changrui Ren, Miao He	IEEE 2013	Hybrid Collaborative Filtering Model (User Based and Item Based CF)
3.	A New Hybrid Algorithm For Business Intelligence Recommender System	P.Prabhu and N.Anbazhagan	IJNSA 2014	Combines knowledge based, profile of the users and most frequent item mining technique
4.	A Behavior Mining Based Hybrid Recommender System	Zhiyuan Fang, Lingqi Zhang, Kun Chen	IEEE 2016	Combines CF , behavior prediction model
5.	An Interactive Personalized Recommendation System Using the Hybrid Algorithm Model	Yan Guo, Minxi Wang and Xin Li	IEEE 2017	Combines Content Based, CF & Association Rule

In this paper, we have compared 5 papers on the basis of the analysis of Hybrid Recommendation System algorithm's for E-commerce application. The first paper is "A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis." Contribution of this paper is combining Collaborative filtering Algorithm and sequential pattern analysis. The second paper is" Hybrid Collaborative Filtering Model for improved Recommendation ". Contribution of this paper is Hybrid model of User based Collaborative filtering & Item based collaborative filtering. The third paper is" A New Hybrid Algorithm For Business Intelligence Recommender System*"*. Contribution of this paper is combining knowledge based filtering, user profile and most frequent item mining technique. The fourth paper is "A Behavior Mining Based Hybrid Recommender System". Contribution of this paper is combining collaborative filtering and behavior prediction model. The fifth paper is "An Interactive Personalized Recommendation System Using the Hybrid Algorithm Model" Contribution of this paper is combining Collaborative filtering, content based filtering and association rule for E-commerce Application.

ALGORITHMS/METHODOLOGIES

Combining implicit rating-based collaborative filtering and sequential pattern analysis

A recommendation system, called HOPE, which integrates CF-based recommendation using implicit rating and SPA-based recommendation. This section presents the overview of the system, followed by the detailed description of each step of the framework [1].

System overview

Fig. 1 shows an overall framework of our recommendation system, HOPE system, which consists of two main processes: CF process and SPA process. The CF process, depicted in the upper left part of the figure, is the same as the traditional CF process, except that an implicit rating derived from transaction data of users is used instead of explicit rating. Thus, it calculates the similarity between a target user and other users using the implicit rating and selects the top k users based on the similarity score as neighbors of a target user. Finally, the predicted preferences of a target user on items purchased by the top k neighbors (CFPP) are calculated based on the ratings of the neighbors. The SPA process, depicted in the upper right part of Fig. 1, derives sequential patterns from transaction data of other users, and predicted preferences on items (SPAPP) are calculated by matching all subsequences of a target user's purchase sequence data with each derived sequential pattern. Finally, the weighted sum of normalized CFPP and SPAPP is calculated as a final predicted preference (FPP) on each candidate item to recommend, and then the top n items with the highest FPP are recommended [1].

