Service Recommendation Technique for Hotel using multicriteria on Hadoop

Prof. S. B. Choudhri[1], Monika Shelar[2]
[1, 2] Department of Computer Engineering
[1, 2] TCOER, Savitribai Phule Pune University, Kondhwa, Pune-48, Maharashtra, India.

Abstract- *In recent years, Service recommendation[11] system is valuable tools for guiding the user search. Since from last decade the information, customer, data duplication are increasing day by day which yields the problem in Big Data. So the traditional recommendation system is not capable of performing the operation on huge data and which in turn suffer from scalability and inefficiency problems when processing on huge data sets. To address the above problem the service recommendation system for hotel using multicriteria propose new approach of algorithm to solve the problem. This is collaborative filtering based approach for recommendation system which takes the user keyword list and matches it with the previous search result of the user and gives the proper recommendation list to the user. To improve the scalability and inefficiency this paper proposed the implementation in Hadoop environment.*

Keywords- Hadoop, Map Reduce, Big Data, Recommendation System, Hotel,

I. INTRODUCTION

In today's world the data is growing[9] rapidly with an enormous speed, these data is called as 'Big Data'[11]. The problem of big data is , traditional system is not capable of analysing ,visualise ,capture or perform any computation on data, so Big Data is tool to perform any operation. Today, the Big Data challenges will stand out as a challenge to provide more and more hardware efficient management to the environment. The big data applications conjointly impact on the present service recommendation system. The service recommendation system is used to give guidance to the user based on their criteria and provide the appropriate result to the user.

Current service recommendation methods are usually used. Content-based approaches recommend services similar to those the user preferred in the past. Collaborative filtering (CF) approaches recommend services to the user that users with similar tastes preferred in the past. Hybrid approaches combine content-based and CF methods in several different ways. In CF based systems, users will receive recommendations based on people who have similar tastes and preferences, which can be further classified into item-based CF and user-based CF. In item-based systems; the predicted rating depends on the ratings of other similar items by the same user. While in user-based systems, the prediction of the rating of an item for a user depends upon the ratings of the same item rated by similar users.

II. EXISTING SYSTEM

The existing service recommendation technique of Hotel reservation, the ratings of service and lists which is presented to users personalized requirements even most of the recommender system single criteria based system. In fact evaluating a service through multiple criteria and taking into account of user feedback can help to make more effective recommendation for the users. The previous approach solves the scalability[13] problems by dividing datasets but their method doesn't have favourable scalability and efficiency if the amount of data grows.

III. PROBLEM STATEMENT

The existing recommendation system[11] provides a list based on the single keyword and concentrate on the single numerical ratings. Using this system the user who can entered single keyword those can get the accurate result but it fails for the multicriteria based system where user can enter two or three keywords. These will eventually increase the problem of the scalability and inefficiency.

IV. PROPOSED SYSTEM

The proposed system used collaborative filtering method to where the active user preference keywords are matched with the other previous user keyword set.

A. Loading and pre-processing

The large collection of data set is collected from open source data sets that are publicly available from major Travel Recommendation application. The Big Data schema is analysed and working rule of the schema is determined. The JAVA API is built using that can process on the CSV (Comma Separated Values) files. The CSV Files in distributed Systems

are invoked through Web Service Running in the Server Machine of the Host Process through a Web Service Client Process in the Recommendation System. The data that Retrieved to the Recommendation Systems are provided with a clean GUI and can be queried on Demand. Each and Every process on the Recommendation Application invokes Web Service which uses light weighted traversal of data using XML. The Users can Review each hotel and can post comments also. The Reviews gets updated to the CSV Files as it get retrieved. After loading the data the pre-processing activity is performed ,we remove the null value, missing tuples and stop word. Using Porter stemming algorithm remove the morphological words. This phase is used to information retrieval for normalization process.

B. Keyword Extraction(Analysis of User Reviews)

The review of the user is transformed into the keyword candidate list and corresponding domain thesaurus. If the reviews contains a word in the domain thesaurus then the word is added to the preference keyword set.

C. Similarity Computation(Evaluation of Recommendation List)

In the similarity computation[1] the keywords of the active user and the previous user is matching with similar taste of the neighbourhood of the user. If the keyword is not matching then there is no similarity between the active user and the previous user. So the comment which are not similar are removed from tha set.

D. Mapper and Reducer Process

To process on the large data set the open source platform Big Data is used. The preference of the active user and the previous user are transformed into the corresponding active user set (APK) and previous preference keyword set(PPK) respectively[1].

Active User

The active user represents the current user needs, they can mentioned the different keywords for hotel they want.

Previous User

The previous user preference keywords set contain the set of the keywords of user who has visited the Hotel and review on the websites. These keyword are extracted and then put it into the corresponding Doman thesaurus.

Each review will be converted into a keyword according to the keyword candidate list and Domain thesaurus.

