An Efficient and Dynamic Service Recommendation by Active Customer Preferences

[1]J. Frieda[2] K.Aravindkumar
[#1] Research Scholar, [#2] Assistant Professor
[#1]Department of Computer Science and Engineering, MS University Tirunelveli,

***Abstract*- Recommendation system act as a device in providing most suitable facility to the user. At present, information through online services increases. This leads to the overhead of data in online and there is a chance of reaching less perfect results. In this paper "AN EFFICIENT AND DYNAMIC SERVICE RECOMMENDATION BY ACTIVE CUSTOMER PREFERENCES" using the active customer preferences. In previous approaches, recommendation of service is stand on the feedbacks and ranking from the preceding user. It doesn't consider the idea of the user at a time, who in need of penetrating for the particular service. The proposed system deals with the implementation of personalized recommendation to provide services for hotel reservation system. Preferences are collected from the active user about particular service for each application. Similar user's opinions are taken from the reviews using keyword extraction method and Supervised learning algorithms are used to identify sentiment orientation. It determines positive or negative opinion along with negation word near to each opinion word and then identifies the number of positive and negative opinions of reviews. Keywords with positive opinion are considered and similarity is calculated between user preferences with reviews of the previous user by accord and cosine measures. From this most similar keywords are provided to the user as recommended service. To provide more accurate prediction of the services needed by the active user the proposed system is implemented using Map Reduce framework. The large amount of data can be recovered by using Map Reduce from the investigation of queries meritoriously. The makeable technique for scheduling diminish responsibilities with the help of joining them into the effective characteristic by the well-organized set of guidelines.**

***Keywords*: — Edge-based active contour, edge-stop function, gradient information, image segmentation**

1. INTRODUCTION

. Now a days Internet has enlarged scalability and development, providing outstanding chances for initiatives to organizing their business at a world level with less or without investment. The Internet collect substantial quantities of business data. Enterprises must be capable to process data on time. Related requests can be detected in scientific and Big Data applications. Hence, handling large data volumes in parallel has become progressively challenging. Cloud computing has developed less amount computing substructure on a pay-as-you-go process. In cloud computing, the Map Reduce is used for concurrently running large data sets and separating them into thousands of dispensation nodes in a cluster is a vital concept. Hadoop introduce the Map Reduce, this is an open-source distributed system this system is used by many enterprises, like Yahoo and Facebook etc., for the purpose of handling large data sets. In recent years, the amount of data in our world has been increasing explosively, and analysing large data sets so called Big Data, which becomes a key basis of competition underpinning new waves of productivity growth, innovation, and consumer surplus. Big Data refers to datasets whose size is beyond the ability of current technology, method and theory to capture, manage, and process the data within a tolerable elapsed time. Today, Big Data management stands out as a challenge for IT companies. The solution to such a challenge is shifting increasingly from providing hardware to provisioning more manageable software solutions. Big Data also brings new opportunities and critical challenges to industry and academia.

Opinion Mining also refers to sentiment analysis is the process of analysing the text in the document and provides the suggestions to the people by extracting opinion through online. Users post their opinion about the services or products in the blogs, shopping sites, or review site Reviews about hotel, automobiles, movies, restaurants are available on the websites respectively. Text analysis in opinion mining is the process of getting high quality information from the text. Approximately, 90% of the world's data is available in unstructured format. By parsing this unstructured data, the patterns involved in it are identified and recommendations are provided.

Traditional system provides recommendation to particular application based upon the ranking given by the personalized user. Now-a-days many application uses recommendation system which includes CDs, books, webpage, hotel reservation system and various. In hotel reservation system, if one user is concerned about particular services and another user is looking for different services in the same hotel. But the same recommendation service is provided for both the user. It is not the good recommendation and the people will not satisfy to the recommendation. Moreover, in hotel reservation system the ratings of services and service recommendation list to the users are same does not consider user preferences. Recommendation system is classified as content based, collaborative based and hybrid based recommendation system. Content based recommendation provides recommendation by taking the user preference from the previous user reviews. Collaborative Filtering (CF) recommends service based on the reviews of the previous user, by checking the similarity with the current user. Hybrid recommendation system combines recommendation of both content and CF.

