RESEARCH ARTICLE OPEN ACCESS

Web Service Recommendation System Incorporating Location Information and Personalized QoS Prediction

Nitin V. Tawar *, Supriya A. Kinariwala **
**(Department of Computer Science & Engineering, Marathwada Institute of Technology, Aurangabad, Maharashtra.*
*** (Department of Computer Science & Engineering, Marathwada Institute of Technology, Aurangabad, Maharashtra.*
Corresponding Author: Nitin V. Tawar

ABSTRACT
With the invent of new technologies and advancements in business, more and more organizations are working towards building automated systems which work over the internet throughout the world for communication purpose and the availability of huge number of web services for different requirements makes the task of business organizations. Due to presence of a wide variety of web services available for the user, it is a difficult task to choose a particular service satisfying the requirements. In this paper we have studied various previous works under recommending user with services which best suits the requirements so that he can select and use the best service available and we then propose a recommendation system, which presents the user with a list of recommended services as per his needs which are near to him physically as well as which satisfy security requirements.

Keywords: Collaborative filtering, QoS Prediction, Personalized influence, Recommendation, Similarity Computation

---

Date of Submission: 30-10-2017 Date of acceptance: 09-11-2017

I. INTRODUCTION

Web services are interoperable components built for internet communication. This communication takes place using XML messaging. Web services are self-contained components which are made for application-to-application interaction over the internet using XML messaging. The web service architecture is composed of SOAP, WSDL and UDDI. Due to wide variety of web services available over the internet and with the increase in the world wide acceptance of the web services, business organizations prefer to use existing web services so as to focus on the business process rather than working on developing a web service which fits into their requirements. If user chooses to build a service, it consumes resources and incurs cost on user as well as provider side. Selecting a high quality web service among the available web services for a particular task is again a tedious job because user has to select a service as per his needs[1]. The user can not test for each and every web service by invoking and recording whether it is performing as needed or not, because invoking such a huge number of services is not possible. Even if some user tries to evaluate the web service, it may happen that the best services are not available at that time or network condition may affect web service response. It brings the necessity of a system which provides you with the recommendations of web service requested based upon parameters like response time or some other mentioned by the user.

To increase the usability of a web service among the available set of web services, it is described by Quality of Service (QoS) parameters which are non functional characteristics of web services. QoS plays a vital role in the selection of high quality web services. These non-functional parameters include response time, throughput, usability, availability, reliability, etc[2]. It is a very difficult process to acquire all the QoS values of all the candidate web services due to network conditions at that instant and unavailability of that service at that time. Along with these issues services may give different QoS for different users. It necessitates the need of location parameter of the user to be considered in recommendation. If a user and web service lie in same location or region they may provide better QoS as compared to the ones lying far away.

Web service recommendation system is one in which high quality web services are selected on the basis of QoS parameters and location and are presented to the user as per his or her requirements. Recommendation systems have become very popular in recent years and are utilized in a variety of areas.

Recommendation system typically includes two kinds of results- through collaborative filtering and content based filtering. Collaborative filtering approaches make recommendations based on the past experiences of QoS parameters of the other users for a particular item requested. It finds similar items based on the similar QoS parameters. First missing values are predicted based on QoS similarity and user similarity as shown in Figure 1.

Some of the web services are available for use only in a particular area. It is irrelevant to use a WS in India that is operable in USA. Existing approaches consider users' past experiences i.e. QoS parameters on web services for web service recommendation. Such methods might provide poor recommendation due to not taking in consideration the parameters like users' locations and the accuracy could be very low. To enhance the prediction accuracy and improve the quality of recommendations we need to take locations into consideration [3].