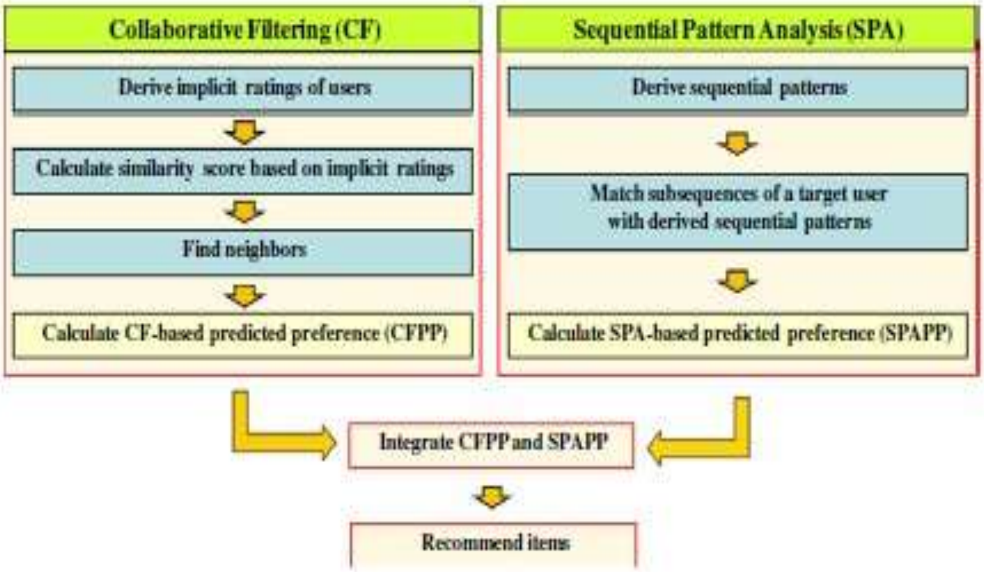

Fig 1. *Overall framework of HOPE system.*

COLLABORATIVE FILTERING-BASED RECOMMENDATION

Deriving implicit ratings of users on items

It is usually difficult to obtain explicit rating information on items. In order to use the CF technique in such circumstance, this paper suggests a method of deriving implicit ratings of users on items from transaction data as an alternative to explicit ratings [1].

Calculating similarity score based on implicit ratings

With the implicit ratings of users on items, similarity between a target user and every other user is calculated, as is done in traditional CF technique, using such similarity function as Pearson correlation coefficient, cosine similarity or distance measures. The Pearson correlation coefficient estimates the similarity based on the rating pattern between two users.

Cosine similarity treats two users as two vectors in the m-dimensional rating vector space, where m denotes the set of all items rated by both users, and estimates the similarity by calculating the cosine value of the angle between the two vectors. Finally, distance measure estimates the similarity between a target user and other user by calculating the absolute magnitude of the similarity between two users in the m-dimensional rating vector space, so that distance-based similarity is defined as an inverse of the distance [1, 9–11].

Finding neighbors

Having calculated the similarity between a target user and every other user using each similarity function, users are sorted by similarity in descending order and then the top k users are selected as neighbors of target user a. We also changed the number of neighbors from 1 to 2 to 3 to 4 to 5 to find the appropriate number of like-minded neighbors.

SEQUENTIAL PATTERN ANALYSIS-BASED RECOMMENDATION

Deriving sequential patterns

In order to calculating predicted preferences of items based on SPA method, sequence data of each user is generated firstly by sorting transaction data for the person according to the transaction date. Sequence data is a series of item sets, ordered by their purchase time stamp. And then, sequential patterns are derived from sequence data of users except a target user using the SPA method.

Matching subsequences of a target user with derived sequential patterns

After deriving sequential patterns, all the subsequences of a target user's sequence data are enumerated. Then, each of these subsequences is matched with each of the derived sequential patterns to find the item(s) to recommend [1].

Integrating CFPP and SPAPP

CFPP and SPAPP are normalized to get N_CFPP and N_SPAPP, respectively, since they are different in the range of values. User a's final predicted preference on item i, FPP(a, I) [1].

RECOMMENDING ITEMS

After obtaining FPP values of items purchased by neighbors of the target user, the top n items are recommended. In this step, unlike usual recommendation systems, items purchased by the target user may be included in recommendation list because users may purchase the same items again [1].

Hybrid Collaborative Filtering Model (User Based and Item Based Collaborative filtering)

Collaborative filtering technique based on users' history in the form of rating given by the user to an item as their information source. It can be accomplished by making relation between the users or between items. Collaborative filtering is categorized into two types: user-based, item-based [12,14].

- ***User-based-***

User-based Approach makes recommendation based on the interest of the user having the similar taste. It correlates user as per the rating given to the items [12,14].

- ***Item-based***

Item-based Approach is based on the items as the user rated items similarly are probably similar [12,14].

Hybrid CF algorithm

Hybrid CF model integrates item-oriented CF algorithm and user-oriented CF algorithm into a unified framework. Hybrid CF algorithm is illustrated in Fig 2.

