International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 8 5215 - 5219

Web service Recommendation by combining QOS and user comments

Sayali Narkhede
Computer Engineering Department
SKN Sinhgad Institute of Technology and Science
Lonavala, Pune, India
Sayali.rameshnarkhede@gmail.com

Prof. Praveenkumar Keskar
Computer Engineering Department
SKN Sinhgad Institute of Technology and Science
Lonavala, Pune, India
praveen.keskar@gmail.com

***Abstract*—** Due to well gaining experience of internet its user's expectation from the search engines increases dramatically. Due to this search engines capability is not only limited to the providing desired URL's to the users query. Moreover to this search engines are expected to provide information by analyzing in proper way like by doing surveys and recommendations. So many recommendation systems are existed which are working on some limited aspect of the parameters for recommending a web service. This paper represents a method of recommendation which considers users opinion and quality of the service parameter of the web service. The proposed idea captures the response time of the user transaction for a web service along with the users opinion comments about the web service. Then by combining both a new hybrid recommendation system is introduced which efficiently provides the recommendation that is more accurate and fine grained. This hybrid recommendation is powered with the Pearson correlation and strong NLP protocols to attain most accurate state.

Keywords:-*Pearson Correlation, Colloborative filtering, recommendation, preference functio, Response time.*

*****

I. INTRODUCTION

In today's era of development more and more peoples are started attracting towards business rather than the job. So to increase the business World Wide Web will be the first choice of the business tycoons compare to the traditional advertising media like newspaper, magazine etc. Because of this trend web services grabbed lots of attention. As the web services gives the data in more convenient manner however user wants. So the number of the users who used these web services is also increasing.

Most of the users are making use of the others comments on the service before availing the service. Since others recommendation gives a valuable opinion about selection or discrimination of that service, So that the recommendation algorithms start gaining lots of attention.

Recommendation systems are the system which comes under the category of information filtration system. Such systems are used to find out the rating or the preferences given by the previous users on the system. Because of their dynamic behavior they have been used in various applications such as news, music, movies, shopping portals etc.

Generally recommendation systems are widely classified in two groups

1. Content based filtering
2. Collaborative filtering
3. Hybrid recommendation systems

Content based filtering : In content based filtering a number of characteristics of the particular items are taken as a base for the recommendation. In simpler word content based filtering makes use of past likes and dislikes of the item by the user. Based on this likes an item having similar characteristics are recommended. So to present the item, Item recommendation algorithms are used. Tf-idf algorithm is one of such algorithm.

Collaborative filtering : Here the last behaviors of the user are taken as a base for the recommendation. E.g. comments of the user on the item can be used to give recommendation. One of the great advantage of the collaborative filtering algorithms are they can give the recommendation without finding the item. Because of this advantage they are widely used in the applications such as movie recommendation.

There are various algorithms used for this purpose like k-nearest neighbor, Pearson correlation etc. Collaborative filtering algorithms are worked on the assumption like if user like one item then in future also they will go to choose the same items only. Number of famous websites such as amazon, Facebook, twitter etc. makes use of this method.

Hybrid recommendation system : A survey was conducted on best method of recommendation, which conclude that there are some application areas where both of the approaches can be combined very easily. The system which combines content based filtering and collaborative filtering are known as hybrid recommendation systems. Netflix is one of the best examples of the hybrid recommendation system. Netflix combines the both examining and searching behavior of the user to recommend the new one.

Generally seven hybridization methods are there

- Weighted
- Switching
- Mixed
- Feature Combination
- Feature Augmentation
- Cascade
- Meta-level

There are number of factors that can be considered for recommendation. Quality of service of the web service is one of such category used for recommending the things. The QOS is measured by finding the response time, failure probability

and throughput. QoS may vary from the system to system as network contributes the result of QoS of the web service.

Preprocessing is widely used in mining domain where there is lots of unnecessary data is presents along with the useful data. Such unwanted data increases the time complexity of the system, so it becomes a must to remove these data. Normally preprocessing is carried out in 3 sub processes.

- Special symbol removal : This process discards the all special symbols from the data to be processed as they are not going to change the meaning of data even if get removed.
- Stemming: Stemming is a well-known method of bringing the derived word to its base form.
- Stop words removal : It is one of the most important methods of preprocessing as it reduces the size of the data to the great extent.

Pearson correlation is one of the best suitable methods used for the purpose of finding linear dependencies between the two entities. It tells the degree to which extent two entities are depends on each other. In case of recommendation systems, Pearson correlation is used to find out the correlation between the ratings of the two users or likes of the two users.