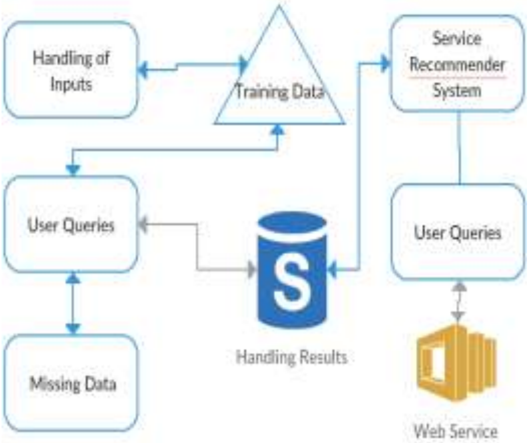

Figure 1: Basic Web Service Recommendation Process

II. LITERATURE SURVEY

2.1 COLLABORATIVE FILTERING

The overwhelming amount of data necessitates mechanisms for efficient information filtering. One of the techniques that are used is collaborative filtering. It is a technique used by the recommender systems to make predictions of missing values due to presence of many missing values in the users' past experiences and recommend potential items of interest to a user by finding similar user to that user. CF is based on user-item matrix. CF algorithms can be classified into two categories: memory based algorithms and model based algorithms.

Collaborative filtering methods have been very successful in recommending services or items to users by computing similarity between users or items or both.

Memory based algorithms include user based approaches, item based approaches or a combination of both. User based approaches compute similarity of the users by incorporating QoS values of users and predicting the missing values. Item based approaches predict the ratings of the users based on the similar QoS values as observed by items. It is simple and low cost techniques but does not scale well when number of user increase due to complex process of prediction computation. Pearson Correlation Coefficient, Cosine Similarity measures [3] are two of the most popular metrics for computing similarities between users and items.

Model based approach builds a model on whole user data using statistical, probabilistic and machine learning techniques such as clustering models, matrix factorization model, latent factor model, etc. It provides good recommendation performance although the models have to rebuild whenever new users or items are added which involves overhead.

2.2 WEB SERVICE RECOMMENDATION SYSTEM

Huge amount of efforts have been put into studying the recommendation of web services. Various web service recommendation algorithms have been implemented for web service recommendation like collaborative filtering [3],[4],[5], content based[6]. CF is most popular because of its simplicity in implementation and effectiveness. Shao [7] proposed a user-based CF method for QoS-aware Web service recommendation. Zheng [2], [5] combined both user-based and item-based CF algorithm to predict Web service QoS values and recommend similar services based on combined results. They wanted to employ both the similarity of active users and similar values of QoS of the targeted web service to predict the values but they failed to implement this approach. Zheng[2] proposed a recommender system to implement location based recommendation via exploiting QoS values.

This approach included creating user region based on the similar user location as that of active user and creating service region based similar values of QoS. Results were then combined to predict the QoS values for active user and recommendations were given to the active user. However this approach does not include location of web services.

Jieming Zhu, Pinjia He, Zibin Zheng, Michael R. Lyu[8] suggested a methodology to preserve privacy of the training users who provided their usage history and QoS values for recommendation purpose. The data obfuscation techniques were used to hide the identity of the users to preserve it from being revealed to the outside users.

Wancai Zhang, Hailong Sun, Xudong Liu [9] put forward the need of prediction of missing values on the basis of a new parameter that was to be time because QoS value of web service predicted could change with time. Temporal QoS-Aware Web Service Recommendation Framework was proposed to make the prediction of missing QoS values using user-item-time model. It used an algorithm namely Non-negative tensor factorization algorithm to deal with that value.

Recently, Matrix Factorization (MF) has been successfully employed for accurate and scalable Web service QoS prediction [10]. However, these model-based CF methods may have difficulties in handling dynamics of the user-service interaction matrix. When new interactions between users and services occur, the MF model has to be recomputed from scratch to perform QoS prediction. Therefore, this work focuses on improving memory-based CF by exploiting the characteristics of Web services and service users.

