Hybrid Movie Recommendation System

Dhrumil D. Shah[1] Dr. Rajesh B. Ingle[2]
[1]Gujarat Technological University PG School, Ahmedabad, Gujarat, India
[2]Pune Institute of Computer Technology, Pune, Maharastra, India

Abstract---Now days, users get so many recommendation websites for movies which suggest/recommends them best movies as per their interest. All these websites have implemented one of the traditional content, context and collaborative recommendations algorithms[3]. Alone, these algorithms are not able to recommend efficient and best recommendations to user. So, there is a need to evolve a new and unique algorithm which combines the features of traditional algorithm along with its new features. This paper describes the new recommendation system, which is a movie recommendation system, based on a unique Hybrid recommendation algorithm, which satisfies a user by providing best and efficient recommendations of movies. Comparative case study of traditional recommendation algorithms to Hybrid movie recommendation algorithm has also been studied and presented in this paper. This case study is based on evaluating criteria of recommendation algorithm i.e. precision, accuracy, recall, F-measure etc. Results of this case study are shown in the form of graphs and tables to clearly specify the need of proposed Hybrid movie recommendation system.

I. Introduction

Recommendation systems are those intelligent algorithms, which can generate results in the form of recommendations to users. They reduce the overhead associated with making best choices among the plenty. Now, Recommender systems can be implemented in any domain from E-commerce to network security in the form of personalized services. They provide benefit to both the consumer and the manufacturer, by suggesting items to consumers, which can't be demanded until the recommendations. Every recommender system comprises of two entities, one is user and other is item. A user can be any customer or consumer of any product or items, who get the suggestions. Input to recommendation algorithm can be a database of user and items and output obliviously will be the recommendations. As in our case, inputs consist of database of customer and database of movies and output denotes the movie recommendations. Generally, Input belongs to recommendation algorithms lies into one of the following categories:

A. Rating Based Input:

It consists of votes of so many people called groups, who rate the particular item on the given scale of minimum to maximum. Collaborative based recommendation systems use this type of input.

B. Content Based Input:

It consists of users' information such as their interest, date of birth, priorities etc. This type of information is hard to find for a particular user, so normally explicitly filled by the user. Content based recommendation systems take this type of inputs.

C. Context Based Input:

This type of input specifies time or behavioral oriented data like date, weather, taste, mood etc. Normally, Context based recommendation lies in this category of input.

II. Data Description

Input to the Hybrid movie recommendation algorithm can be classified into two parts:

A. Movie Database:

The Movie database contains information of various Movies, consists of attributes like Title, ID, Studio, Price, Rating etc. The data layout of Movie database is as follows:

Column name	Content
ID	Unique Identification no. of Movie
Title	Title of the Movie
Studio	Studio Name
Price	Price of Movie DVD
Rating	Movie Rating
Year	Year of Release
Category	Category of Movie
User	No. of users rated the movie

Table. 1: Data Layout of Movie Database

B. User Database:

The User database contains number of user's information, consists of attributes like. The data layout of User database is as follows:

Column name	Content
UID	Unique Identification no. of User
Name	Full Name of User
Address	Full Address of User
DOB	Date of Birth
Email	User's Email ID
PR1	User's Movie Category Priority 1
PR2	User's Movie Category Priority 2
PR3	User's Movie Category Priority 3
MBase	Movie Base Year
VI	User's Visit Counter

Table. 2: Data Layout of User Database

III. Exiting Movie Recommendation System

This study specifies the various conventional algorithms that are still used by some of the most top rated movie purchasing websites. This case study specifies those algorithms along with their flow charts as follows:

A. Content based Movie Recommendation System:

This type of system generates recommendations from source based on the features associated with products[1] and the user's information. Content-based recommenders treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features. So, in case of recommending movies

All rights reserved by www.ijsrd.com

below figure describes the flowchart of content based movie recommendation algorithm