Pearson correlation can gives three values as a output

1. Positive correlation
2. Negative correlation
3. No correlation

Figure 2 elaborates these 3 value n very simple manner

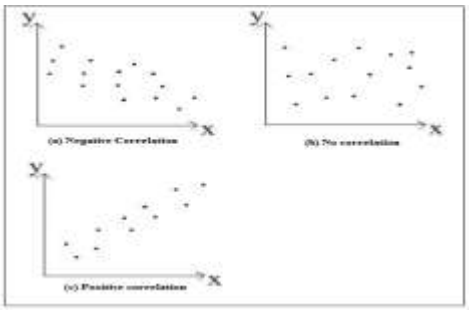

Figure 1 : Correlation types

The rest of the paper is organized as follows. Section 2 discusses some related work and section 3 presents the design of our approach. The details of the results and some discussions we have conducted on this approach are presented in section 4 as Results and Discussions. Sections 5 provide hints of some extension of our approach as future work and conclusion.

II. LITERATURE SURVEY

This section represents all the related works of technologies used in our proposed method.

[1] Present's a deep study on the different stemming methodologies used to obtain the stems. Figure 5 explains all the algorithms that come under the discussion category with their some methodologies also. As figure 3 shows, stemming algorithms are generally classified in three main methods truncating, statistical and mixed.

Truncating methods removes the suffixes and prefixes also known as affixes of the word, statistical methods removes the affixes by applying some statistical methodologies whereas mixed approach makes use of both to become more powerful.

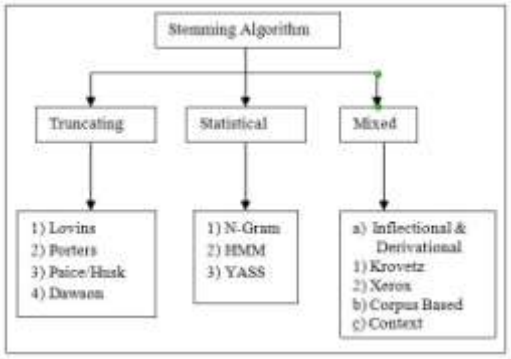

Figure 2: stemming algorithms

[2] Gives a technique to find out the association between the different properties of the QoS parameters. In advance to this author also find out the users those have insufficient QoS information and labeling them as malicious. For proper management and discovery of the web services an intelligent recommender system is developed. Also side by side they focused on the searching and ranking algorithms which contributes the result of the recommendation.

To recommend the precise web service according to the interest of the user is very difficult task. By considering this problem as a base [3] proposed a recommendation theory. Here a semantic matcher framework was designed for matching of the web services. Also the framework consists of recommendation algorithm which recommends the best web service from the list of matched services. Developer designs the system in such way that it can easily adopt the dynamic nature of the user requirement. Since the system not generates the DAML description of the web service, they kept it as a future work of this study.

In [4] Pradip gives a good survey on collaborative filtering algorithm used for the purpose of recommendation. Here author had a short look on the QoS characteristics by clustering the users into five regions. [5] Narrates a new approach for recommendation that makes use of both i.e. ratings and profile content.

To do so a relationship between the ratings is find out by using some dynamic features. To bring the idea into reality developer makes use of Top-k algorithm, Pearson correlation, model based algorithm, memory based algorithm and collaborative filtering.

[6] adds more preciseness in to the answer nikita et. Al demonstrates a theory that makes use of location to give recommendations. Again the collaborative filtering algorithms are at the base of the system. To address the problem of QoS

based recommendation, [7] explores a skyline based QoS approach. Author states that in spite of being large size of the dataset system behaves well for recommendation. Here they find out the local quality level of the web service which improves the factors responsible to attract the user towards the site.

[8] Exposed an importance of knowledge on the recommender systems and there relative study. After studying vast set of related literatures authors find that the semantic of the texts has a great impact on the result of the recommender systems. Also they states that how the different knowledge categories can be used to overcome the drawback of the recommendation systems.

III. PROPOSED METHODOLOGY

Proposed system designed a scenario where some 5 web service for bus booking services is been incorporated . Where every user first creates his user profile and then book the bus to his desired destinations.

So for every transaction carried out by the user is recorded with its response time at the web server. Then after getting some threshold number of response times our system will generate the recommendation purely based on response time of the user. This we call as Fresh recommendation.

Then user enjoys this fresh recommendation and then goes ahead at the end to give his opinion in the survey. Then this opinion will be used for the collaborative recommendation. Then by finally merging and measuring the weight of the two recommendations, system will suggest the best web service to the user.

So here proposed system of hybrid recommendation system for web services using QoS parameter and user comments with the below mentioned steps as shown in figure 3.