2. *HADOOP*

Hadoop is generally a system that allows high Presentation between dispersed nodes hadoop is also a multi-tasking system so it allow to process multiple data sets in multiple jobs for multiple users within the same time

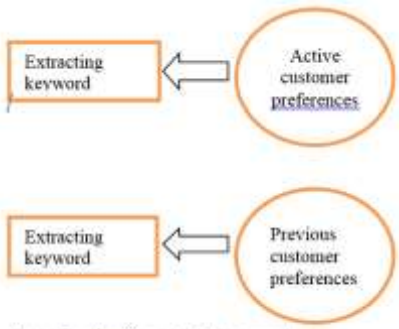

Figure1: nodes for extracting keywords

Hadoop is a popular open source framework for storing and processing large data sets across clusters of computers. Hadoop allows for distributed processing of data across multiple local sites. it has built in libraries. It also uses apache hive data warehouse. This is used for querying and managing large datasets in distributed storage. this using the query and data in SQL. it allow to use the map reducing programming

3. *MAP REDUCE*

. Keywords with positive opinion are considered and similarity is calculated between user preferences with reviews of the previous user by accord and cosine measures. From this most similar keywords are provided to the user as recommended service. To provide more accurate prediction of the services needed by the active user the proposed system is implemented using Map Reduce framework. The large amount of data can be recovered by using Map Reduce from the investigation of queries meritoriously. The makeable technique for scheduling diminish responsibilities with the help of joining them into the effective characteristic by the well-organized set of guidelines

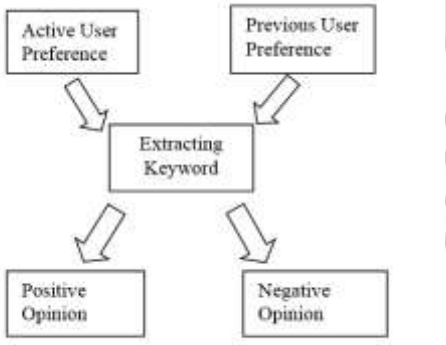

Figure2: Working Of Recommendation System

4. *LITERATURE SURVEY*

1. *MULTI-CRITERIA USER MODELING IN RECOMMENDER SYSTEMS*

Author -

Data center-scale c Kleanthi Lakiotaki, Nikolaos F. Matsatsinis

Year: 2011

In parallel, Multiple Criteria Decision Analysis (MCDA) is a well established field of Decision Science that aims at analyzing and modelling decision maker's value system, in order to support him/her in the decision making process. In this work, a hybrid framework that incorporates techniques from the field of MCDA, together with the Collaborative Filtering approach, is analyzed. The proposed methodology improves the performance of simple Multi-rating Recommender Systems as a result of two main causes; the creation of groups of user profiles prior to the application of Collaborative Filtering algorithm and the fact that these profiles are the result of a user modelling process, which is based on individual user's value system and exploits Multiple Criteria Decision Analysis techniques. Experiments in real user data prove the aforementioned statement.

2. *RANKING PREDICTION FOR CLOUD SERVICES*

Author – Zibin Zheng

Year: 2013

QoS rankings provide valuable information for making optimal cloud service selection from a set of functionally equivalent service candidates. To obtain QoS values, real-world invocations on the service candidates are usually required. To avoid the time-consuming and expensive real-world service invocations, this paper proposes a QoS ranking prediction framework for cloud services by taking advantage of the past service usage experiences of other consumers. The proposed framework requires no additional invocations of cloud services when making QoS ranking prediction. Two personalized QoS ranking prediction approaches are proposed to predict the QoS rankings directly.

3. ***TOWARDS TV RECOMMENDER SYSTEM: EXPERIMENTS WITH USER MODELLING***
Author - Milan Bjelica
Year: 2010

In this paper, author analyzes such recommender system design under the broadcast scenario, where uplink connection to the network center is not available. Author put special emphasis on user modelling algorithm that would be able to efficiently learn the user's interests. The system proposal applies the elements of machine learning and pattern recognition, as well as the information retrieval theory, like vector spaces and cluster hypothesis. The derived algorithm is computationally simple, while experimental results show high acceptance ratio of the proposed recommendations. To the majority of users, the most important feature of digital television is probably a large number of available programs. Indeed, cable operators often advertise their services by stating the number of programs they offer. Surprisingly, the users do not necessarily benefit from this over abundance of available content; instead, they repeatedly find it difficult to retrieve the right program at a given time and consequently watch only few programs from the available myriad on a regular basis. Clearly, this outcome satisfies neither service/content providers nor the users.

5. *SYSTEM DESGIN*

. Hadoop supports the Map Reduce programming model originally proposed by Google and it is a convenient approach for developing applications. The system discusses the core algorithms for collaborative filtering and traditional means of measuring their performance against user rating data sets. The system will then move on to discuss building reliable, accurate data sets, understanding recommender systems in the broader context of user information needs and task support and the interaction between users and recommender systems..