Both item similarity matrix and user similarity matrix contribute to

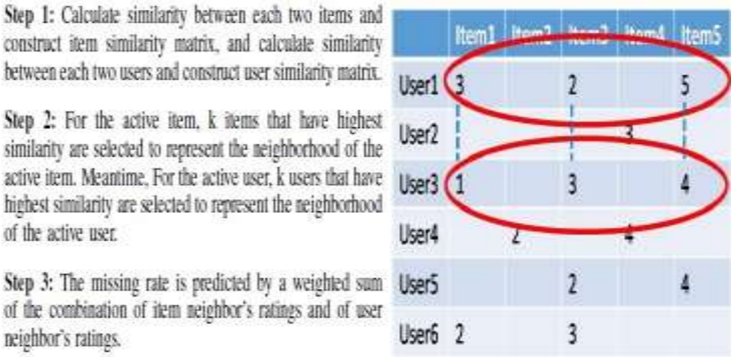

Fig 2. *Hybrid CF Algorithm (Item based + User based)*

the final rating prediction. Two neighbourhood relationships, user-user relationship and item-item relationship, are retained in the same time. This integration of item-oriented CF algorithm and user-oriented CF algorithm not only reward a improved prediction accuracy, but also bring greater robustness of the recommendation system to sparseness problem [2].

Combines Knowledge based, profile of the users and most frequent item mining technique

In this business world, there exists a lot of information. It is necessary to maintain the information for decision making in business environment. The decision making consists of two kinds of data such as OnLine Analytical Processing (OLAP) and OnLine Transactional Processing (OLTP).The former contains historical data about the business from the beginning itself and the later contains only day-to-day transactions on business. The proposed new hybrid algorithm design is shown in Fig 3. Based on these two kinds of data, decision making process can be carried out by means of a new hybrid algorithm based on frequent item sets mining and clustering using k-means algorithm and knowledge of users in order to improve the business intelligence [3].

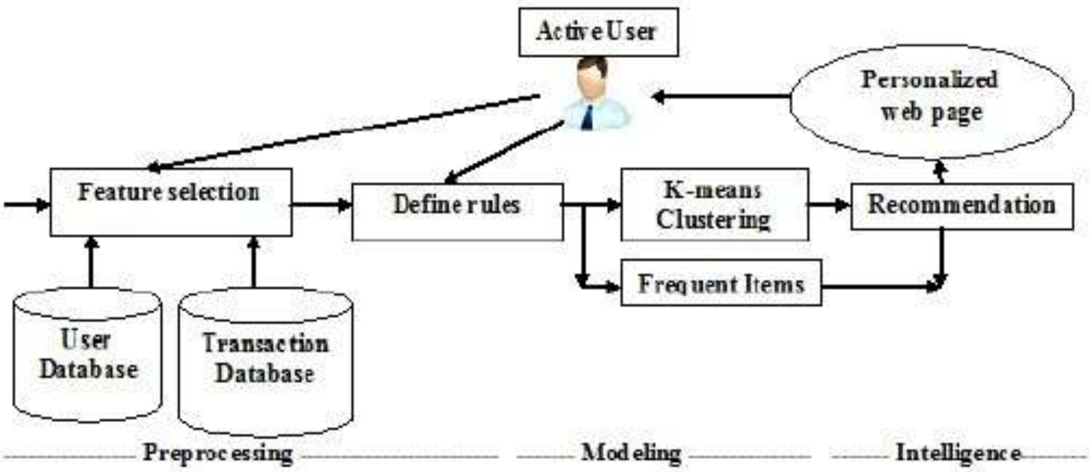

Fig 3. *A New Hybrid algorithm design*

Combines CF, Behavior prediction model & Recommending Phase

Architecture of the Hybrid Model

Fig.4 shows the systems architecture for the hybrid recommending mechanism. The workflow of the system consists of three phases: Behavior Prediction Phase, CF Phase and Recommend Phase [4].