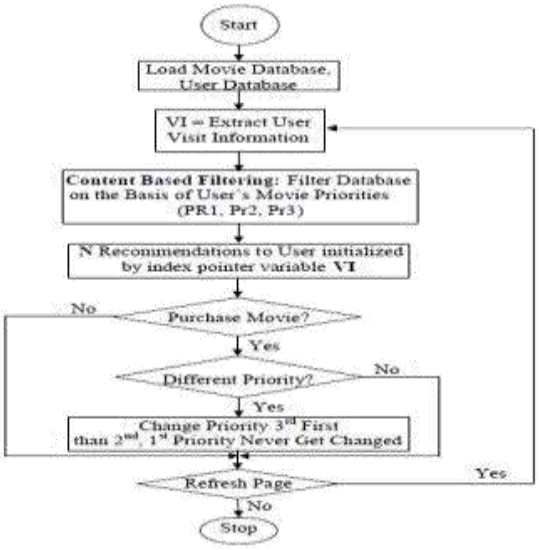

Fig. 1: Flow Chart of Content based Movie Recommendation Algorithm

B. Collaborative based Movie Recommendation System:

In Collaborative recommendation engines, suggestions are generated on the basis of ratings given by group of people[6]. It locates peer users with a rating history similar to the current user and generates recommendations for the user. Most of the movie recommendation engines based on this algorithm, described by the flowchart of collaborative based movie recommendation algorithm:

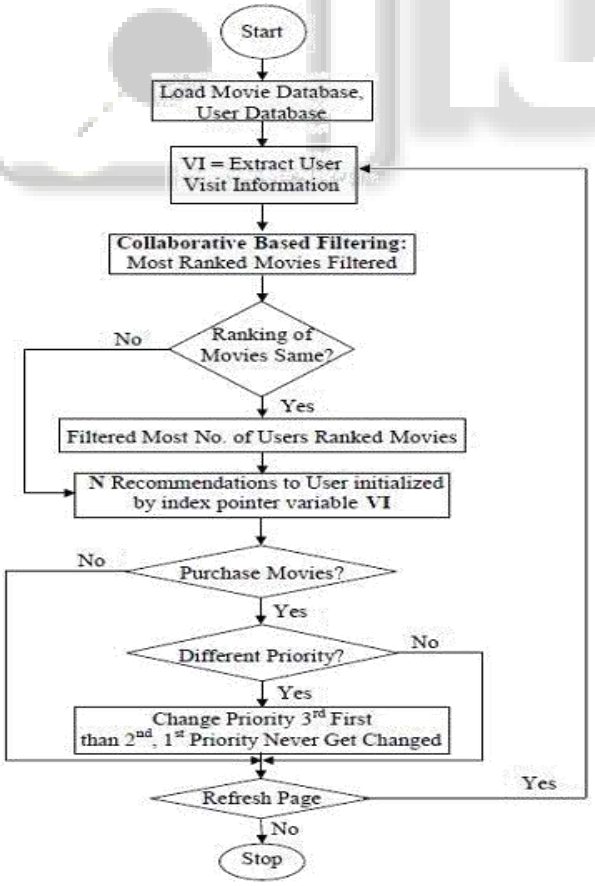

Fig. 2: Flow Chart of Collaborative based Movie Recommendation Algorithm

C. Context based Movie Recommendation System:

This type of recommendation requires the additional data about the context of item consumption like time, mood and behavioral aspects[4]. These data may be used to improve the recommendation compared to what could be performed without this additional source of information. Very rare movie recommendation system based on this algorithm, specified by the flowchart of context based movie recommendation algorithm:

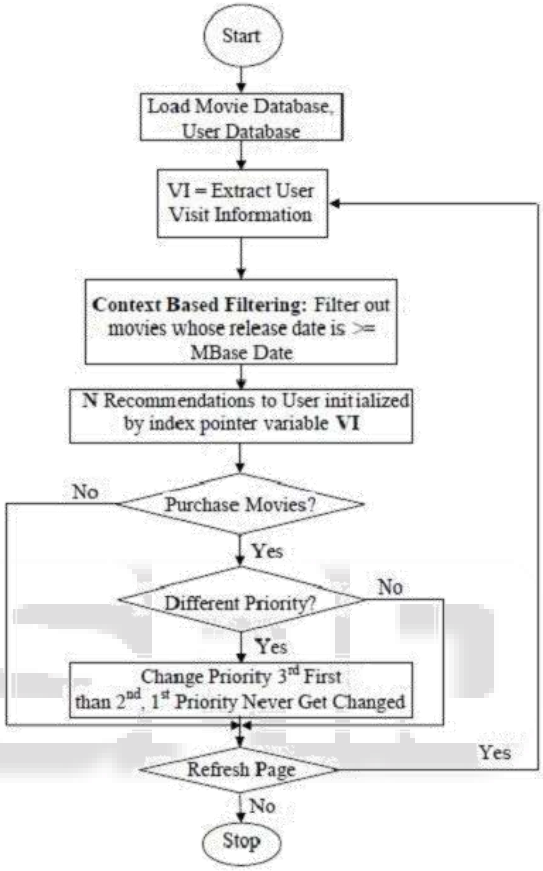

Fig. 3: Flow Chart of Context based Movie Recommendation Algorithm

IV. PROPOSED HYBRID MOVIE RECOMMENDATION SYSTEM

All conventional recommendation algorithms suffer from the limitation of quality, accuracy, precision of recommendations criteria. The proposed system is a Hybrid Movie recommendation system which aims at combining the various features of content, collaborative filtering and context based recommendation system.

It provides an easy to use graphical user interface for user profiles and Movie information management. It generates optimal recommendations for the people that have not sufficient personal experience or competence to evaluate the, potentially overwhelming, number of alternatives offered by a website.

Following figure represents the flow chart and idea of evolution of novel and unique recommendation algorithm:

To overcome the existing Movie recommendation problem Hybrid Recommendation Algorithm has been proposed. It is a fusion of Content, Context, and

All rights reserved by *www.ijsrd.com*

Collaborative Recommendation algorithms. Complete description of proposed algorithm is as follows:

- Input: User Database & Movie Database
- Output: Movie Recommendations
- Steps of Algorithm:

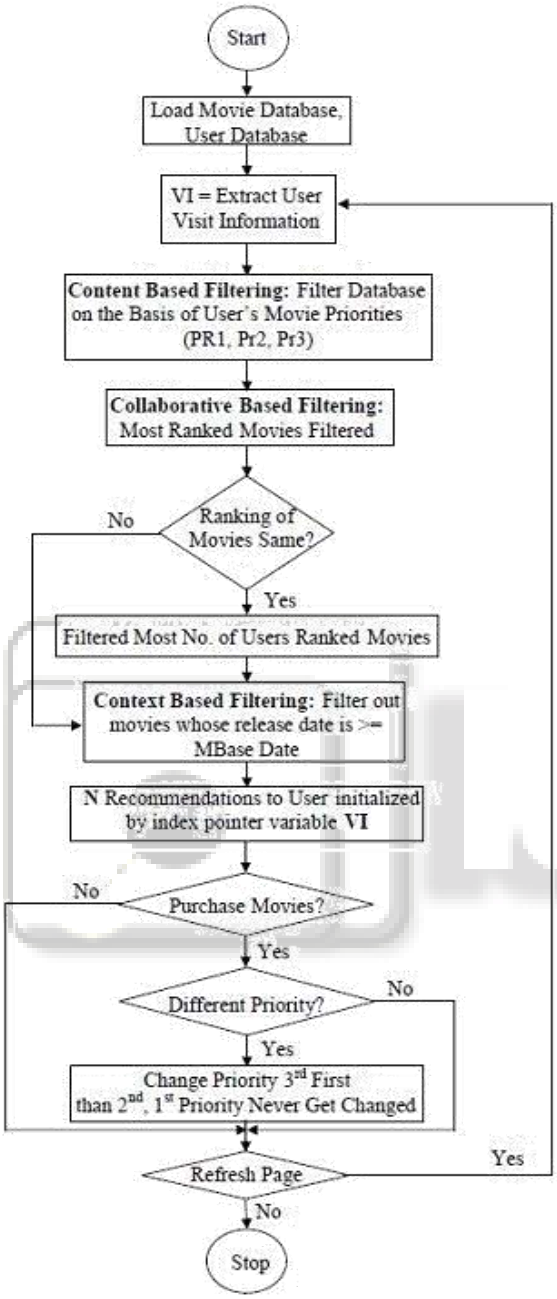

Fig. 4: Flow Chart of Hybrid Movie Recommendation Algorithm

Step. 1 : Initialize variable visit information VI with user's website visit attribute.