Collaborative filtering (CF) is a popular recommendation algorithm that bases its predictions and recommendations on the ratings or behaviour of other users in the system. The fundamental assumption behind this method is that other users' opinions can be selected and aggregated in such a way as to provide a reasonable prediction of the active user's preference.

The focus of this survey is on collaborative filtering methods, although content-based filtering will enter our discussion at times when it is relevant to overcoming a particular recommender system difficulty. The majority of collaborative filtering algorithms in service today, including all algorithms detailed in this section, operates by first generating predictions of the user's preference and then produces their recommendations by ranking candidate items by predicted preferences.

CF is a straightforward algorithmic interpretation of the core premise of collaborative filtering find other users whose past rating behaviour is similar to that of the current user and use their ratings on other items to predict what the current user will like. For example, to predict Mary's preference for an item she has not rated, user–user CF looks for other users who have high agreement with Mary on the items they have both rated. These users' ratings for the item in question are then weighted by their level of agreement with Mary's ratings to predict Mary's preference.

Pre-computation and truncation is essential to deploying collaborative filtering in practice, as it places an upper bound on the number of items which must be considered to produce a recommendation and eliminates the query-time cost of similarity computation. It comes with the small expense of reducing the number of items for which predictions can be generated.

The system deals with the implementation of personalized recommendation to provide services for online shopping system. Preferences are collected from the active user about particular service for each application. Similar user's opinions are taken from the reviews using keyword extraction method and supervised learningalgorithms are used to identify sentiment orientation. It determines positive or negative opinion along with negation word near to each opinion word and then identifies the number of positive and negative opinions of reviews. Keywords with positive opinion are considered and similarity is calculated between user preferences with reviews of the previous user by accord and cosine measures. From this most similar keywords are provided to the user as recommended service. To provide more accurate prediction of the services needed by the active user the proposed system is implemented using Map Reduce framework.

6. *METHODOLOGY*

1. *NCHANNEL CLUSTERING*

Stemmer is used to remove the inflected part of the word to get their root form. It is used to reduce the word to its root form. Different variants of a term can be conflated to a single representative form. It saves storage space and time. A stemming is a technique used to reduce words to their root form, by removing derivational and inflectional affixes. The stemming is widely used in information retrieval tasks. Many researchers demonstrate that stemming improves the performance of information retrieval systems. Stemmer is the most common algorithm for English stemming. Stemming is a technique to detect different inflections and derivations of morphological variants of words in order to reduce them to one particular root called stem. A word's stem is its most elementary form which may or may not have a semantic interpretation. In documents written in natural language, it is hard to retrieve relevant information. Since the Languages are characterized by various morphological variants of

words, this leads to mismatch vocabulary. In applications using stemming, documents are represented by stems rather than by the original words. Thus, the index of a document containing the words "computing", "compute" and "computer" will map all these words to one common root which is "compute". This means that stemming algorithms can considerably reduce the document index size, especially for highly inflected languages, which leads to important efficiency in time processing and memory requirements.

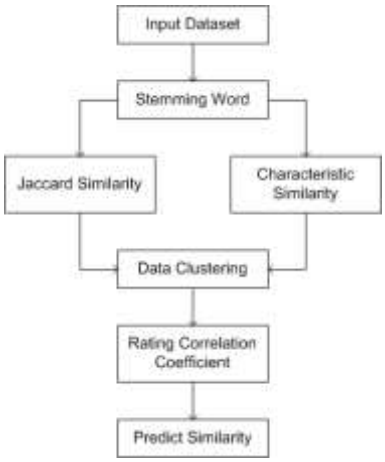

Fig 3 Architecture diagram

2. ***SIMILARITY MEASURES***

Jacquard and characteristic similarity has been processed between the set of services. In-order to enhance the frequency rate mechanisms the system find the weights of attributes and ranking it there by improve the Search scenario. Web Services data has to be categorized according to the set of open service descriptions and their properties. String matching mechanisms usually consist of keyword based search mechanisms and their degree of matching

3. ***COLLABORATIVE FILTERING***

Clustering of web documents enables (semi-)automated categorization, and facilitates certain types of search. Any clustering method has to embed the documents in a suitable similarity space.

4. ***RATING SIMILARITY AND PREDICTED RATING***

PCC is applied to compute rating similarity between each pair of services in CF. Ranking algorithm compute similarity between document and query vectors to yield a retrieval score to each document. According to the relevance with the user query retrieved document are ranked. Based on the enhanced rating similarities between services, neighbours are predicted.

7. *PERFORMANCE EVOLUATION*

Collaborative based Service clustering achieves less number of clusters compare to whole system of clusters. Proposed system achieves less executional time. Performance is measured in terms of (Parameters) computation time, no of clusters and memory usage.

