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. 4, Issue 5, May 2016

Diversity Based Web Service Recommendation using Clustering

Pooja Kumbhar, Prof G. S. Mate

PG Student, Department of Computer Engineering, JSPM's RSCOE Savitribai Phule Pune University, Pune, India

Associate Professor, Department of Information Technology, JSPM's RSCOE SavitribaiPhule Pune University, Pune, India

ABSTRACT: While dealing with the web services Quality of service (QoS) is most known and important aspect. QoS gives non-functional feature and also acknowledgment of the web services all over the World Wide Web. We make use of various filtering schemas together based on QoS values as well as pros of the present use of analysis of services. QoS explains non practical web service execution and gives service selection [3]. QoS is classified as group of user attentive attributes with reaction time, accessibility and reputation etc. The most common and preferable technique is collaborative filtering method. Here in this paper we introduce a new method for web recommendation which makes sure users QoS recommendations and also type of features in which user are interested web services. From observations of last few years' utilization of online services shown that the users are very much interested and also QoS recommendations on web services area. At last, users ranking of web services is arranged by finding their scores and also quality of degrees occupied depending on online services graph through applying diversity aware web service ranking algorithm [6]. In projected system, web services clustering are done by making use of hierarchical clustering algorithm as well as web services which are organized by diversity based web service ranking algorithm. After the ranking of web services top k list of web services is generated to fulfill user's requirement and QoS recommendations. Our method used for recommending web services to increase the performance and also quality of recommendation outcomes depending on hierarchical clustering.

KEYWORDS: Web service recommendation, hierarchical clustering, service usage history, diversity, web service

I. INTRODUCTION

Web services are being more popular in last few years and also hasan important part with different applications such as business application, e-commerce etc. there are number of web services present on web also the number of web applications will become a problem having tendency of service computing group [1]. Also the count of web services is present with comparable workings and different without any functional quality. To choosing effective web service which is able to fulfill the users non-functional as well as functional needs.

The recommendation system is concentrated on the concept of giving the proper recommendation that meets with the user's needs. A common recommendations system must give small amount of regular papers that is further reduces the user's interest. Though in realty user strictly entered by the papers written on the subject further the recommendation system us efficient to positioning them depending on references [4]. This system is not proper which outcome demonstrated similar pages more than once and not the different one.

If the k top recommendations are same with others, in that much of them are not beneficial for the user in the system. That is why the utility of k recommendations can be drastically small. If the system gives various set of cases to the user as optimal scope of the data area. In systems like Zhou et al. which explained the diversity accuracy perplexity of recommendation system considered as crucial as similarity. Karl Waldemar Ziegler et al. introduced a recommendation is frequently maximized by subject diversity. These convictions, author has a mind tosolve that various diversities are additionally needed attribute in service recommendation systems. In this analysis [5], we provide a recommendation system which is based on diversities. QoS utility and types of selection of web services which well diverse top-k services is already used for exhorted to users[2].

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. 4, Issue 5, May 2016

In this paper, the top-k recommendation web service list is given to the user with the diversity of web services. The system takes into consideration functional and non-functional requirements of the user with the diversified recommendation list. User's usage history is mined, through the text similarity and operation similarity. Then active user's potential interest is calculated through the collaborative filtering approach. QoS preferences are mined from user's usage history. Scores are calculated for each web service candidate through interest relevance and QoS utility. Ranking of optimal web services is done through diversity-aware web ranking algorithm.

II. RELATED WORK

Guosheng Kang et al. [1] illustrated a new technique for recommending web services to the clients. In this authors incorporated user's probable QoS recommendations and also various quality features of user's interest in Web services depending on exploring user's history of the Web Service.

Min Gong et al. [2], gives service recommendation algorithm known as URPC-Rec (User Relationships Preferences Clustering and Recommendation). In this algorithm, basically, requested services depending on their history behaviorsused by clusters and after that specifications are given.

Yutu Liu, Anne H.H. Ngu, LiangzhaoZeng [3] developed method to overcome the issue of multi-user web service selection. System identifies the missing multi-QoS values based on historical QoS experience of user as well as after that select generally utilized technique for multi-user by our quick match technique.