We use the desensitized transaction records provided by T-mall, Alibaba which contains ten thousands users' twenty millions historical purchasing data. We divide the consumer's different behaviors into 4 categories: Click, Collect, Add to cart and Payment. The hybrid model consists of CF (collaborative filtering) and behavior prediction model. The BPM (Behavior pattern model) is in charge of payment behavior prediction work. BPM utilize the Prefix-span algorithm to extract the most prevailing purchasing sequences from the warehouse in real time, and match the sequences with the customer's behaviour pattern that is browsing or adding an item to cart. The real time BPM will return a set of the potential purchasing behavior and the category of the purchasing item. When the recommender system's behavior monitoring part detects the users' potential purchasing tendencies, the system will fetch the user's historical behavior record from sequence database and build an item-user rating matrix [4].

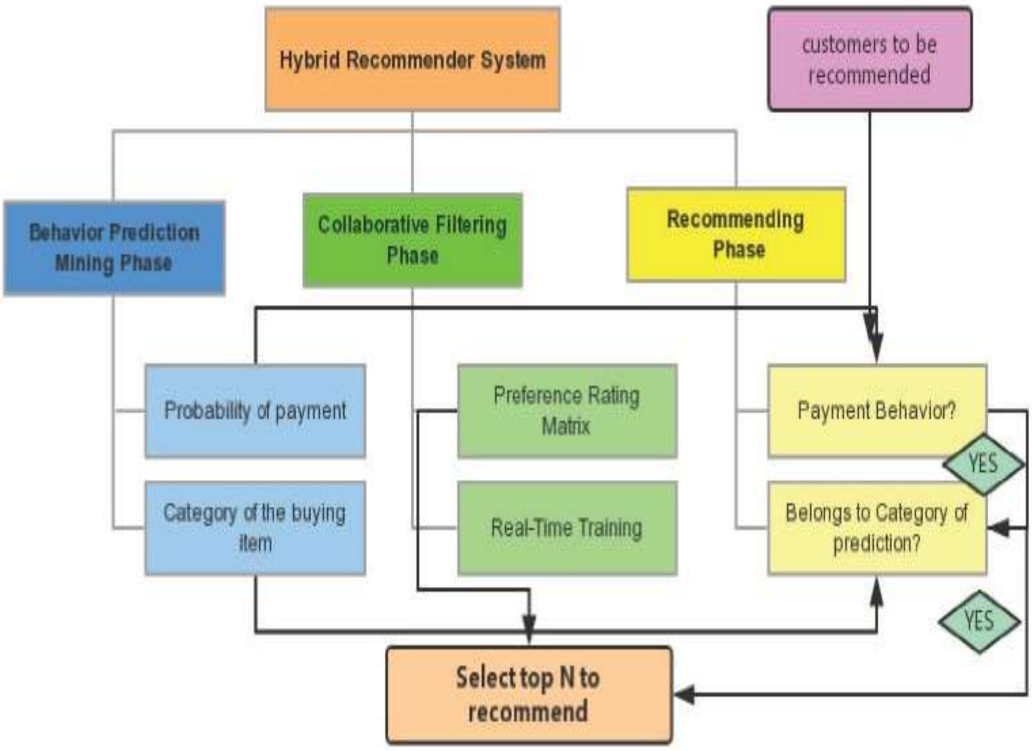

***Fig 4.** The hybrid recommender system's Combines CF, Behavior prediction model & Recommending Phase*

Combines Content Based, CF & Association Rule

To construct an interactive recommendation system based on a hybrid algorithm of multiple recommendation algorithms, we followed the framework as outlined in Fig 5. The first phase is to pre-process the information of products and consumers, where retaining useful information and building relationships are the main tasks. The second phase obtains original recommendation results through various recommendation algorithms. The third phase calculates the weights for each result. A weight measures the importance of a recommended result. The fourth phase is to get the final recommendation results through fusing the original results of various recommendation algorithms. The last phase is to show the recommendation results to the customer through an interactive interface, and record the customer's feedback information to correct the recommendation weights [5].

