International Conference on Innovative Research in Engineering, Management and Sciences
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 4 | Issue 9 | ISSN : 2456-3307

Association Rule Based Recommendation System Using Mapreduce

Mrs. Divya K V [*1]

[*1] Assistant Professor, Department of Information Science and Engineering New Horizon College of Engineering Bangalore, India
divya.k.vasudevan@gmail.com[1]

ABSTRACT

Recommender systems are integral part of any ecommerce store in order to withstand and compete with other growing businesses. There are various recommendation techniques which are used to appropriately recommend a product to the active user. The recommendation system has to analyze large amount of data to provide better recommendation and such important issue can be addressed using Hadoop ecosystem. In this paper, a recommendation system for product based on Hadoop framework is proposed. The proposed system recommend products to the user depending on the products present in the user cart. First, it uses framework to import the product transactions. Furthermore, the Apriori for finding frequent itemsets and Association rules are implemented in Hadoop and processing the data using MapReduce.

Keywords : Recommendation system, Apriori, Frequent itemsets, Association rule.

I. INTRODUCTION

The effectiveness of a product recommendation system for a particular active user is very significant in today's world. Two types of recommendations can be provided: personalized recommendations and non-personalized recommendations. Examples of personalized recommendations are the 'People you may know' section based on mutual friends in Facebook or the 'Recommended videos' section based on the previous browsing history in Youtube. Examples of non-personalized recommendations are the hotel recommendations, movie recommendations which are based on other similar users' interest. Recommendation systems uses the features of information retrieval and using data mining techniques. These recommendation carry out in three steps: data pre-processing, data analysis and result interpretation.[1] Recommendation systems are portion of information filtering system which recommends information items, social objects based on the user's interest. These systems have changed the technique the user look for information and product. [2] It is a matter of custom-made, interesting and usefulness that separates the recommendation systems from information filtering system or search engines. [3]

Recommendation systems are stated as scheme which recommends the products to particular user by considering the user's interest and recommending items on prediction based on items, users and item user interaction. The main aim of recommendation systems is to lower the information overload issue by providing user needed information from huge amount of data. [4] From the definition of recommendation system it says the basic unit of each and every recommendation system which takes user

preferences and by evaluating it helps in predicting the products to the user. [5]

One major problem with recommendation systems is to combine different recommendation systems in order to attain high performance. Each and every recommendation systems have advantages and disadvantages in which disadvantages can be reduced by merging features of different systems [3].

This paper presents an overview of some of the recommendation system techniques which can be used in various applications for the purpose of appropriate recommendations. Further explained the proposed system aims at implementing a product recommendation system, which appropriately recommends products to the active user. When a user logs in through an ecommerce web portal and begins shopping, depending on the products present in the cart, the system analyses and recommends the products accordingly. To make our system scalable, we use the Apriori algorithm implemented as a MapReduce program for processing the Big Data on HDFS. We also show that the proposed recommendation system solves the coldstart problem, which is a drawback of all the existing recommendation systems mentioned above. Unlike the medium to large retailers, small retailers have certain limitations like small data pools and limited computing resources. Hence, the algorithms used to process the continuously growing data should support scalability in large amounts. The proposed system, aims to use the Hadoop environment and the MapReduce concepts.

II. RELATED WORK

There are many recommendation algorithms that are commonly used. For example content-based recommendation, collaborative-filtering recommendation, association rule- based recommendation, utility-based recommendation and knowledge-based recommendation.

The Product recommendation system requires the interaction of the retailer, all of the retailer's past customer base and the active customer who has currently made a selection for his or her shopping cart. [1] The retailer wishes to accurately predict which product recommendation will be most likely to result in the active customer making an extra purchase during this current transaction. In order for a small retailer to implement a product recommendation system such system must be efficient when running on a server machine with modest computing capability, as small businesses normally do not have the financial capacity to invest in a large infrastructure.

Location aware Recommendation system user are associated with a home city, and alert friends when visiting a location by checking in the application. during check in user can also write reviews which are free text notes describing what they liked about the location. The user reviews about items in a location have been utilized to model the data .The spatial user rating for spatial items are extracted such as user id, user location, rating, item, item-location, reviews. Each user visit is mapped to a single location based rating.[5]

Location Based Context Aware Recommender System dealing with a data mining based approach named as Preference oriented location based search (POLS) to efficiently search k nearby stores that are preferred by the user based on the user's location, query time and preference. In POLS, two preference learning algorithms are proposed that automatically learns user preference and ranking function is also proposed to rank the nearby stores based on the user's location, query time and preference knowledge-based recommendation recommend items based on specific domain knowledge about how certain item features meet the user needs and preferences and utility, how the item is useful for the user. Knowledge based recommender system are case based. In this, similarity function estimates how much the user needs match the recommendation.

