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 4, April 2016

# Web Service Recommendation by Predicting QoS and User's Location

Ekta K. Chainani, Prof. Rajeshri R. Shelke

Student (M.E), Dept. of Computer Science and Engineering, HVPM's College of Engineering and Technology, Sant Gadge Baba Amravati University, Amravati, Maharashtra, India

Associate Professor, Dept. of Computer Science and Engineering, HVPM's College of Engineering and Technology, Sant Gadge Baba Amravati University, Amravati, Maharashtra, India

**ABSTRACT**: Web services are integrated software components for the support of interoperable machine to machine interaction over a network. Web services have been widely employed for building service-oriented applications in both industry and academia in recent years. The number of publicly available Web services is steadily increasing on the Internet. However, this proliferation makes it hard for a user to select a proper Web service among a large amount of service candidates. An inappropriate service selection may cause many problems to the resulting applications. In this paper, we propose a novel collaborative filtering-based Web service recommender system to help users select services with optimal Quality-of-Service (QoS) performance. Our recommender system employs the location information and QoS values to cluster users and services, and makes personalized service recommendation for users based on the clustering results. Compared with existing service recommendation methods, our approach achieves considerable improvement on the recommendation accuracy.

**KEYWORDS***:* Web Services Recommendation, QoS Prediction, Collaborative Filtering; Location Aware

## I. INTRODUCTION

The past several years have witnessed an increasing abundance of open services on the Web. This calls for more effective approaches for Web service discovery, which is a critical issue in services computing [1]. Recently, there has been an increasing interest in employing Collaborative filtering (CF) to recommend Web services to service users [3,4,5,7]. CF has been widely used in commercial recommendation systems [8].

CF algorithms can be divided into two categories: memory-based and model-based [9]. Depending on characterizing relationships between users or product items, Memory-based CF has two kinds of approaches: user-based approaches [9] and item-based approaches [11]. The user-based approach recommends to a user product items collected by other users sharing similar tastes; while the item-based approach recommends to a user those items similar to the ones the user preferred in the past. Since number of Web services with similar functionalities is rising, it's quite important to recommend services considering their non-functional properties. QoS of Web services [12], is mainly comprised of performance factors that include availability, response time, reliability, throughput, and etc. Values of these QoS factors are usually highly dependent on the network distance between services and service users, i.e. the locations of services and users, which are not fully incorporated in the existing CF recommendation algorithms. Motivated by this, we propose a location aware collaborative filtering method to predict missing QoS values and recommend Web services to users.

## II. RELATED WORK

Numerous research works have been done on collaborative filtering, recommendation systems and Web service discovery, here we only concentrate on QoS-aware Web service recommendation and QoS prediction.

The first work making QoS prediction using collaborative filtering technique was conducted by Shao et al. [3]. They proposed a user-based CF algorithm to predict QoS values. Zheng, et al. [4] proposed a hybrid user-based and item-based CF algorithm to recommend Web services, and carried out a series of large-scale experiments based on real

Web services dataset. They also developed an enhanced Pearson Correlation Coefficient (PCC) measurement for user similarity computation, which addressed the problem that PCC often overestimates similarities of users of services who are actually not similar but happen to have similar QoS experience on a few co-invoked Web services. users have similar QoS experiences on same services, i.e. the services they used have similar QoS values, the two users are similar. However, they failed to recognize and exploit the characteristics of QoS in similarity computation.

In order to improve their accuracy of QoS prediction of Web services, several new enhanced methods are proposed [5,10,11]. Jiang, et al. [10] proposed that the influence of personalization of Web service items should be taken into account when computing degree of similarity between users. Zhang et al. [11] suggested that it was better to combine users' QoS experiences, environment factor and user input factor to predict Web services QoS values. But how to obtain environment factor and user input factor were not discussed. Chen et al. [5] were the first to recognize the influence of user location in Web services QoS prediction and proposed a novel method.

Compared with prior related works, this paper makes the following major contributions. First, we consider not only locations of users but also of services, and present a hybrid location-aware QoS prediction method. We show that both user location and service location can be used in improving performance and accuracy of QoS prediction significantly. Second, we are the first to use large-scale real Web services QoS dataset to show that there exists a relationship between QoS similarity and user (service) location.

## III. OVERVIEW OF OUR METHOD

Fig. 1 is an overview of our Web service recommendation method. Suppose that the active user's interest is known, and a service list matching his functional interest is identified. We focus on predicting missing QoS values of the service candidates, which is critical in QoS based service recommendation.

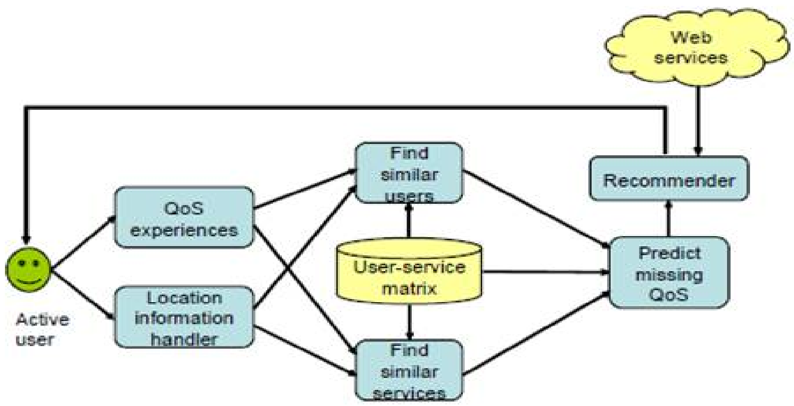


**Figure 1. Overview of our recommendation method**

In the method, we first acquire historical QoS data and the location information of the active user. A location information handler deals with the location information of both the active user and the target service whose QoS values are missing to the active user. The user-service matrix records every user's QoS experiences on Web services he invoked. To find similar users, user similarity measurement will be computed based on the historical QoS data of the users who are located close to the active user, determined by the location information handler. Likewise, services similarity measurement is computed based on the QoS records of the services which are located close to the target service, also determined by the location information handler.