Shagging Wang et al. [4] presented a new concentrated QoS expectation structure with location based regularization (LBR). Author initially added up in most popular Matrix Factorization (MF) method for values which is not given in expectation.

Wei Lo et al. [5] gave a methodto solve the issue of the "ensured" quality. To overcome this issue authorgiven a history record- based service minimization method called as Hire Some. Integrity of service arrangement is maximized bythistechniqueas well as by taking advantage of a web service's QoS history records instead of developing the given QoS values endorsed by the provider of the service.

Lin Wenmin et al. [6], gives a method depending on idea of horizon to efficient and successfully select services for structure. Also minimize the amount of applicant services to be determined.

R.-H. Li, et al. [7], evolved an effective greedy algorithm to get near-optimal diversity depending on ranking along with the linear time and space complexity respectively size of the graph. Mohammad Alrifai et al. [8] gave a public, genuine and dynamic QoS calculation method for web services. Web services are chosen by executing and also by observing with a QoS registry in an assumed telephone service giving business application.

III. PROBLEM DEFINITION

In the system the web service recommendation with the verity by developing clustering method. Based on hierarchical clustering to the clustering method to cluster the service as well as diversity based web service ranking algorithm to rank the web service is used by system.

IV. IMPLEMENTATION DETAILS

A. *Proposed System:*

Input to the system is the web service WSDL files. Also the hierarchical clustering algorithm is created on the dataset. For determining the user interest in particular web service we have to study user history. By applying collaborative filtering (CF) method similarities are searched depending on possible related interest. QoS are calculated on the web services dataset from internet. For generating web service graph, ranking is performed based on QoS and user interest also depending on that outcomes are evaluated. At the end of process we will have the top k web services recommendations as an outcome of the system. The developed hierarchical clustering algorithm is designed on the cluster of the web services and WSDL files. It maximizes accuracy of a system.

- Clustering:Hierarchical clustering is a method for cluster analysis that creates a hierarchy of clusters developing proper measure in data mining.

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. 4, Issue 5, May 2016

- Web services ranking: depending on the web service graph top-K ranked web services are gathered [1]. Web service graph has nodes and edges where node shows web service and edges are the network in them.

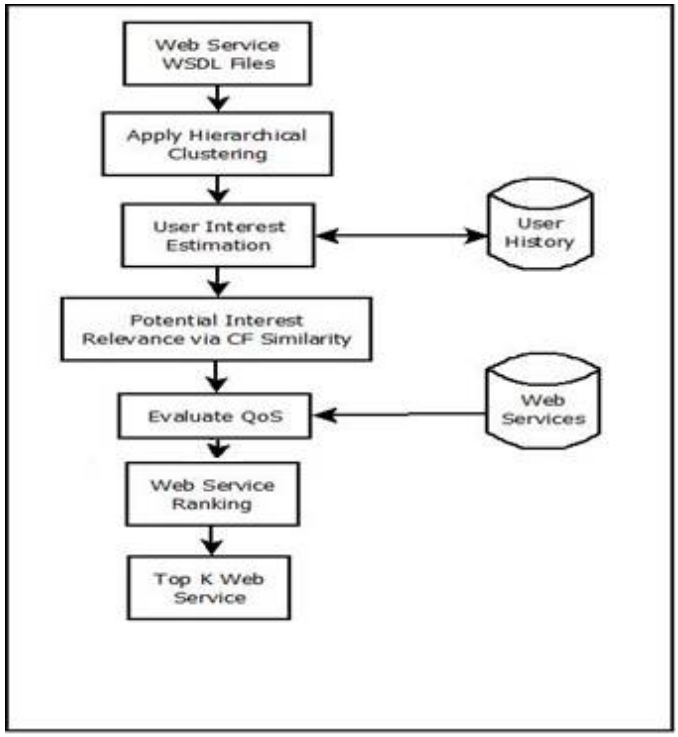

Fig. 1 Proposed System Architecture

B. *Description of the Proposed Algorithm:*

The proposed system implements the selection of web services through functional and non-functional evaluation using clustering.