Step. 2 : Content Based Filtering: Filter Movie database on the basis of user's movie priorities (PR1, PR2, PR3); that gives the user's priorities wise movies list

Step. 3 : Collaborative Based Filtering: Filter the above database on the basis of most rated movies information from movie rating field

Step. 4 : If two or more movies share the common rating than extract the databases on the basis of most no. of user ranked, otherwise forward to next step

Step. 5 : Context Based Filtering: Now, filter the above database on the basis of movies release date which is >= MBase, gives user specified date wise movies

Step. 6 : Recommend N Movies to user, initialized by index pointer variable VI where N= Number of recommendations

Step. 7 : If user purchase movie, having different priority than already given priorities, then change first 3rd priority and then second priority, 1st priority never get changed. Otherwise forward to next step

Step. 8 : If user refreshes page than increment the VI++ and go to step 3, otherwise VI++ and update the user database

Step. 9 : Stop

V. Comparative Study of Various Movie Recommendation System

To evaluate any recommendation algorithm and their top N-recommendations, there are certain evaluating criteria[2]. A commonly used measure to evaluate performance of recommendation algorithm is accuracy, which defines the fraction of correct recommendations to total possible recommendations as:

$$Accuracy = Correct\ Recommendations / Total\ Possible\ Recommendations$$

To evaluate information retrieval performance measure, best suitable criteria are precision and recall[5] as follows:

$$Precision = Correct\ Recommended\ Items / Total\ Recommended\ Items$$

$$Recall = Correct\ Recommended\ Items / Total\ Useful\ Recommendations$$

To clearly specify the trade-off between precision and recall, a popular single-valued measure is the F-measure. It is defined by the harmonic mean of precision and recall as:

$$F - measure = 2/(1/Precision) + (1/Recall)$$

On the basis of selected evaluation parameters discussed above, following shows the comparative study of Content, Collaborative, and Context movie recommendation algorithm to proposed unique Hybrid movie recommendation algorithm:

Content Based Recommendation Technique				
No. of Inputs	**Accuracy**	**Precision**	**Recall**	**F-Measure**
60	0.6	0.57	0.44	0.49971429
80	0.57	0.54	0.44	0.46875
100	0.53	0.5	0.41	0.42714286
120	0.52	0.48	0.36	0.40615385
140	0.49	0.46	0.34	0.385
160	0.48	0.44	0.32	0.36363636
180	0.44	0.41	0.29	0.33105263
200	0.4	0.37	0.25	0.286
220	0.4	0.36	0.24	0.27428571
250	0.38	0.35	0.23	0.26230769

Table. 3: Evaluation of Content Based Recommendation System on the basis of various Recommendation Parameters

All rights reserved by *www.ijsrd.com*

Context Based Recommendation Technique				
No. of Inputs	**Accuracy**	**Precision**	**Recall**	**F-Measure**
60	0.66	0.61	0.61	0.6052174
80	0.66	0.58	0.53	0.6026279
100	0.63	0.58	0.5	0.5558974
120	0.63	0.57	0.48	0.5156757
140	0.57	0.55	0.48	0.4954286
160	0.57	0.53	0.45	0.4851515
180	0.51	0.5	0.42	0.4546667
200	0.46	0.51	0.38	0.4138462
220	0.47	0.46	0.37	0.4138462
250	0.47	0.47	0.36	0.3933333

Table. 4: Evaluation of Context Based Recommendation System on the basis of various Recommendation Parameters