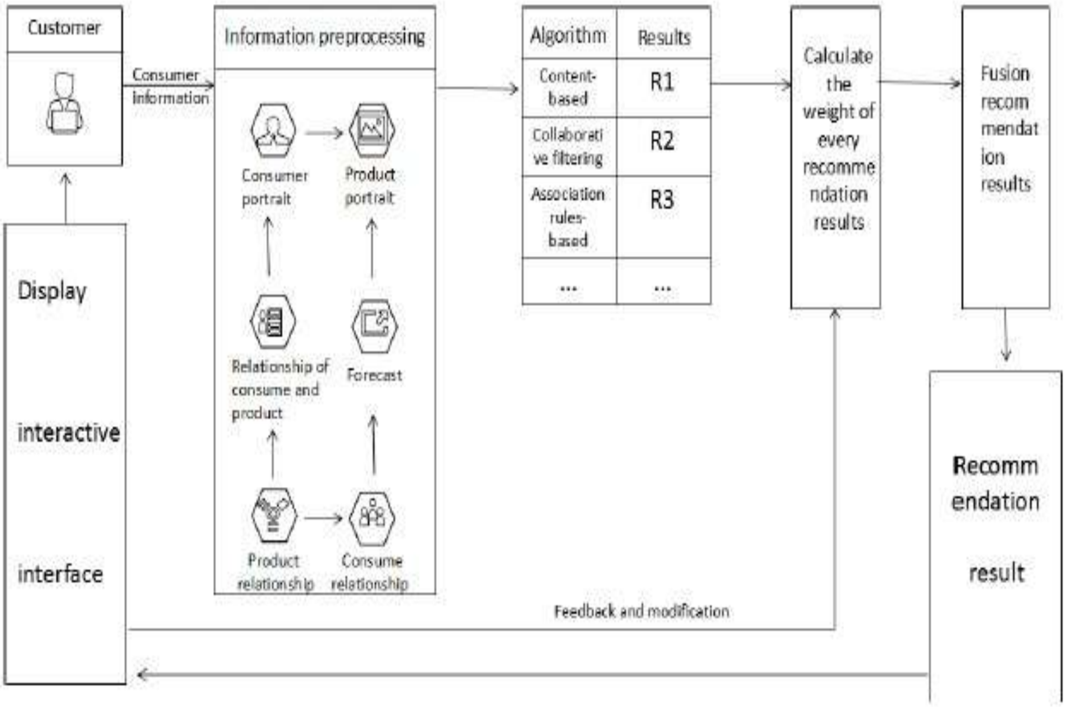

Fig 5. The Combination of CBF, CF, Association Rule recommendation framework.

A new method to enable customers to interact with recommendation systems and to create a feedback loop that incorporates customer feedback and input. Detailed description about the process of interactive interface and procedures with customers is shown in Fig 6. The customer-behavior node refers to the customer's current browsing/search products, click, evaluation and so on, which is used as a basis for calculating personalized recommendations. The information such as time, location and weather is denoted by the context node. The recommendation engine node gets the information from the customer behavior node and the context node to calculate the data for the medium node. The medium node represents data inferred from customer behavior and context data by the recommendation engine: a list of customers that are similar to the active customer is a typical example of such data [5].

Measure the Weights of Each Recommendation Result

After recommendation results of four algorithms are generated, the next step is to determine the weights of the corresponding results. In decision analysis, different results have different importance, and they play different supporting roles in our decision making [5].