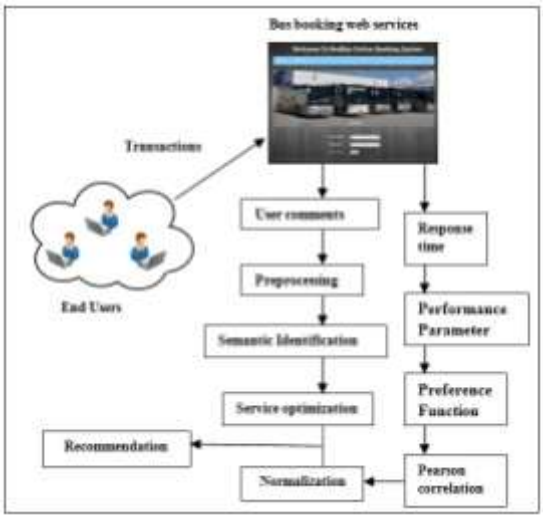

Figure 3: Overview of the proposed work

Step 1*:* This is the initial step our project where many users are accessing the different web applications of bus booking sites and performs the transactions.

Where on each transaction the response time is captured at the server's end and store in the database. And after completing the user transaction a survey is conducting to the user where the respective users are been asked for their opinion about the web services and their opinions are been stored in the database promptly.

Step 2: Here in this step mean and standard deviation of the repose time is been calculated as the performance parameter.

Step 3: In this step a set of ranges of the response times are been calculated which are oscillating by the means of standard deviation along the average response time considering the preference function

(A) PREFERENCE FUNCTION

$$\Psi(i,j)= \sum_{V \epsilon N(u)^{i,j}} W_v (q_{v,i} - q_{v,j}) \quad \text{----------(1)}$$

$\Psi(i,j)$ = Preference function
$N(u)^{i,j}$ = Subset of similar user
V= Similar user of the current web service u
W_V = Weighting factor
q_v = Quality of service of two web services

Step 4: Here in this step correlation is calculated for the web service indices based on their aggregation values of the other users by applying Pearson correlation.

(B) PEARSON CORRELATION

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \quad \text{(2)}$$

N = number of pairs of scores
X,Y = two variance entities

Step 5: Here filtering is done by setting a threshold value of 0.5 for correlation score to get fine grained web service names for the hybrid recommendation in the coming steps.

Step 6: In this step all the users comments are been fetched from the database for preprocessing job to get useful text for content recommendation.

Step 7: Semantic recommendation is the process of identifying good and bad words by incorporating the strong NLP protocols. The result of this Best web service are selecting based on the weight of good semantics.

And then finally the web services collected in the step 5 and web services collected from the NLP implementation are

merged to get the best recommendation that is representing both system performance and also it is the user's favorite.

IV. RESULTS AND DISCUSSIONS

To show the effectiveness of proposed system some experiments are conducted on java based windows machine using Apache tomcat as the server. And system is set many tests to evaluate its performance under different parameters.

4.1 Mean Absolute Error: Mean Absolute Error (MAE) metric is frequently using parameter to measure the error rates of the recommendation systems, which is defined as:

$$MAE = \sum_{i,j} | r_{i,j} - r'_{i,j} | / N \quad \text{-------(3)}$$

Where $r_{i,j}$ denotes the expected recommendation of Web services , whereas , $r'_{i,j}$ denotes the predicted recommendation of the web services, and N denotes the number of predicted values. When our model of recommendation designed by using QOS and user comments parameter is compared with other systems for the measured MAE the results can be shown in the below table 1.

Methods	MAE
UserBsd	0.7062
SVD	0.7124
BPMF	0.6944
PFM	0.69
Our Approach	0.6435

Table 1: Performance Comparison

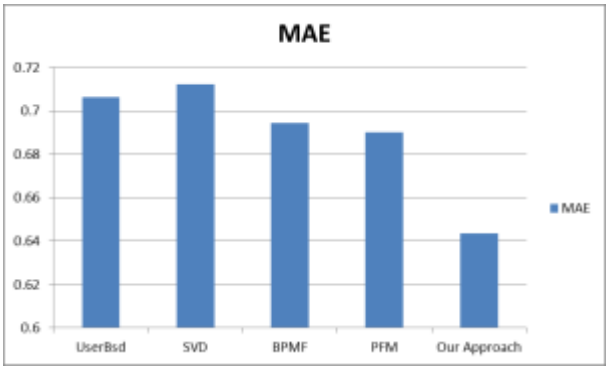

Figure 4 : Performance Graph

The plot in the figure 4 clearly indicated that recommendation of our system which is achieved by the efficient combination of user comments and QOS parameters is having the lesser MAE than the other systems as stated in [10].