III. SYSTEM ARCHITECTURE OF WEB SERVICE RECOMMENDATION

The advent of more and more web services on the internet makes it a complex task for finding appropriate service as per requirements. To avail business organizations with a fast recommendation system of web services as per his or her needs, the demand for efficient evaluation of service is becoming strong. To facilitate user with ease of getting optimal web services on a single click, we define a web service recommendation system which makes use of both locations and personal QoS values provided by training users.

In this section we design a system which takes into account the user as well QoS experiences of past users to predict the values of the user and recommend the most rated web services to the active user. Existing QoS prediction methods do not take into account the personalized influence of the users and services while computing similarity between the users and services. We have taken this concept into consideration for proposing this system. The proposed method first selects the groups of users based on location parameters because we consider that the web services or which are lying physically close, they can provide better QoS experience rather than those which are far away from each other. In the next step we compute similarity between services and users using QoS values by considering personalized influence of individual services and user on the QoS data to predict missing QoS values and prepare and present the recommendations.

We consider the situations where an active user logs into the system and searches for high-quality Web service and the system recommends high-quality Web services to an active user. In these situations, first of all QoS values of web services which are unknown to the current user are predicted; then, web services with high QoS values are selected and recommended to the active user.

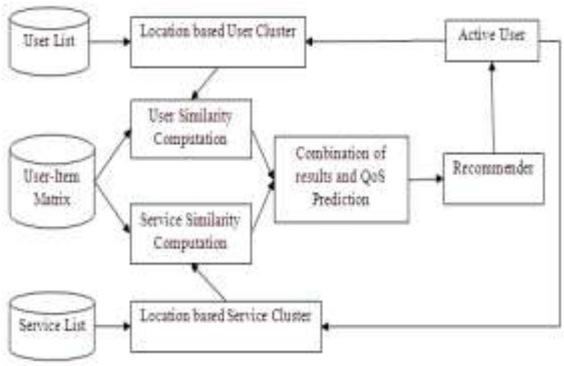

Figure 2: System Overview of Proposed Methodology

The main focus is to predict the QoS values for the missing values of the web services for the active user. As shown in Fig. 2, our system to recommend services consists of following modules:
1. The user log into the system and its details like IP address, country, ASN are obtained from the net package in java and registration database.
2. Based on the active user's ASN(Autonomous System), we find a group of users.
3. The user selects a service from service category and based on the category and ASN of active user, we find a group of services which are having same ASN as that of active user in the whole country.
4. The similarity between the active user and every user of selected group is computed using PCC and the pairs with high values of similarity coefficient are selected for further use.
5. In the same way, the similarity among the searched service and every service of the selected group of web services is computed and the top results are selected.
6. The values of selected users are used to predict the QoS values for the services of active user.
7. The values of selected services are to predict the QoS values for the services of active user.
8. The predictions done in the step 7 and 8 are combined to give the final set of predicted values.

IV. IMPLEMENTATION

In this section, an online service recommendation scenario is depicted to show the implementation of the problem domain of this paper. The basic idea of this approach is that users located close to each other are more likely to have similar service experience than those who stay far away from each other. We employ the idea of user-collaboration based on AS or country in our web service recommendation system. The more QoS information the user contributes, the more accurate service recommendations the user can obtain, since more user characteristics can be analyzed from the user contributed information. We are using response time of web service observed by users in our implementation. Based on the collected QoS data of response time, our recommendation approach is designed as a two-phase process. In the first phase, we divide the users into different groups based on their physical locations(AS or country) and historical QoS experience(response time)[15] on web services. In the second phase, we find similar users for the current user and make QoS prediction for the unused services. The predictions are done on the basis of both the user similarity as well as service similarity and results are combined to form the final set of predictions. In the end, services with the best predicted QoS will be recommended to the current user.

