ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Collaborative Filteration for user behavior Prediction using Big Data Hadoop

Jayshri Manohar Somwanshi, Prof. Y. B. Gurav

Department of Computer Engineering, P.V.P.I.T Bavdhan, Pune, Maharashtra, India

Department of Computer Engineering, P.V.P.I.T Bavdhan, Pune, Maharashtra, India

ABSTRACT: The large number of online video-sharing systems, You-Tube is the most popular web site, provides features that allow internet users to post large number of video on social networking site and video-sharing website as a response. The mobile user is very difficult to find out their favorite video. The proposed work define collaborative filteration technique to help the mobile users can easily find out their favorite videos. In this paper , video-based recommendation system using big data of hadoop which can reduce the network and fast of collaborative filteration system. The mobile user can easily find out their favorite videos. The recommendation system mainly described on content-based filteration, collaborative-based and context-aware filteration. The collaborative filteration used as user-based recommendation and item to item based recommendation system. The result shows that the proposed approach collaborative filteration desired services with similarity and ratings.

KEYWORDS: collaborative filtering; Mobile user behavior analysis; Multimedia service recommendation; hadoop; Map-Reduce

I. INTRODUCTION

Recommendation system all forms of like that TV, music, books, video on user demand, online Internet video.The internet user post the large number of video clips on the social networking website like that facebook and video-sharing websites. The large number of videos are available on online. The number of video content are copied, little bit different or similar. The online users are suffering very difficult to find out their favorites videos because to obtain large number of video content list from the amount of web pages in short time. This Situation is very hard for the mobile users because of the screen size and its bandwidth, so mobile users very hard to find out their favorite videos. In the recommendation system video-sharing websites filtered their favorite videos content list, their watching video history, video description tags and video classification. The user wants to search their desired favorite video quickly for search engine to improve this situation for the recommendation system. The user search their favorite video on search engine, searching is based on the keywords. Mostly mobile users dont have any keywords, so mobile users do not search their desired video instantly. The recommendation system are very usefull for the mobile multimedia users application. The main aim of the recommendation system is to use the high quality and top rated videos to the user. When the user looking for information about any movie, music, video, the internet users come across a number of options to fetch previous data from the recommended video list. YouTube developed in early 2005 has been one of the successful user-generated video sharing and video recommendation website, which become a most attractive and a most popular destination for all kind of users to find most videos online. Mostly online users to use to see their favorite videos and download their mostly popular or highly rated videos. The google search engine used the content-based filteration (CB), because content based filteration system makes on the similarities of content rating, titles, description and tags. In video recommendation system, the datasets are crawled from YouTube, a well-liked online video-sharing community to suggest videos with high rating along with minimum rated large videos [3].The user gives the rating of their favorite video then next time they search video they can easily find their favorite video by using rating because they gave the rating of this video then recommendation automatically recommend that video of this user. Nowadays, allover every person has use mobile device such as a smart phone.With the rapid growth of the technology, there are large number of communicating applications are available in mobile device [2]. For example in mobile device can be equipped with lots of application like Wi-Fi, Bluetooth, GPRS, mobile data. So mobile users use their device as little computer to using the internet, as watching video, listen to music [2]. In filteration system content-based filteration

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

(CB) systems determine item descriptions to user identify their favorite items of their interest to the user. In content-based recommendation system ,available a summary list of item and description of item and user choose more information on an item or to communicate with the item [4]. The item to item recommendation find out the similarity of video ratings.

II. RELATED WORK

Recommendation systems have changed the way people find their favorite video, books, products, music, information, and even other people. The mobile user easily find their favorite video ,product and other information by using recommendation system. The filteation system based on the users choice have been used to user favorite choice of filteration. In recommendation services for several of online readers. YouTube uses the recommmender system to help user find their favorite video or watching history to previous [1]. The recommendation system mainly used for the content base filteration, Collaborative filteration, context-aware filteration and graph-based filteration.

A. CB Filteration:-

In content based filteration system, the learning model of users interest with content-based filteration, the content-based filteration system can give very good recommendations if the content does not contain enough information to distinguish items the user favorite from items the user does not like[4]. Content-based filteration systems filtered an item to a user depends upon a explanation of the item and a profile of the users interests. Then a user profile are easily entered by the user, then it is learned from previous the user provides on items. The content based filteration are easily developed. A large number of learning algorithms have been adapted to learning user profiles, and the favorite of learning algorithm depends upon the representation of content.