STEP I: User Query Processing
Usage history dataset is mined to get the relevant web services related to the user query
Similarity is evaluated as:

1) Term Frequency

$$tf(t_j, WSDL_i) = \frac{freq(t_j, WSDL_i)}{|WSDL_i|}$$

where t_j is the j^{th} term in the corpus; $WSDL_i$ is the WSDL document of the i^{th} Web serviceWS_i; $freq(t_j, WSDL_i)$is the occurrence number of t_j in $WSDL_i$; $|WSDL_i|$ is the total occurrence number of all meaningful terms in $WSDL_i$.

2) Inverse Document Frequency

$$idf(t_j, WSDL_i) = log_2 \frac{|WSDL|}{|\{WSDL_i : t_j \epsilon WSDL_i\}|}$$

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. 4, Issue 5, May 2016

Where$|WSDL|$ is the total number of WSDL documents, which equals to the number of Web services, i.e.,$|WSDL| = N$; $|\{WSDL_i : t_j \epsilon WSDL_i\}|$ is the number of documents where the term t_j appears.

3) Weight Calculation

$$\omega_{i,j} = tf(t_j, WSDL_i) * idf^2(t_j, WSDL_i)$$

A high weight in TF/IDF is achieved through a high term frequency and a low document frequency of the term in the whole collection of documents. The weights hence tend to filter out very common terms.

4) Cosine Similarity

$$texSim = \cos(W_i, W_j) = \frac{w_i . w_j}{w_i \times w_j}$$

Where $|w_i|$ and $|w_j|$ are the Euclidean length of the vector w_i and w_j respectively, and the numerator is the dot product of w_i and w_j .

5) Web Service Similarity

$$wsSim(WS_i, WS_j) = \varphi texSim + \emptyset opSim$$

STEP II: Collaborative filtering Similarity

Potential user interest is calculated using collaborative filtering. Web service recommendation system not only takes user's usage history into consideration, but experiences of other web service users are also considered. Experience of other web service users can be used to predict the potential interest of the active user. Collaborative filtering approach is used to predict the potential interest of active user. In collaborative filtering approach, user similarity is calculated based on the web service invocation records of a set of users. Similar users share the common interests, so likely to use the web services with same functionality. The more commonly invoked web services two users have in their invocation records, the larger the user similarity between them.

The user similarity is computed by,

$$User\ Sim(u_i, u_j) = \frac{2X|cs_{ij}|}{|s_{u_i}| + |s_{u_j}|}$$

Where $s_{u_i}$and $s_{u_j}$are the sets of web services used by user ui and uj respectively, cs_{ij} is the set of web services used by both user $s_{u_i}$and s_{u_j}. If $|cs_{ij}| = 0$, then user Sim(ui,uj) = 0. In collaborative filtering approach, web services used by similar users are recommended to the active user.

STEP III: Apply Hierarchical clustering

Clusters of web services which are closely related and similar to each other are created from historical user interest and potential user interest.

C. *Clustering Algorithm*

Agglomerative Hierarchical clustering

Nearest distance as well as pair-wise distance calculation between the data points are determined to cluster each data.There is question arises such as distance between the data point is recalculated but which distance is considered and when the groups has been generated. Number of solutions is available to overcome this question. Some as follows: complete farthest distance or complete linkage, single nearest distance or single linkage, average distance or average linkage, centroid distance. This process is continuous till single cluster being generated.

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. 4, Issue 5, May 2016

Algorithm 2: Hierarchical Clustering Algorithm-

Input: Data vectors $X_{n_{n=1}^{N}}$ =1, group wise distance DIST (G, G')

Output: Cluster of web services
Process:

1) $A \leftarrow \phi$ // Active set starts out empty.
2) For $n \leftarrow 1...N$ do // Loop over the data.
3) $A \leftarrow A \cup \{xn\}$ // add each datum as its own cluster.
4) end for
5) $\tau \leftarrow A$ A // Store the tree as a sequence of merges. In practice, pointers.
6) while |A| > 1 do // Loop until the active set only has one item.
7) $G_1^*, G_2^* \leftarrow$ min DIST (G, G')// choose pair in A with best distance.
8) $A \leftarrow (A \setminus G_1^*) \setminus G_2^*$ // Remove each from active set.
9) $A \leftarrow (A \cup G_1^* \cup G_2^*)$ // Add union to active set.
10) $\tau \leftarrow (\tau \cup G_1^* \cup G_2^*)$ // Add union to tree.
11) end while
12) Return: Tree τ

D. *Detail flow of proposed system*
Proposed System Flow:

Step 1: User inputs the query (Handling inputs).
Step 2: Usage history is checked and also check active user
Session history.
Step 3: Historical user interest is collected from similarity evaluation.
Step 4: Apply hierarchical clustering on web services.
Step 5: Internet is searched to get the QoS preferences and potential user interest.
Step 6: Potential interest evaluation is done through collaborative filtering approach.
Step 7: Clusters of web services are created from collaborative filtered web services.
Step 8:QoSof web services is evaluate. It is consider in evaluation of diversity of web services.
Step 9: Diversified web service ranking is calculated.
Step 10: Top-K diversified web service list is generated.

Algorithm 3: Diversified Web Service Ranking

Input: Web Service Graph $G = (V, E)$, parameter λ, adjacency matrix A.
Output: A set S of k ranked Web services.
1: $S = \emptyset$;
2: while $|S| \leq k$ do
3: $find\ v_{max} = \arg_{v \in (V-S)}^{max}(1-\lambda)Score_v + \frac{\lambda}{K}|N\{v\} - N\{S\}|$;
4: $S = S \cup \{v_{max}\}$;
5: end while
6: return S;

Diversified Web Service Ranking selects a node having maximum marginal gain $\rho vS = (1-\lambda)Score_v + \frac{\lambda}{K}|N\{v\} - N\{S\}|$ at every rotation which is line 3 in above given algorithm, as well as inserts it into the output set S which is line 4 in given algorithm. The above procedure from line 2 to 5 will executed reputedly for retrieving the top-k ranking list. This algorithm will create an ordered ranking list based on ρvS. Indeed the ρvS fulfills the non-decreasing

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. 4, Issue 5, May 2016

properties; algorithm presented is in charge of ranking procedure in which node having a maximum ranking score will be shown in top ranking list [1].

E. *Experimental Setup*

This system developed on Java framework (version jdk 6) and Netbeans (version 6.9) is used as a development tool on Windows platform. System is able to run on common machine also it does not require any specific hardware.

V. RESULT AND DISCUSSION

A.Dataset

This system experimented over real-world Web Service Dataset. WS-DREAM engine is aggregate commonly available WSDL file on the Internet. Web services are analyzed by different distributed nodes by implementing aggregated QoS data. Normally previous work of systems is used real-world dataset to calculate performance by on web service.

B.Outcome

Figure 2 shows the precision graph of the proposed and present system. We can see by making use of clustering we have achieved higher precision while the previous system in which they have not made use of clustering are not more precise.

Figure 3 shows the Mean Absolute Error Graph which show the comparison in the systems i.e. the system with clustering and system without clustering. Due to the clustering the MAE rate of our system is reduced and it produces results faster than the other system.