1. Location based user cluster: A group of users is selected on the basis of user location. Location of a user is represented in terms of IP address, Autonomous System Number (ASN) and Country. All the users are selected with the same ASN as that of the active user. If there is less number of users in ASN set then all the users belonging to the country of active user are collected and stored in a list. We don't cluster users by collaborating IP address due to repetition of IP addresses in AS which is in turn due to short range address availability of IPv4. We also avoid grouping users based on longitude and latitude because the users might be located close but on the internet they might be on different ASs [1]. Therefore we find country and ASN to be the best measure to group users.

2. Location based Service Cluster: The similar services are clustered into one cluster for the service searched by the active user on the basis of matching country as that of active user.

3. User Similarity Computation: The group of similar users formed is used to compute similarity between every pair of users that is the active user and each user from selected users. The neighbors with high value of similarity are selected. The similarity is computed uses weighted PCC [2],[3] because we consider the individual influence of users on the QoS values. The similarity is calculated using PCC as follows:

$$\boldsymbol{PCC(u,v)} = \frac{\sum_{i\in I_u\cap I_v}(\boldsymbol{r(u,i)} - \overline{\boldsymbol{r(u)}})(\boldsymbol{r(v,i)} - \overline{\boldsymbol{r(v)}})}{\sqrt{\sum_{i\in I_u\cap I_v}(\boldsymbol{r(u,i)} - \overline{\boldsymbol{r(u)}})^2}\sqrt{\sum_{i\in I_u\cap I_v}(\boldsymbol{r(v,i)} - \overline{\boldsymbol{r(v)}})^2}};$$

….1.1

where *u* is the active user and *v* is a user from the group of selected users, *i* is the set of services observed by user, *r(u,i)* is the response time of the service *i* observed by user u and $\overline{\boldsymbol{r(u)}}$ is the average response time of all the services observed by user *u*.

4. Web Service Similarity Computation: The web service cluster is used to find the similarity between the pair of services to predict and recommend the services based on weighted PCC incorporating personalized influence of web services on the QoS parameters. The web services with high similarity coefficients are selected.

$$\boldsymbol{PCC(i,j)} = \frac{\sum_{u\in U_i\cap U_j}(\boldsymbol{r(u,i)} - \overline{\boldsymbol{r(i)}})(\boldsymbol{r(v,i)} - \overline{\boldsymbol{r(j)}})}{\sqrt{\sum_{u\in U_i\cap U_j}(\boldsymbol{r(u,i)} - \overline{\boldsymbol{r(i)}})^2}\sqrt{\sum_{u\in U_i\cap U_j}(\boldsymbol{r(u,j)} - \overline{\boldsymbol{r(j)}})^2}}$$

….1.2

where *i* and *j* are the services from the group of selected services, *u* is the set of all training users available in the dataset, *r(u,i)* is the response time of the service *i* observed by user u and $\overline{\boldsymbol{r(i)}}$ is the average response time of the service *i* as observed by the user set *u*.

5. User based and Web Service based QoS Prediction:

The QoS values for the web services are missing for active user because the active user might not have invoked and experienced all the web services matching his requirements. Therefore it becomes necessary to predict the missing values before recommendations. This is done with the help of similarity values computed for services as well as users. All the values of QoS observed by all the user for a particular service are used to predict the missing values for that user and same technique is applied for the web services where all the QoS parameters of a particular service obtained by all the users is to predict the value for the active user's missing QoS values. Then the results of user and service based QoS prediction are combined to find the accurate predictions according to both user locations as well as personal influence [1] of web services.

Based upon the similarities computed between users and confidence of active user for PCC value in that AS or country and previous QoS values, we find new values of QoS for all services of active user by using following mathematical equation[1]:

$$\widehat{\boldsymbol{r_u}}(\boldsymbol{u,i}) = \frac{\sum_{v\in N(u)} \boldsymbol{conf(u,v)\ X\ PCC(u,v) X\ r(v,i)}}{\sum_{v\in N(u)} \boldsymbol{conf(u,v)\ X\ PCC(u,v)}}$$

….1.3

where *v* is the neighbor of *u* which belongs to the set of selected users with highest similarity value, $\widehat{r_u}(u,i)$ is the user based predicted value of service i for active user u, $PCC(u,v)$ is the similarity between active user *u* and *v*, $r(v,i)$ is the response time service *i* as observed by user *v* and *conf(u,v)* is the confidence of *PCC(u,v)* in the view of *u*.