Collaborative Based Recommendation Technique				
No. of Inputs	**Accuracy**	**Precision**	**Recall**	**F-Measure**
60	0.79	0.82	0.81	0.8340741
80	0.8	0.81	0.78	0.804031
100	0.81	0.8	0.75	0.7639669
120	0.78	0.77	0.75	0.7439316
140	0.75	0.74	0.68	0.7238938
160	0.76	0.73	0.67	0.7038532
180	0.74	0.7	0.64	0.6737864
200	0.68	0.66	0.6	0.6336842
220	0.68	0.65	0.6	0.6336842
250	0.66	0.65	0.59	0.6136264

Table. 5: Evaluation of Collaborative based recommendation System on the basis of various Recommendation Parameters

Hybrid Recommendation Technique				
No. of Inputs	**Accuracy**	**Precision**	**Recall**	**F-Measure**
60	0.92	0.87	0.86	0.8747239
80	0.89	0.86	0.83	0.8447134
100	0.85	0.81	0.83	0.804698
120	0.83	0.8	0.79	0.7846897
140	0.81	0.78	0.75	0.7646809
160	0.79	0.76	0.75	0.7446715
180	0.76	0.71	0.7	0.7146565
200	0.71	0.69	0.67	0.6746341
220	0.71	0.68	0.68	0.6646281
250	0.7	0.68	0.68	0.6546218

Table. 6: Evaluation of Hybrid Recommendation System on the basis of various Recommendation Parameters

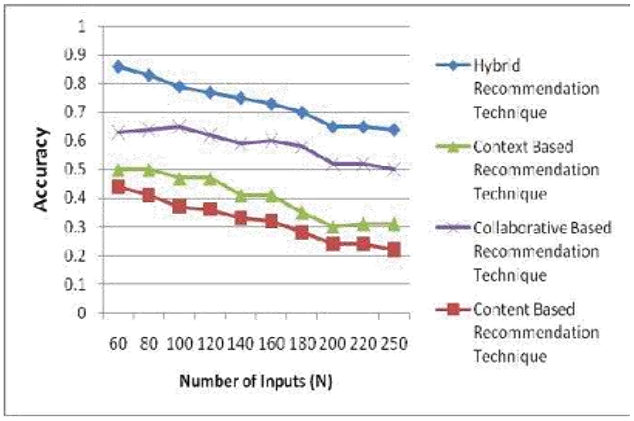

Fig. 5: Comparative analysis of various recommendation techniques on the basis of Accuracy

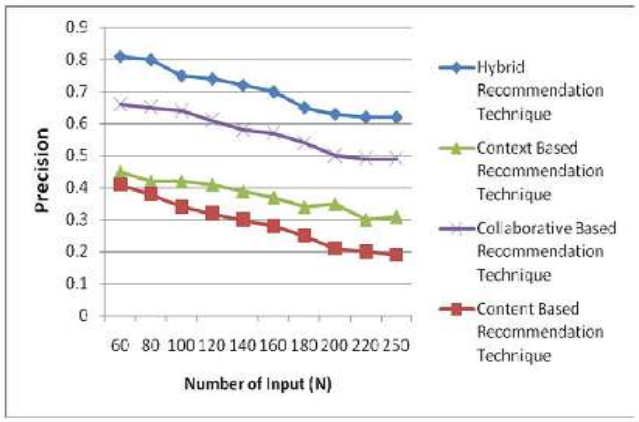

Fig. 6: Comparative analysis of various recommendation techniques on the basis of Precision

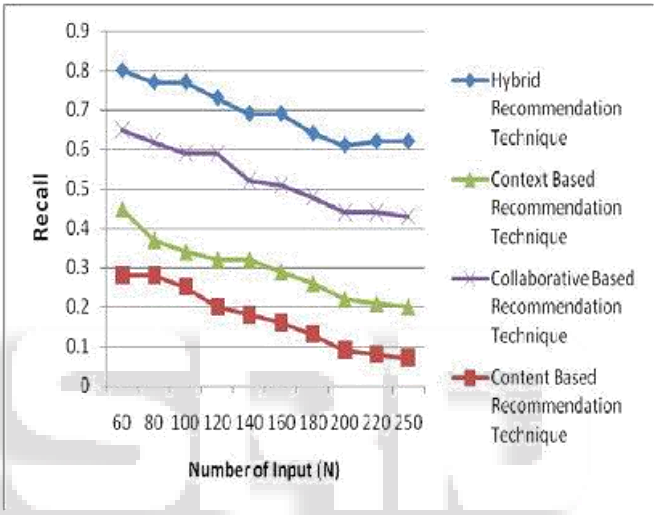

Fig. 7: Comparative analysis of various recommendation techniques on the basis of Recall