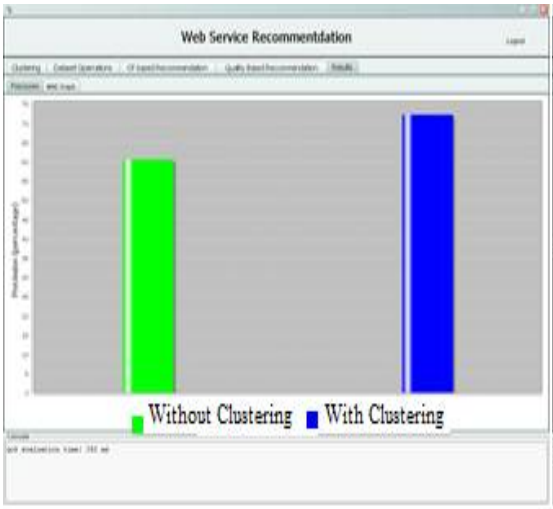

Fig2.Precision Graph

Fig 3. Mean Absolute Error Graph

VI. CONCLUSION AND FUTURE WORK

Here in this paper, we have maximized web service recommendation efficiency by qualities, QoS utility, useful association and the heterogeneous ranking analysis. Different problems present in Web Service recommendation system for example Collaborative Filtering problem specifically cold-start drawback, gray sheep drawback, word drawback, ramp up drawback, shillings attack, data integrity and quantifiability. Also, it gains maximized accuracy precision-recall, outcome accuracy and makes sure of retrieval time. This paper demonstrates QoS-based web service selection, ranking algorithm having trust and reputation administration support. Web Services and WSDL files are clustered by using hierarchical clustering algorithm in proposed system and after which search out the top k web service recommendation depending on user interest and web services. Clustering is created for similarity computation

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. 4, Issue 5, May 2016

of web services for better efficiency. Ranking computation basically concentrates on the immediate neighborhood data of 'S' in web service graphs. Thus, further system requires implementing different ranking computations along with k-hop nearest neighbors.

In a future work, we can use social network history of present user's for recommendations for the new user which will increase the scope of system.

REFERENCES

[1] Guosheng Kang, Mingdong Tang, Jianxun Liu, Xiaoqing (Frank) Liu, Buqing Cao, "Diversifying Web Service Recommendation Results via Exploring Service Usage History", IEEE TRANSACTIONS ON SERVICES COMPUTING on Volume: PP; Year 2015.
[2] Min Gong, ZhaoguiXu, Lei Xu, Yanhui Li, "Recommending Web Service Based on User Relationships and Preferences", Web Services (ICWS), 2013 IEEE 20th International Conference on Date of Conference:June 28 2013-July 3 2013.
[3] Yutu Liu, Anne H.H. Ngu, LiangzhaoZeng, "QoS Computation and Policing in Dynamic Web Service Selection", WWW2004, May 1722, 2004, New York, New York, USA. ACM 1-58113-912-8/04/0005.
[4] Shangguang Wang, Ching-Hsien Hsu, Zhongjun Liang, Qibo Sun, Fangchun Yang, "Multi-user web service selection based on multi-QoS prediction", Information Systems FrontiersMarch 2014, Volume 16, Issue 1, pp 143-152 First online: 04 October 2013.
[5] Wei Lo, Jianwei Yin, Shuiguang Deng, Ying Li, Zhaohui Wu, "Collaborative web service qos prediction with location-based regularization", doi.ieeecomputersociety.org/10.1109/ICWS.2012.49.
[6] Lin Wenmin, Dou Wanchum, LuoXiangfeng, Jinjun Chen, "A History Record-Based Service Optimization Method for QoS-Aware Service Composition", Web Services (ICWS), 2011 IEEE International Conference on Date of Conference:4-9 July 2011.
[7] R.-H. Li, J. X. Yu. "Scalable diversified ranking on large graphs". IEEE Transactions on Knowledge and Data Engineering, Vol. 25, No. 9, pp. 2133-2146, 2013.
[8] Mohammad Alrifai, DimitriosSkoutas, Thomas Risse, "Selecting Skyline Services for QoS-based Web Service Composition", WWW 2010, April 2630, 2010, Raleigh, North Carolina, USA.ACM 978-1-60558-799-8/10/04.

BIOGRAPHY

Ms.Pooja T. Kumbharbcurrently pursuing a ME degree in computer science technology from SavitribaiPhule Pune university. Completed B.E degree in computer engineering from PimpriChinchwad college of engineering under the SavitribaiPhuleuniversity,Pune June 2013.

Prof. Gitanjali S. Mate received M.E. degree in Information Technology from D.Y.Patil college of engineering, akurdi, Pune ,India. Currently she is working as Associate Professor in Dept. of Information Technology, RajarshiShahu College of Engineering, Pune, India. She is having 15 years of teaching experience in various engineering college under SPPU Pune. Her current research area is nano technology.