Conf(u, v) is calculated as follows:

$$conf(u,v) = PCC(r(u), \bar{r}(U^a))$$

….1.4

where $r(u)$ is the response time of all the web services observed by user u and $\bar{r}(U^a)$ is the average value of response time of all the similar user in a particular ASN a.

In the same way, based upon the similarities computed between selected web services and confidence of searched service for PCC value in that AS or country and previous QoS values, we find new values of QoS for all services of active user by using following mathematical equation:

$$\hat{r_i}(u,i) = \frac{\sum_{j\in N(i)} conf(i,j)\ X\ PCC(i,j) X\ r(u,i)}{\sum_{j\in N(i)} conf(i,j)\ X\ PCC(i,j)}$$

….1.5

conf(i, j) is calculated as follows:

$$conf(i,j) = PCC(r(i), \bar{r}(I^a))$$

….1.6

6. Combination of results using HLACF:

We have obtained the predicted values for all services on the basis of similar users and also on the basis of similar services. These results are combined to calculate the values of all the services for the active user using following techniques:

$$\hat{r}(u,i) = \lambda\hat{r_i}(u,i) + (1-\lambda)\hat{r_u}(u,i)$$

….1.7

We have taken λ as 0.5. These predicted values are used to recommend the services to the user as per his search.

7. Web Service Recommendation: After having the predicted QoS values for missing parameters for active user, we can select the most promising web services by selecting the services with high quality QoS values and recommend these services to the user.

In this way a hybrid approach is used to predict QoS with high accuracy and recommend the potential services to the active user. The accuracy depends upon the personalized prediction of QoS values along with location taken into account.

V. INTEREST BASED RECOMMENDATION

To make system more interactive to the user, another module which can be added to the existing system as an enhancement is to predict and recommend the services based on the user interest. The user interests are recorded every time a search is made by the user for a web service. The user can be recommended a list of services matching all his interests stored the dataset as and when required.

Whenever user makes a search for a web service, the search text is saved in database along with user details. High interest is found from the count of a particular web service category recommends the web services to the users which match its ASN and interest.

This module is added as an enhancement to the existing system.

VI. PERFORMANCE ANALYSIS

We have conducted two experiments to evaluate the performance of our prediction method and time of recommendation procedure. All the experiments and implementation is done in J2EE using JSP and servlet. We have used a dataset of 130 users and 123 web services where there are 3 files associated with the dataset. The userlist contains all the details of training users such as user ID, IP address, country, ASN. Another file contains the list of 123 web services data such as web service ID, service Category, link, IP, country and ASN.

1. Prediction Accuracy Evaluation

The Mean Absolute Error (MAE) is often used in collaborative filtering methods to measure the prediction accuracy.

$$MAE = \frac{\sum_{u,i}|r(u,i) - \hat{r}(u,i)|}{N}$$

….1.8

It finds and sums the absolute difference between old value and predicted value of response time of web service and averages it.

Because different Web service QoS factors have distinct value ranges, we also used the Normalized Mean Absolute Error (NMAE) metric to measure the prediction accuracy.

$$NMAE = \frac{MAE}{\sum_{u,i} r(u,i)/N}$$

….1.9

A smaller NMAE value represents higher accuracy.