This architecture consists of a user interface model a interface engine, acknowledge base of the product domain, a customer database. The user interface module interacts with users. It asks users about what features of the product they need and collects the answes from the users.to the user ,recommendation system appear as a guided interrogation supported by the image that illustrate the answer alternatives.

Hybrid Product Recommender System propose a recommender system based on RFID technology for VIP customers in apparel retailing store.[2] The application of the recommender system will improve the quality of customer service via automatically generating of recommender list from the view of customers' interests. The contents in the list may guide the shop assistant to know what kind of products may interest the customer, and the assistant could introduce the corresponding products to the customer based on the recommender list. In apparel retailing stores, new shopping experience provided to customers such as automatic product recommender could improve the brand loyalty.

Content-based recommendation is one of the oldest recommendation system technique. To recommend an item using a content based recommendation system, a user profile is first generated. This profile is built using the data that the user provides. The user may either provide the data explicitly(rating an item) or implicitly(clicking on a link). This data is then processed, interpreted and organised into relevant information so as to generate the corresponding user profile. Once the user profile is built, the content based recommendation system analyses the contents/features of the items along with the information in the user profile and then recommends a similar non-rated item from the retailer's database, that might be of interest for that particular user.

Usually Collaborative recommender system recommends choices to people based on the opinions of other people who have similar taste. The customer gets recommendations that he/she hasn't rated previously, but that were rated by his/her neighborhood. Basically aggregation of ratings giving by the people (neighborhood) is done, identify similarities based on their ratings available and recommendation is provided. In this paper, a performance enhancement of content-based filtering using diverse collaborative prediction. Especially in movie domain, where there are multiple releases of movies every day, acquiring enough ratings so as to recommend movies to other users is very difficult. Hence collaborative method becomes inefficient. On the other hand, using content-based method and recommending movies similar to the one the user has rated before tends to be inefficient too. Since, we want a system that works efficiently in both condition. [9]

III. PROPOSED METHOD

User adds products to the cart. Based on the products that are frequently purchased and the products that are currently present in the cart will be displayed as recommendations for the active user on the same web page.Registration is done using the web application, customer enters his personal information like name, Mobile number and password.

The customer adds products to the shopping cart on the e- commerce website. The product ids of these products in the cart are passed to the PHP module. The flat file where the association rules are stored, is searched for the corresponding rule whose antecedent matches the product ids in the cart. If a match is found, the product ids from the respective consequent part of the rule is returned. These returned product ids are queried from the product database to fetch the product details of these products which are then displayed as recommendations to the active user. [13]

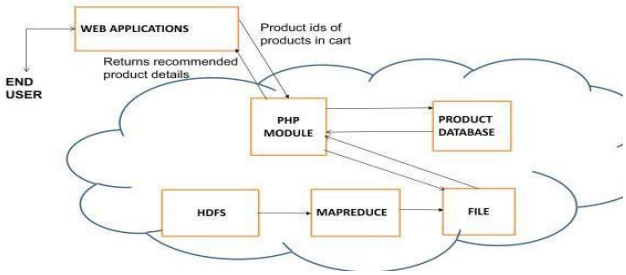

Fig. 1 System Architecture

Caching Association Rules

A Cron job that runs periodically calls a PHP function which invokes the Java MapReduce jobs that compute association rules. The output from these MapReduce jobs are the association rules which are stored in a flat file. This caching of association rules at different intervals of time rather than executing the MapReduce jobs to generate association rules every time a user adds products to the cart improves the performance of the recommendation system.

Association Rules Extraction

HDFS file contains the transaction data of all the previous transactions done in an e-commerce store. It is given as the input to the FrequentItemsetMapper function which is a mapper function written in Java that produces [key,value] pairs where, key is the product id or product ids and value is set to 1.After this step, shuffling of the [key,value] pairs is done thereby, rearranging the [key,value] pairs according to the increasing order of the keys.

The output of the shuffling process is given as the input to the FrequentItemsetReducer function which is a reducer function that calculates and updates the value attribute of the [key,value] pairs based on the number of times [key,value] pairs are repeated. It further eliminates the [key,value] pairs whose value attribute is below the minimum threshold value.AssociationRuleMapper and AssociationRuleReducer functions compare the frequent itemsets to the transactions and builds rules which have a confidence of more than threshold of 0.5 and the output of the above process is stored in flat file.[15]

IV. EXPECTED OUTCOME

On successful registration or login, the user is redirected to the Home Page of the e-commerce website. The page is divided into three sections. There is a section where the list of products that are available are displayed. The two other sections display a cart and the recommendations.