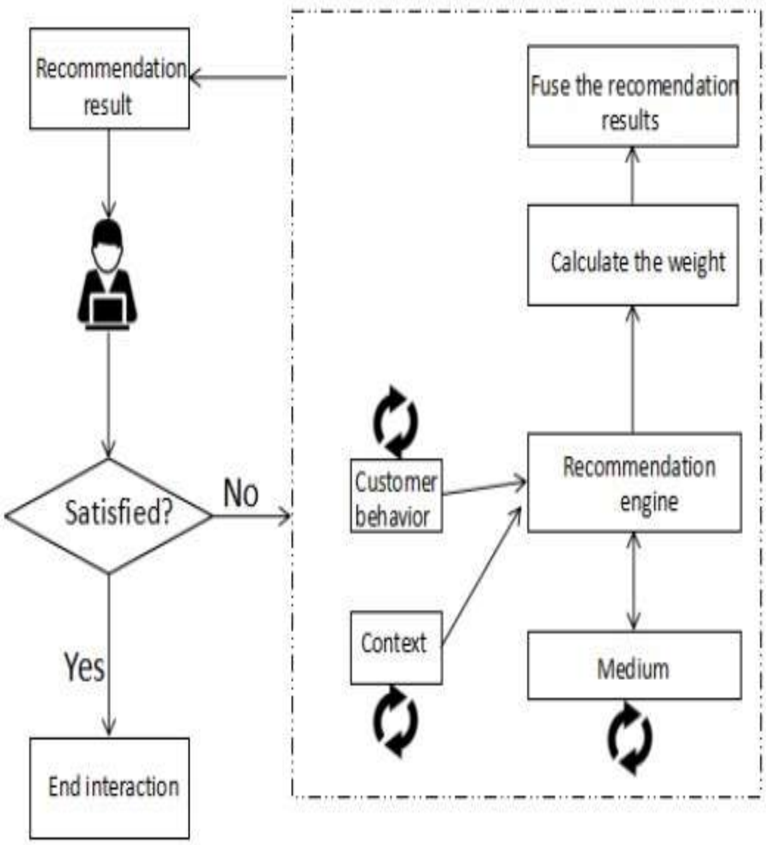

Fig 6. *The framework of the interactive process.*

The weight of the original recommendation result is closely related to the accuracy of the final recommendation result. If the result of a recommendation algorithm has higher accuracy, then the algorithm should be more reliable and have greater weight in the decision process. For recommended data sources I, I =< U, A [P >, the importance of a subset of non-null conditional attributes [5]. The weight of the result is

$$\sigma(SA,P)=\frac{\gamma_A(P)-\gamma_{A-SA}(P)}{\gamma_A(P)}=1-\frac{\gamma_{A-SA}(P)}{\gamma_A(P)},$$

Where *g*A(P) is the classification accuracy of attributes set A relative to class attributes P, and *s*(SA, D) describes the impact of deleting attributes subset SA on the classification of conditional attributes. The greater the impact, the more important the attribute subset. The recommendation capabilities of the recommendation algorithms are different under different situations. Therefore, the weights should be different in the personalized recommendation process. A recommendation algorithm R is given the basis of the distribution of credibility, that is, the rule condition consists of a subset of conditional attributes [5].

APPLICATIONS

Three applications have proven most successful:

Direct product recommendation

The most direct applications of a recommender system to E-commerce is to make direct product recommendations to help individual customers find products they would like to purchase. For instance, Amazon.com (www.amazon.com) has BookMatcher recommender system in which they ask customers to rate a number of books [11,14].

To gift centers

A second application of recommender systems to Ecommercies to gift centers. This application is intended for a customer who visits a site with the intent of purchasing a product as a gift for someone else [11,14]. The challenge in this application is to learn enough about the blessing beneficiary to viably make a suggestion, without requiring excessively exertion from the client. One such framework is utilized as a part of the album now blessing focus (www.cdnow.com)[14].

Cross-sell Recommendation

The final application of recommender systems to Ecommerce at we will discuss is cross-sell. Cross-sell in E-commerce involves recommending items to a customer based on other items that that user has already selected for purchase in the current visit [11,14].

ISSUES IN RECOMMENDATION SYSTEM

Data Collection

The data used by recommendation engines can be categorized into explicit and implicit data. Explicit is all data that user themselves feed into the system. The collection of explicit data must not be intrusive or time consuming [12,14]. Implicit data source in online business is the exchange information including the buy data. Understood information should be broke down first before it can be utilized to depict user highlights or user-item ratings [14].

A. Cold Start

The cold start problem occurs when too little rating data is available in the initial state. The recommendation system then lacks data to produce appropriate recommendations. Two cold start problems are new user problem and new item problem [12,14].