4.2 RMSE (Root mean square Error) :

The use of RMSE is very common and it makes an excellent general purpose error metric for numerical predictions.

Compared to the similar Mean Absolute Error, RMSE amplifies and severely punishes large errors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{-(4)}$$

On plotting of the RMSE values for different number runs in between our method and adoptive weighting algorithm as discussed in [9] we found the graph as shown in figure 5.

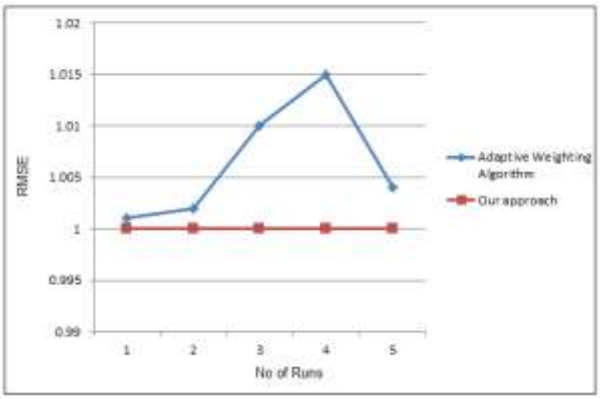

Figure 5: Accuracy of the proposed algorithm comparison

Above graph indicates that proposed system is having less RMSE value than the adaptive weighting algorithm model .This shows the better performance of our idea of using hybrid recommendation with the QOS parameter and user contents.

V. CONCLUSION AND FUTURE SCOPE

In this paper an enriched hybrid recommendation method is proposed. This recommends the web services for the bus booking sites hosted locally based on the response time of the user transactions and users opinion about the web services. Proposed method successfully captures the response time for the each performed transactions and then by proper incorporation of preference function and Pearson correlation best web service are categorized based on the quality service parameter.

Then to achieve more accuracy system collects the user's comments from the web pages where they are written after availing the service as a part of the survey by the user. These comments are properly scrutinized using strong NLP protocols for service optimization. Then optimized services are combined with categorized web services based on QOS parameter to yield best hybrid recommendation system which eventually yields the fine grained results in all aspects.

The proposed system can be enhance to consider more technical parameters of the web services like probability of failure, throughput to enrich the recommendation system.

REFERENCES

[1] Ms. Anjali Ganesh Jivani "A Comparative Study of Stemming Algorithms" Anjali Ganesh Jivani et al, Int. J. Comp. Tech. Appl., Vol 2 (6), 1930-1938 IJCTA | NOV-DEC 2011

[2] "An Enhanced Web Service Recommendation System with Ranking QoS Information" S.Suria1 and K.Palanivel2 International Journal of Emerging Trends & Technology in Computer Science Volume 4, Issue 1, January-February 2015

[3] Manikrao, Umardand Shripad, and T. V. Prabhakar. "Dynamic selection of web services with recommendation system." *Next Generation Web Services Practices, 2005. NWeSP 2005. International Conference on*. IEEE, 2005.

[4] "Review on Collaborative Filtering and Web Services Recommendation" Pradip N. Shendage, International Journal of Engineering Research and General Science Volume 2, Issue 6, October-November, 2014 ISSN 2091-2730

[5] "Web Service Recommendation Using QOS Parameters and Users Location" Asst.Prof. N.K. Patil1 , Priyanka pawar2 , Shankar More3 , Bhausaheb Tupe4. International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 3, March 2015.

[6] "Personalized QoS-Aware Web Service Recommendation via Exploiting Location and Collaborative Filtering" Nikita R.Gurjar1 , Sandeep V.Rode2 Volume 5, Issue 1, January 2015 . International Journal of Advanced Research in Computer Science and Software Engineering

[7] "Selecting Skyline Services for QoS-based Web Service Composition" Mohammad Alrifai , Dimitrios Skoutas, Apr 26, 2010

[8] Amini, Bahram, Roliana Ibrahim, and Mohd Shahizan Othman. "Discovering the impact of knowledge in recommender systems: A comparative study." *arXiv preprint arXiv:1109.0166* (2011).

[9] Xiangyu Tang and Jie Zhou , " Dynamic Personalized Recommendation on Sparse Data ", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 12, DECEMBER 2013.

[10] Kai Zhang, Keqiang Wang, Xiaoling Wang, Cheqing Ji and Aoying Zhou , " Hotel Recommendation based on User Preference Analysis" , IEEE ,ICDE Workshops 2015.