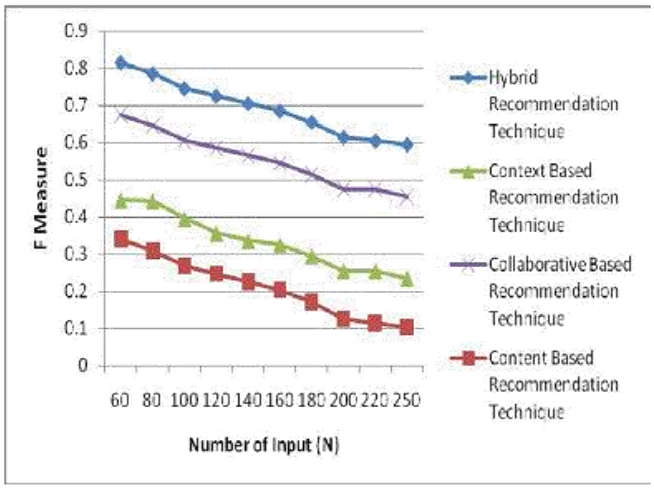

Fig. 8: Comparative analysis of various recommendation techniques on the basis of F-measure

On the basis of calculations and results concluded on all of four recommendations algorithms and on the basis of recommendation evaluation parameters like accuracy, precision, recall and F-measure, it can be concluded that the proposed Hybrid movie recommendation algorithm is best among the others.

All rights reserved by *www.ijsrd.com*

VI. CONCLUSION AND FUTURE RESEARCH WORK

We have covered the core of recommendation experiments, and have examined how to incorporate information from metadata into recommendation algorithms. In addition, we evaluates the possibility to combine the various recommendations techniques and incorporated with proposed hybrid algorithms to improve recommendation performance by exploiting the complementarities of different algorithms. It describes the traditional Content, Collaborative and Context Filtering recommendation approaches along with their precision, recall and accuracy parameters.

On the bases of this study, Hybrid algorithm approach has been proposed in order to improve the basic algorithm. At this point, it is essential to include a categorization of the previously described algorithms based on different criteria. Finally, this we have presented a number of utilized evaluation metrics, from which some were used to measure quality, while others to measure performance.

Recommendation systems are an extremely strong tool utilized to assist the selection process easier for users. This analysis authenticated that the hybrid recommendation algorithm is able to recommend Movies for web-users, whereas the other recommendation algorithms are quite slow with inaccuracies. This recommendation system will assuredly be a great web application, which can be club with today's high demanding online movie purchasing web sites.

REFERENCES

[1] G. Linden, B. Smith, and J. York, "Amazon recommendations: Item-to-item collaborative filtering," IEEE Internet Comput., Feb. 2003.
[2] Michael Hashler, "Recommender Lab: A Framework for Developing and Testing Recommendation Algorithms" Nov. 2011.
[3] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Trans. Knowl. Data Eng.
[4] C. N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, "Improving recommendation lists through topic diversification", WWW '05: Proceedings of the 14th international conference on World Wide Web, New York, NY, USA, 2005, ACM.
[5] O. Celma and P. Herrera, "A new approach to evaluating novel recommendations", RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems, New York, NY, USA, 2008, ACM.
[6] R. Bell, Y. Koren, and C. Volinsky, "Modeling relationships at multiple scales to improve accuracy of large recommender systems" KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, New York, NY, USA, 2007, ACM.

All rights reserved by *www.ijsrd.com*