We have compared our method with several well-known prediction methods for collaborative filtering. These methods include user-based methods using PCC (UPCC) [14], item-based method using PCC (IPCC) [17], hybrid CF method WSRec [8] and location-aware method RegionKNN [29]. The performance comparison among various techniques is given in the following table 1.1 where rttmae means the MAE for response time and rttnmae represents NMAE for response time:

Table 1.1 Performance Comparison of various methods for MAE

Method	Metrics (rttmae)	Metrics (rttnmae)
UPCC	0.5972	1.046
ULACF	0.4162	0.7479
IPCC	0.8223	1.5836
ILACF	0.5620	1.0099
WSRec	0.5220	0.9384
Region-KNN	0.4905	0.8814
HLACF	0.1749	1.0

The comparison among various prediction methods for MAE and NMAE can be shown with the help of graphs as given below in figure 3 and figure 4 respectively:

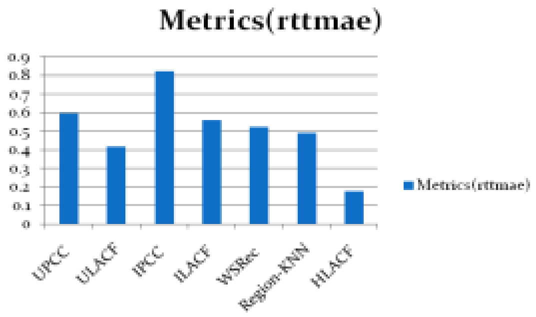

Figure 3: Performance comparison of various prediction methods for MAE

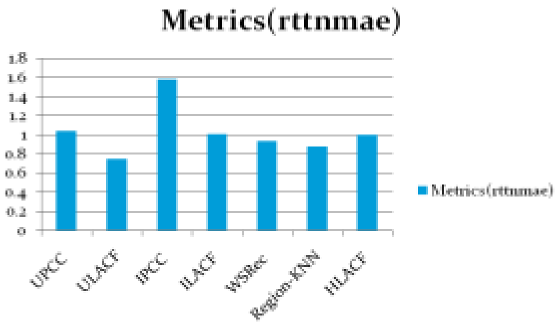

Figure 4: Performance comparison of various prediction methods for NMAE

2. HLACF Computing Time

HLACF (Hybrid Location Aware Collaborative Filtering) involves combining prediction values given both by user based and service based prediction methods. The computing time taken by this technique to predict the values for services and then selecting the optimal web services can be considered as a performance measure because users always choose a system which provides results very fast.

We find time required to make predictions and recommendations in NetBeans using profiler which runs live test and computes the time required by every invocation of method and JSP pages. The time varies due to network conditions. We have compared this time with many other previous techniques like UPCC, IPCC, WSRec, ULACF, HLACF, and Region-KNN etc. The results are listed below in the table 1.2 and the comparison can be graphically depicted for HLACF computing time in figure 5 below:

Table 1.2: Performance Comparison of various methods for HLACF Computing Time

Method	Computing Time
UPCC	0.0549
ULACF	0.0021
IPCC	0.0261
ILACF	0.0078
WSRec	0.0816
Region-KNN	0.0045
HLACF	0.136

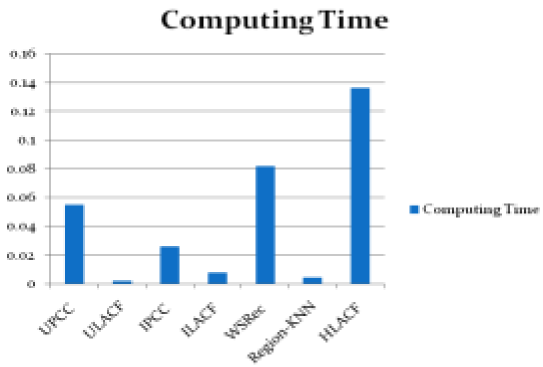

Figure 5: Performance Comparison of various methods for HLACF Computing Time

VII. CONCLUSION

With the aim of enhancing the prediction and recommendation performance, the basic idea is to predict Web service QoS values and recommend the best one for active users based on historical Web service QoS records. We combine prediction results generated from service clusters and user clusters, which achieves better results than existing approaches.