User can shop by adding the products to the cart. When the user adds products to the cart, the recommendations are displayed in the recommendation section of the Home page. The user has added toothbrush to his cart. Hence, toothpaste is shown as recommendation to the user which is the appropriate recommendation as per the association rules generated using the previous transaction data. When user clicks on the checkout button, he is redirected to the checkout page where the user has to enter the shipping address to complete the transaction.After the user clicks on place order, the order is confirmed and the confirmation page displays the order confirmation message. A new record is added to the Transaction.

The accurate product recommendations were made even to a first time user. Therefore, the proposed recommendation system solves the cold-start problem. The proposed recommendation system is scalable and also performs better than the existing recommendation systems. However, it does not provide personalised recommendations as per the user's preferences. Due to this, accuracy of recommendations is low when compared to collaborative recommendation system. The proposed recommendation system can hence be combined with Collaborative recommendation system to form a Hybrid recommendation system. This will solve the cold-start problem which is a major drawback of collaborative recommendation system. The proposed recommendation system can be used to display the

recommendations to the user until a model of the user's preferences is built. Once the model of user's preferences is built, Collaborative recommendation system can be used to display the recommendations.table of the database and the MapReduce Job is run to generate the updated Association rules.

V. CONCLUSION

Recommendation system is an innovative interactive technology for fetching information that can provide a different platform for growth by assisting customer in searching items of their need on the basis of their environment and behaviour. Recommendations on item set that occurs frequently will add new demission by providing associativity in item (or any subset) occur at least as frequently. Associated items in an item set will provide scope for recommending item set to a customer in place of individual items by helping customers to find products which they want to buy plus enabling then to pick product which they must buy. Conversely, they also help business by generating more sales, increasing their revenue.

VI. REFERENCES

[1] Mukta Kohar and Chhavi Rana, "Survey Paper on Recommendation System", International Journal of Computer Science and Information Technologies(IJCSIT), Vol.3(2), pp.3460-3462, 2012.

[2] Jinal S Chauhan, "Survey on Hybrid Recommendation System", International Journal of Advance Engineering and Research Development, Vol.3, Issue 5, pp.347-351, 2016.

[3] Robin Burke, "Hybrid Recommender Systems: Survey and Experiments", Kluwer Academic Publishers Hingham, MA, USA, Vol.12, Issue 4, pp.331-370, 2002.

[4] Jie Lu, Dianshuang Wu, Mingsong Mao, Wei Wang, and Guangquan Zhang, "Recommender System Application Developments: A Survey", Elsevier Science Publishers B.

[5] Amsterdam, The Netherlands, Vol.74, Issue C, pp.12-32, 2015

[6] Gurpreet Singh and Rajdavinder Singh Boparai, "A Survey on recommendation system", IOSR Journal Of Computer Engineering(IOSR-JCE), Vol.17, Issue 6, pp.46- 51, 2015.

[7] Jeff M. Phillips, "L13: Frequent Itemsets", Data Mining; Spring 2013, University of Utah.

[8] Karandeep Singh Talwar, Abhishek Oraganti, Ninad Mahajan, PravinNarsale,"Recommendation system Using Apriori Algorithm","IJSRD - International Journal for Scientific Research &Development,pp.183-185

[9] M. Boley and H. Grosskreutz. Approximating the number of frequent sets in dense data. pages 65–89, 2009.

[10] Rakesh Agrawal and RamakrishnanSrikant, "Fast Algorithms for Mining Association Rules", IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120.

[11] Recommender Systems Handbook, Ricci, F.; Rokach, L.; Shapira, B.; Kantor, P.B. (Eds.), 2011, Springer.

[12] Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl,J., "Analysis of Recommendation Algorithms for E-Commerce. In Proceedings of the ACM EC", Conference. Minneapolis, MN. pp. 158-167, 2000.

[13] Takeaki Uno, Masashi Kiyomi, Hiroki Arimura, "Efficient Mining Algorithms for frequent or closed or Maximal Itemsets", Information.

[14] Agrawal, R., Imielinski, T., and Swami, A. N., "Mining association rules between sets of items in large databases.",In Proceedings of the 1993 ACM SIGMOD International Conference on Management of data pp. 207-216, 1993.

[15] Baralis, Elena; Cagliero, Luca; Garza, Paolo; Grimaudo, Luigi (2015). PaWI: Parallel Weighted Itemset Mining by means of MapReduce. In: 2015 IEEE International Congress on Big Data, New York USA), 26-30 giugno 2015. pp. 25-32

[16] Sudakar singh , Raki Garg,P.M.Mishra :Performance Analysis of Apriori Algorithm with Different Data Structures on Hadoop Cluster. in:International Journal of Computer Applications (0975 – 8887)

[17] Kiran chavan, S.N Patil :Frequent Itemset Mining for Big Data .IEEE 2015

[18] Abishek sakshena,Naveen K Gaur:Frequent Item Set Based Recommendation using Apriori .International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 5.