E. Analysis(Evaluation Process)

After performing the Map Reduce task of the process ,the result will aggregate to generate the recommendation list to the user. The process which is used is collaborative filtering algorithm to generate the result of the recommendation list.

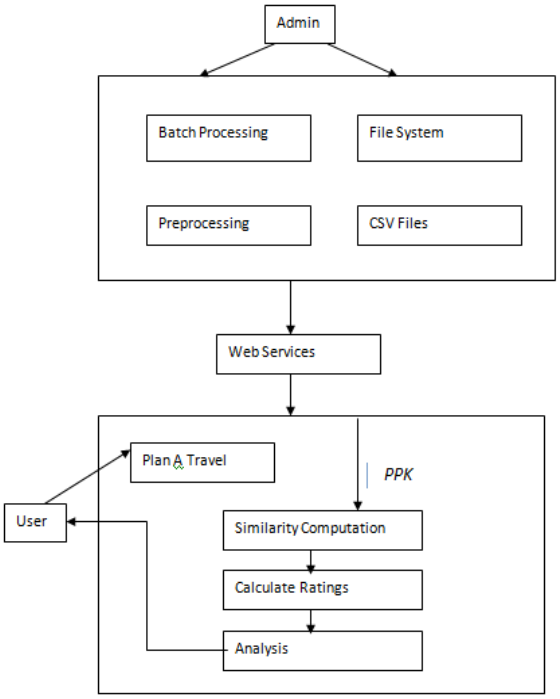

Fig. 1. Service Recommendation Architecture

V. EXPECTED RESULTS

The proposed system is expected to have a comparative analysis of the recommendation system with and without semantic analysis. Also the analytical reports will be generated consisting of accuracy and performance evaluation. Accuracy evaluation will be done for the approximate and exact recommendation system by using the F-measure. The response time and the scalability of the system can be analyzed by judging the performance of the system.

VI. ANALYSIS

The evaluation process takes place to predict the accuracy of the recommendation list. The Evaluation will be in the form of graph. To evaluate the performance we will compared this with the other two well known recommendation method i.e Pearson Coefficient method and item based algorithm[2].

VII. CONCLUSION

Collaborative filtering based method is adopted to improve the scalability and inefficiency of the system using Keyword candidate list. The keyword candidate list and domain thesaurus are used to help to give the preferences to the user. The active user gives their preferences and using collaborative filtering approach it will match with the previous user reviews according to the keyword provided by the active user set. To improve the scalability and inefficiency in the Big Data environment we implemented on the Hadoop framework.

ACKNOWLEDGMENT

I consider myself most fortunate and to have worked under guidance of Prof. S. B. Choudhri Faculty and Computing Department. I would like to thank to all other teachers and friends who are really helping me to make project successfully.

REFERENCES

[1] Shunmei Meng, Wanchun Dou, Xuyun Zhang, Jinjun "KASR: Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications," IEEE Trans. On parallel and distributed system, March 2014.

[2] Greg Linden, Brent Smith and Jeremy York "Amazon.com Recommendations Item to Item Collaborative filtering," IEEE Internet computing, February 2003.

[3] Pasquale Lops, Marco de Gemmis and Giovanni Semeraro "Content-based Recommender Sytems: State of the Arts and trends," Springer Science + Business Media, LLC 2011.

[4] Badrul Sarwar, George Karypis, Joseph Constan and John Riedl,"Item-based Collaborative Filtering Recommendation Algorithm," ACM 1-58113,348-0/01/0005, 2001.

[5] Zhi-Dan Zhao, Ming-Sheng Shang, "User-based Collaborative filtering Recommendation Algorithms on Hadoop," Third international conference on Knowledge Discovert and Data Mining, 2010.

[6] Robin Burke, "Hybrid Recommender Systems: Survey and Experiments".

[7] Yoav Freund, Raj Iyer, Robert E. Schapire, Yoram Singer, "An efficient Boosting Algorithm for Combining Preferences," Journal of Machine Learning Research, 2003.

[8] J. Manyika, M. Chui, B. Brown, et al, "Big Data: The next frontier for innovation, competition, and productivity," 2011.

[9] C. Lynch, "Big Data: How do your data grow?" Nature, Vol. 455, No. 7209, p. 28-29, 2008.

[10] Tom white, \Hadoop the de_nitive guide", O'Reilly 3rd Edition.

[11] Avram Shinar ,David Cunninghma,Benjamin Herta,' M3R:Increased perfor-

[12] mance for In-Memory Hadoop Jobs', IEEE 2012

[13] Radheshyam Nanduri, Nitesh Maheshwori,Reddy Raja, Vasudeva Varma,' Job Aware scheduling Algo for MapReduce framework', IEEE 2014

[14] Toby Segaram, "Programming Collective Engine", O'Reilly.

[15] F. Chang, J. Dean, S. Ghemawat, and W. C. Hsieh, "Bigtable: A distributed storage system for structured data," ACM Transactions on Computer