We also find that the combination result is much better than the result from any single method, either the prediction generated from user regions or the one generated from Web service regions. This is because these two methods analyze the problem from different aspects and the combination of them counteracts the error of individual methods.

We have also developed a technique which records user's interest every time a search is made. Then the user is recommended which matches the high interest of the user. It facilitates the user to have other options in front of him based on the interest.

REFERENCES

[1]. Jianxun Liu, Mingdong Tang, Member, IEEE, Zibin Zheng, Member, IEEE, Xiaoqing (Frank) Liu, Member, IEEE, Saixia Lyu, "Location-Aware and Personalized Collaborative Filtering for Web Service Recommendation", IEEE Transactions On

Services Computing, Vol. 9, No. 5, September/October 2016

[2]. X. Chen, Z. Zheng, Q. Yu, and M. R. Lyu, "Web service recommendation via exploiting location and QoS information", IEEE Trans. Parallel Distrib. Systems, vol. 25, no. 7, pp. 1913–1924, 2014.

[3]. L. Shao, J. Zhang, Y. Wei, J. Zhao, B. Xie, and H. Mei, "Personalized QoS prediction for Web services via collaborative filtering," in Proc. Fifth Int. Conf. Web Serv., pp. 439–446, 2007.

[4]. Z. Zheng, H. Ma, M. R. Lyu, and I. King. "WSRec: A collaborative filtering based web service recommendation system," in Proc. 7th Int. Conf. Web Serv., Los Angeles, CA, USA, pp. 437–444, 2009.

[5]. Z. Zheng, H. Ma, M. R. Lyu, and I. King "QoS-Aware web service recommendation by collaborative filtering", IEEE Trans. Service Computing, vol. 4, no. 2, pp. 140–152, Apr. 2011.

[6]. G. Kang, J. Liu, M. Tang, X. Liu, B. Cao, and Y. Xu, "AWSR: Active web service recommendation based on usage history," in Proc. Int. Conf. Web Serv., pp. 186–193, 2012.

[7]. X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, ''Personalized QoS-Aware Web Service Recommendation and Visualization,'' IEEE Trans. Serv. Comput., vol. 6, no. 1, pp. 35-47R1st Quart., 2013.

[8]. Jieming Zhu, Pinjia He, Zibin Zheng, Michael R. Lyu, "A Privacy-Preserving QoS Prediction Framework for Web Service Recommendation", IEEE Trans. Serv. Comput., 2 May, 2015

[9]. W. Zhang, H. Sun, X. Liu, and X. Guo, "Temporal QoS-aware web service recommendation via non-negative tensor factorization," in Proc. of the 23rd International World Wide Web Conference (WWW), 2014, pp. 585–596.

[10]. W. Lo, J. Yin, S. Deng, Y. Li, and Z. Wu, "An Extended Matrix Factorization Approach for QoS Prediction in Service Selection", in Proc. Int. Conf. Serv. Comput., Honlulu, Hawaii HI, USA, pp. 162–169, 2012.

Author Profile

Nitin V. Tawar is pursuing ME in Computer Science and Engineering from Marathwada Institute of Technology, Aurangabad, Maharashtra, India.

Supriya S. Kinariwala is working as a professor in Marathwada Institute of Technology, Aurangabad, Maharashtra, India since 2004.

International Journal of Engineering Research and Applications (IJERA) is **UGC approved** Journal with Sl. No. 4525, Journal no. 47088. Indexed in Cross Ref, Index Copernicus (ICV 80.82), NASA, Ads, Researcher Id Thomson Reuters, DOAJ.

Nitin V. Tawar, "Web Service Recommendation System Incorporating Location Information and Personalized QoS Prediction" International Journal of Engineering Research and Applications (IJERA) , vol. 7, no. 11, 2017, pp. 35-41.