B. CF Filteration:-

In collaborative filteration, system makes filteration depends on abundant user previous histories and content choice. CF filteration system requires number of records and previous feedback. Otherwise user prediction feedback and user own opinion description method should be filteration [1]. In collaborative filtering, real-time and location recommendation . The main aim of collaborative filterating solution is to develop the diversity of filteration. The key challenge is to recommend candidate in real-time, with minimum data.The CF recommendation system is important for filtering items to users depends on various similarity content. The collaborative filtering mainly focused on user based recommendation and item based similarity. Forever the social network application like video, with a large number of video post by users.to slove this problem by collaborative filteration, in determine user rating and item to item based similarity such as MapReduce [5]. The CF recommendation system can decrese popularity by using item to tem based recommendation of algorithm in the filteration.

C. Context-aware Filteration:-

In context-aware recommendation, system make easily provide constant filteration do not consider user context information.However user interest differ according to real-time, location-aware and their emotions. context-aware recommendation is very hard concept to briefly explain;semantic context-aware are used to enrich the explanation of context. Quality of context-aware are considered user not satishfied factory for the start of research in context-aware systems[6]. Initially context-aware systems have also needed to collect and model additional data with context. Quality of context defined as QoC is any inherent information that explain context information and can be used to find out the worth of the information for a specific domain. QoC shows the accurate quality of context information that depicts that context-aware information is free of error and explain the recent situation in the environment.

D. Graph-based Filteration:-

Graph-based filteration is used in the recommendation system to calculate the similarity and correlation between user video rating [1]. The user based recommendation and item to item based recommendation ,links on video rating pages are converted to undirected weighted graph. With the large number of user id different categories of videos, user gives the different video categories of different rating, graph based filteration system require more and more computation capacity[1]. To resolve the large computation requirement, user based and item to item based recommendation algorithms have been implemented on hadoop platforms to improve performance and scalability of the filteration system.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

III. PROPOSED SYSTEM

A. *Proposed Work:*

The proposed system work for the users are easily search their favorite video by using collaborative filteration. The user easily analse the video by browsing the data. Then the user-based recommendation system recommend user favorite video by video average rating. The item to item based recommendation system recommend the user similarity of video ratings. We discuss the system in remaining sections.

B. System Architecture:

As shown in the below fig. the system will work as follows:

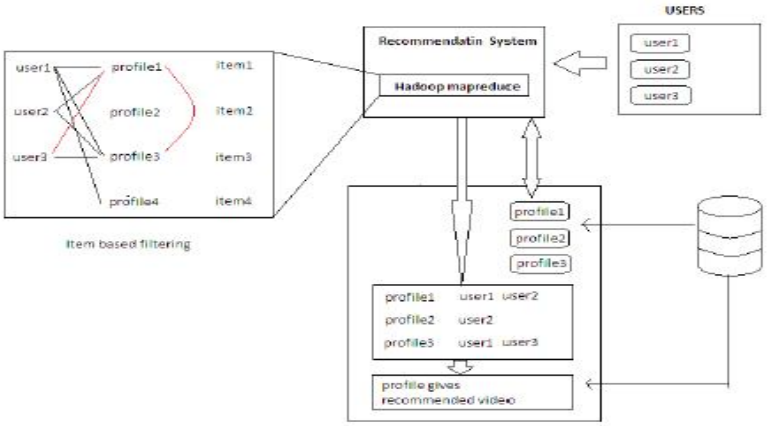

Fig. 1. System Architecture

The system architecture as shown in fig. the system architecture divided in to two parts user based recommendation and item to item based recommendation.

In user based recommendation, user analyse the video and number of user gives the different rating of different categories of videos. The videos displayed in yearly. The user gives the high rating of their favorite video then next time that video easily search. The user based recommendation ,user does not get recommend their own video. The recommendation is based on calculating video average.

In item based recommendation system, user find out the similarity of user gives the different rating of different categories of videos. Similarity calculate between items are more likely to converge over time than similarities between users. Item based recommender begin with a list of user's preferred item and therefore do not need a nearest item neighborhood as user based recommender do. In item based recommendation one of the video does not have any rating this video automatically deleted.

C. Algorithm:

(1) System (user_id, video_id, video_rating, time)

(2) Initialize Req=[ ] //Registeration

R = {U, Pw}

U = {u1, u2,……un}

Pw = Password

(3) Analyse = video count + yearly and video type.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