7.1 CPU USAGE

Result = end_usage – (dataset * (allotted_ usage***)***

The chart display the difference between the average CPU usage of analyzing datasets using existing and proposed systems

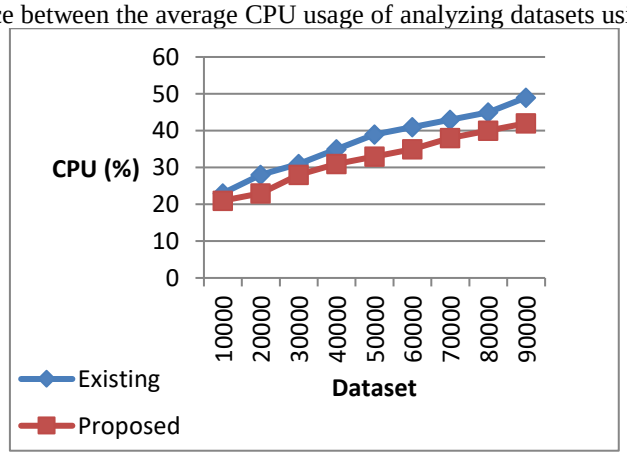

Fig 7.1 CPU Usage

7.2 MEMORY USAGE

Result = (dataset / current) + (startMemory – endMemory

The chart display the difference between the average memory usage in MB of analyzing datasets using existing and proposed systems.

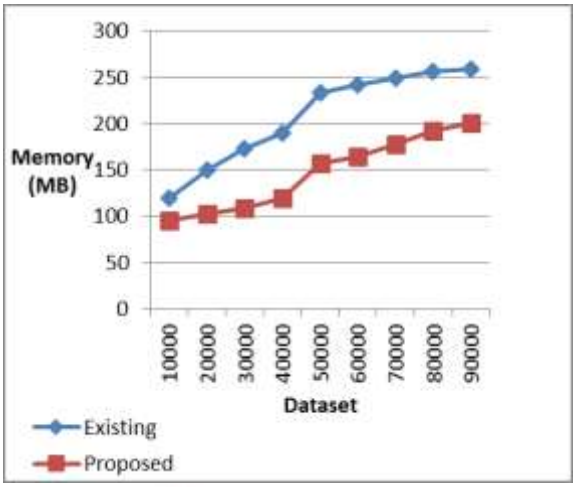

Fig 7.2 Memory Usage

7.3 PREDICTION ACCURACY

Result = ((predicted result) / (original result)) * 100

The chart display the difference between the accuracy of prediction level in percentage for usage of analyzing datasets using existing and proposed systems.

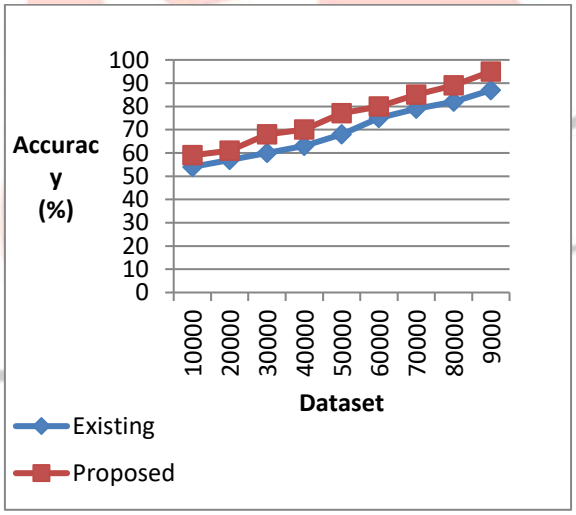

Fig 7.3 Prediction Accuracy

8. *REFERENCES*

[1] Ghodsi, A., Zaharia, M., Shenker, S., and Stoica, I., "Choosy: max-min fair sharing for datacenterjobs with constraints," In Proceedings of the 8th ACM European Conference on Computer Systems,pp. 365–378, 2013.
[2]I Hammoud, M., and Sakr, M. F., "Locality-aware reduce task scheduling for MapReduce," In Proceedingsof IEEE Third International Conference on Cloud Computing Technology and Science (Cloud-Com), pp. 570–576, 2011.
[3] Ghoshal, D., Ramakrishnan, L.,"Provisioning, Placement and Pipelining Strategies for Data-IntensiveApplications in Cloud Environments," Cloud Engineering (IC2E), 2014 IEEE International Conference, pp. 325–330, 2014.
[4]," NMR Biomed., 2011. Joe Weinman, "Cloud Computing is NP-Complete" Working Paper, 2011