B. Stability vs. Plasticity

The converse of the cold start problem is the stability vs. plasticity problem. When consumers have rated so many items, their preferences in the established user profiles are difficult to change [12,14].

C. Sparsity

In most utilize cases for recommendation systems, because of the index sizes of e-business merchants, the check of evaluations as of now acquired is little identified with the tally of appraisals that should be anticipated. But collaborative filtering techniques focuses on an overlap in ratings across users and have difficulties when the space of ratings is sparse (few users have rated the similar items). Sparsity in the user-item rating matrix degrades the quality of the recommendations [12,14].

D. Performance & Scalability

Performance and scalability are important issues for recommendation systems as e-commerce websites must be able to determine recommendations in real-time and often deal with huge data sets of millions of users and items [12,14].The huge development rates of e-business are influencing the sets much bigger in the client to measurement [14, 15].

E. Gray-sheep problem

Focused specialist organization may give poor appraisals to its aggressive administrations along these lines diminishing its shot of being prescribed. Also, aggressive specialist organization may give great evaluations to its own administrations along these lines expanding its shot of being prescribed. Even some users‟ acting as malicious users‟ might provide inappropriate ratings to products or services. For new users‟, user profile will be initially created with no rating of targeted users‟ by other users‟. This approach is similar for items or services too [13,14].

COMPARATIVE STUDY

***Table 3.** Comparative Table(5 Algorithms)*

Algorithms	Superiority	Robustness	Performance	Decision Making	Quality	Accuracy	Cold Start Problem	Gray Sheep Problem	Sparsity Problem
CF & SPA	Low	Low	High	Low	High	77%	Yes	Yes	Yes
Hybrid CF	High	High	Good	Low	Good	78%	Yes	Yes	Yes
KB, User Profile	Low	Low	High	High	Good	80%	Yes	Yes	Yes
CF & BP	Better	Better	High	Better	High	95%	Yes	Yes	Yes
CBF,CF & AR	High	High	High	Good	High	81%	No	No	No

CBF,CF & AR --Speed Low, Time consuming

Table III shows that comparative study of the 5 Hybrid Recommendation algorithms for E-commerce applications. In this table Collaborative Filtering and Sequential Pattern Analysis, Hybrid CF(User based + Item based),Knowledge based and User Profile, CF & Behaviour Pattern, Content Based Filtering ,CF and Association rule are compared on the basis of Superiority, Robustness, Performance, Decision Making, Quality, Accuracy, Cold Start Problem, Gray Sheep Problem, Sparsity Problem.

From this table we analysed that, From these 5 algorithms the Hybrid approach of CBF,CF & AR provides moderate 81% accuracy with no limitations.And Combination of CBF,CF & AR technique although executes without problems still it faces the limitation of time consumption due to low speed.