(4) add *U* to relevant rating

(5) *avrg = video_id + video_rating*

(6) *item = (item_id, item_count, sum,rating, key, value)*
(7) *item_count =0*
(8) *item_sum = 0*
(9) *Final = []*
(10) *For item_id, rating in value*
(11) *Item_count + = 1*
(12) *Item_sum + = rating*
(13) *Key = user_id*
(14) *Value = item_count, item_sum, final*
(15) *Emit (key, value)*
(16) sum_xx, sum_xy, sum_yy, sum_x, sum_y, n = (0.0, 0.0, 0.0, 0.0, 0.0,0)
(17) sum_x + = item_x
(18) n+ = 1
(19) item_pair, co_ratings = pair_key, lines
(20) item_xname, item_yname = item_pair
(21) item_x, item_y, in lines
(22) sum_xx+ = item_x * item_x
(23) sum_yy+ = item_y * item_y
(24) sum_xy + = item_x * item_y
(25) sum_y + = item_y
(26) similarity = normalized_coorelation(n, sum_xy, sum_x, sum_y, sum_xx, sum_yy)
(27) key = (item_xname, item_yname)
(28) value = (similarity)
(29) key = (item_id, item_id)
(30) value = (similarity, n)
(31) Emit = (key, value)

IV. EXPERIMENTAL RESULT

(a) Registration Process:-
(1) The first starts the registration process it gives the user name and password then it sign in the system and also in signup we can create a new user and clear buttons are present as shown in fig. 2.

(b) Analyse Data:-
(1) Then it browse the data from usr/local/video/video.dat and click the show dataset and click the Copy to HDFS. Then click on analyse button, go to localhost:70 input 1 and output1 file generated and also go to localhost:30 video analysis completed. Click on show o/p yearly, all video shows in yearly and also video display on video types as shown in fig 3.

(c) Recommendation System:-
(1) In recommendation, firstly browsing rating from usr/local/video/rating.dat and click on Copy To HDFS then all data copy to HDFS.
(2) In user based recommendation, click on Run button then go to localhost:70 Input 2 and Output 1 And Output 2 file generated and also go to localhost:30 userbased recommendation completed. Then click on recommend button user rating average calculated as shown in fig. 4.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

(3) In item based recommendation, click on Run button then go to localhost:70 Output 3 and Output 4 file generated and also go to localhost:30 item based one and item based two recommendation completed. Then click on recommend button item to item similarity calculated as shown in fig. 5.

(d) Rate Video:-

(1) First enter user id then enter video id and then gives the rating and then submit and finally video rating completed as shown in fig. 6.

Fig. 2. Registration process

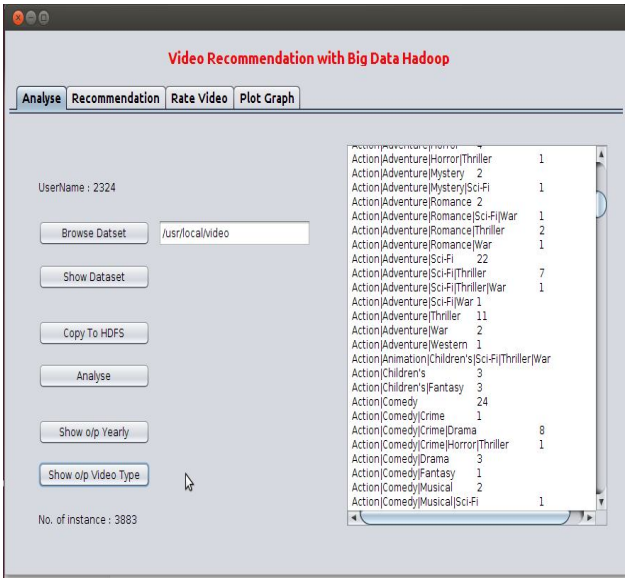

Fig. 3. Analyse the video.

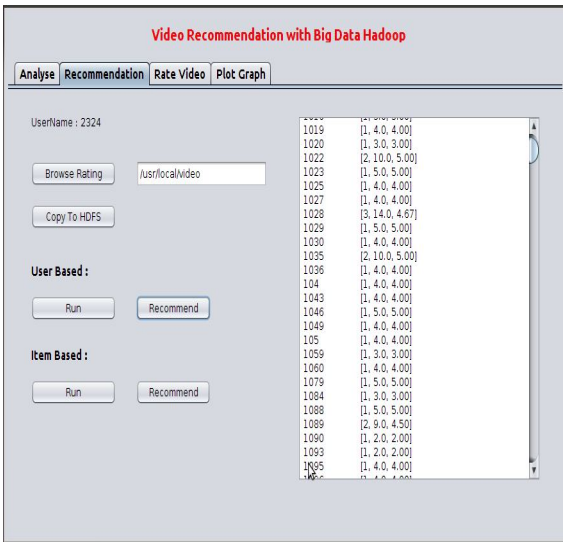

Fig. 4. User based recommendation.

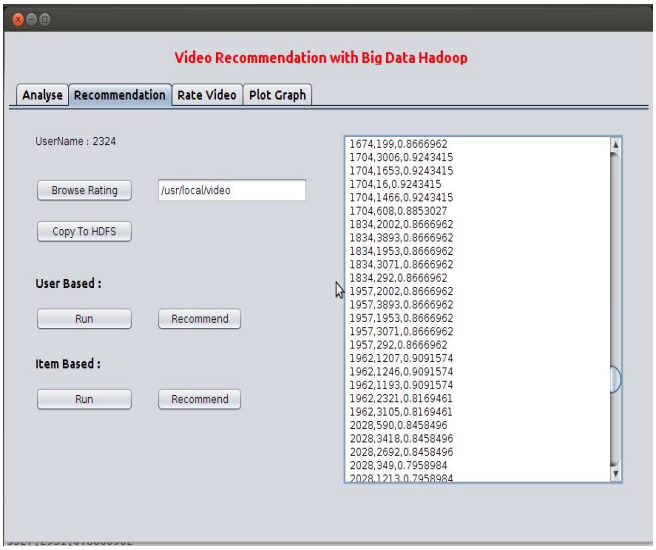

Fig. 5. Item based recommendation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

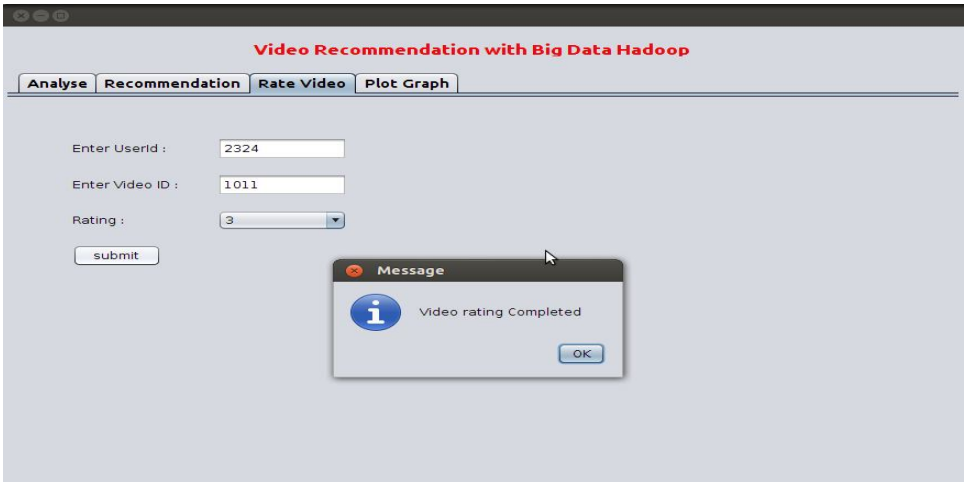

Fig. 6. Video rating

V. CONCLUSION

Recommendation systems are a very powerful technology for personalization. Thus the users are easily recommending the video by using collaborative filteration for user behavior prediction using big data hadoop.The collaborative filterating proves to be the well recommended video to the user, valid recommendation and most important time efficiency.

REFERENCES

[1] Yijun Mo, Jianwen Chen, Xia Xie, Changqing Luo, and Laurence Tianruo Yang, Cloud-Based Mobile Multimedia Recommendation System With User Behavior Information,IEEE System Journal & Vol.8, 1,March 2014.
[2] K.-D. Chang, C.-Y. Chen, J.-L. Chen, and H.-C. Chao, Challenges to next generation services in IP multimedia subsystem, J. Inf. Process. Syst & Vol 6, no. 2 , pp. 129-146, Jun. 2010.
[3] X.Wu, Y. Zhang, J. Guo, and J. Li, Web video recommendation and long tail discovering, Proc. IEEE ICME, 2008, pp. 369-372.
[4] M. J. Pazzani and D. Billsus, Content-based recommendation systems,in TheAdaptiveWeb. Berlin,Germany: Springer - Verlag, 2007, pp. 325-341.
[5] Z.-D. Zhao and M.-S. Shang, User - based collaborative-filtering recommendation algorithms on Hadoop,in Proc. WKDD, 2010, pp. 478-481.
[6] P. Pawar and A. Tokmakoff, Ontology -based context-aware service discovery for pervasive environments ,, in Proc. IEEE Int. Workshop Service Integr. Pervasive Environ ., Jun. 2006, pp. 1-7.
[7] D. Li, Q. Lv, X. Xie, L. Shang, H. Xia, T. Lu, and N. Gu, Interest based real - time content recommendation in online social communities ,Proc. Knowl.-Based Syst., vol. 28, pp. 1-12 Apr. 2012.

BIOGRAPHY

JAYSHRI MANOHAR SOMWANSHI student of ME Computer Engineering second year from the college TSSM's Padmabhushan Vasantdada Institute of Technology, Bavdhan, Pune.