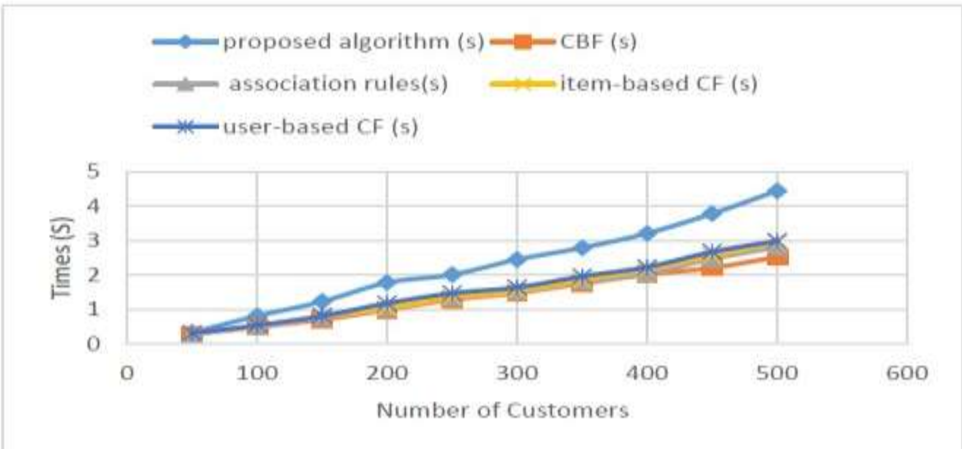

***Fig 7.** A comparison of the recommendation speed according to the number of customers.*

CONCLUSIONS

1. Various Hybrid Recommendation Algorithms are analyzed for E-commerce Application.
2. Various Hybrid Recommendation Algorithms like CF & SPA, Hybrid CF, KB & User Profile, CF & Behavior Prediction, Combination of CBF, CF & Association Rule are compared on various parameters (Superiority, Robustness, Performance, Decision Making, Quality, Accuracy,

Cold Start Problem, Gray Sheep Problem, Sparsity Problem) for E-commerce Application.
3. As per the comparison between 5 algorithms the Hybrid approach of CBF, CF & AR provides moderate 81% accuracy with no limitations.
4. Combination of CBF, CF & AR technique although executes without problems still it faces the limitation of time consumption due to low speed.

REFERENCES

1. Keunho Choi , Donghee Yoo , Gunwoo Kim , Yongmoo Suh,” **A hybrid online-product implicit recommendation system : rating-based collaborative filtering and sequential pattern analysis “,**ScienceDirect,2012.
2. Hao Ji, Jinfeng Li, Changrui Ren, Miao He**, “Hybrid Collaborative Filtering Model for improved Recommendation”,**IEEE,2013.
3. P.Prabhu1 and N.Anbazhagan2,**” A New Hybrid Algorithm For Business Intelligence Recommender System”,** International Journal of Network Security & Its Applications (IJNSA), 2014.
4. Zhiyuan Fang,Lingqi Zhang,Kun Chen,**” A Behavior Mining Based Hybrid Recommender System”,** *IEEE* 2016.
5. Yan Guo, Minxi Wang and Xin Li,” **An Interactive Personalized Recommendation System Using the Hybrid Algorithm Model”,** MDPI/ Symmetry,2017Debashis Das,Laxman Sahoo,Sujoy Datta, **“A Survey on Recommendation System”,***International Journal of Computer Applications*,2017.
6. Yu Lia,b, Liu Lub, Li Xuefeng **“A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce”**,*IEEE*,2005.
7. Ganesan R, Dr. K. Vivekanandan**,” A Novel Hybrid Security Model for E-Commerce Channel”** , International Conference on Advances in Recent Technologies in Communication and Computing,2009.
8. Sahin Renckes, Huseyin Polat, and Yusuf Oysal **“A new hybrid recommendation algorithm with privacy”**, Blackwell Publishing Ltd/Expert System,2010.
9. Babak Sohrabi,Payam Mahmoudian,Iman Raeesi**,” A framework for improving e-commerce websites usability using a hybrid genetic algorithm and neural network system”**, Springer,2011.
10. S. Shruthi, Dr. J. Viji Gripsy **,“An Effective Product Recommendation System for E-Commerce Website Using Hybrid Recommendation Systems ”**,IJCSC,2017.
11. Obert Driskill and John Riedl,**” Recommender Systems for E-Commerce: Challenges and Opportunities”**, *AAAI Technical Report*,1999.
12. Tejal Arekar Mrs. R.S. Sonar Dr. N. J. Uke, **“A Survey on Recommendation System**”, IJIRAE,2015.
13. Karishma Singh, Analp Pathak**, “A Survey on Personalized Recommendation System for Web Services”,** IJSR.2016.
14. Mrs Jayashree B. Salunke , Ms.Anagha N. Chaudhari**, ” Classification of Recommendation System for E-commerce Application”,** *MAT Journals (Computer Science Engineering and Software Testing)*, Dec 2017.
15. Mrs Jayashree B. Salunke , Ms.Anagha N. Chaudhari, **”Implementation of Friendbook: A Recommendation System for Social Networks**”, *MAT Journals (Web Development and Web Designing),* 29 Dec 2017