After finding similar users and similar services for the active user and target service respectively, both user-based CF and item-based CF algorithm are used to predict the missing QoS values of the target service. We synthesize

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 4, April 2016**

them into a hybrid algorithm to improve the prediction accuracy. After predicted QoS values of all service candidates are computed, we recommend Web services with top QoS values to the active user.

## IV. LOCATION INFORMATION REPRESENTATION, ACQUISITION, AND PROCESSING

We represent a user' location as a triple (*IPu, ASNu, CountryIDu*), where *IPu* denotes IP address of the user host, *ASNu* denotes number of the Autonomous System (AS)1 that
*IPu* belongs to, and *CountryIDu* denotes the ID of the country that *IPu* belong to. Note that users are within a same AS does not definitely mean they are close geographically, and vice versa. Generally speaking, intra-AS traffic is much better than inter-AS traffic regarding transmission performance such as response time [6]. Therefore, even if two users are located in the same city, they may seem far away from each other in terms of network distance if their computers are within different ASs. That's one reason why we choose AS instead of state (province) or city to represent user location. Similarly, we model a Web service's location as (*IPs, ASNs, CountryIDs*).

For the location information of services, its acquisition is similar to users. Since the services' URLs or DSNs are already known, only a prior DSN name to IP address translation is required, which is also straightforward. The key issue in handling location information is how to measure user similarity or service similarity regarding their locations. In this aspect, our method is also different from the work by Chen et al. [5]. In their method, location similarity is computed based on IP similarity. That is, if two users have similar IP addresses, they are deemed as physically close. This seems to be reasonable but may cause inaccuracies in reality. Due to several factors, such as IPv4 address shortage and multi-homing, IP prefixes (i.e., IP address blocks assigned to networks) are constantly divided into finer granularities [14]. Therefore, two IP addresses with similar values do not necessarily belong to the same network or AS.

TABLE I. EXAMPLES OF IP TO ASN AND IP TO COUNTRY MAPPING

| Start IP Address | End IP Address | AS Number | Country Name |
|---|---|---|---|
| 4.56.0.0 | 4.67.63.255 | AS863 | Canada |
| 4.67.64.0 | 4.67.67.255 | AS9996 | Japan |
| 4.67.68.0 | 4.68.247.255 | AS863 | Canada |
| 4.68.248.0 | 4.68.249.31 | AS1148 | Netherlands |
| 4.68.294.32 | 4.71.36.3 | AS863 | Canada |
| 4.71.36.4 | 4.71.36.7 | AS1148 | Netherlands |

Table I shows some examples of real IP to ASN and country mapping. Each row of the table contains an IP address block, AS, and country it is assigned to. We can see that the IP addresses possessed by a network (e.g. AS863) are unnecessarily continuous, and the IP addressed with similar values are unnecessarily belonged to the same network or country (e.g. 4.67.68.0 and 4.67.64.0). In this situation, as a contrast, it's better to use ASN and country name to identify location-based similar users. If two users with different IP addresses are within a same AS, they are similar in terms of location; likewise, if two users with different IP addresses and different ASNs are within a same country, they are still similar. However, the latter should have a less similarity than the former. For the convenience of searching a user's or a service's location information and finding physically nearby users, we use a dada structure of hash tables to organize and store the location information of all users and services.

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 4, April 2016

## V. THE QOS PREDICTION ALGORITHM

### *A. Notations and Definitions*

*The following are important notations used in the rest of the paper:* $U = \{u_1, u_2, ..., u_m\}$ is a set of service users, where $m$ is total number of service users registered in the recommendation system.

$S = \{s_1, s_2, ..., s_n\}$ is a set of services, where $n$ is total number of services collected by the recommendation system.

$M\ \{r(u_i, s_j)\ |1 \le i \le m, 1 \le j \le n\}$ is the user-service matrix, where $r(u_i, s_j)$ is a vector of QoS attribute values acquired from $u_i$ invoking $s_j$. $r(u_i, s_j) = \varphi$ if $u_i$ has no experiences on $s_j$.

$A = \{a_1, a_2, ..., a_k\}$ is a set of ASNs, where $k$ is total number of ASs detected from all users and services. Let $U_{ai}$ denote the subset of users located in an AS with number $a_i$. Similarly, $a_i\ S$ denotes the subset of services located in an AS with number $a_i$.

$C = \{c_1, c_2, ..., c_l\}$ is a set of country IDs, where $l$ is total number of countries detected from all users and services. Let $U_{Ci}$ denote the subset of users located in a country with ID $C_i$ Similarly,
$C_i\ S$ denotes the subset of services located in a country with ID $C_i$.

$\bar{r}\ (U_{ai}, s) = avg\{r(u, s)\ |\ u \in U_{ai.} \wedge r(u, s) \neq \varphi$ is a vector of average QoS values of service $s$ observed by users in $U_{ai}$. .

Similarly, $\bar{r}\ (U, s)\ c_i$ is a vector of average QoS values of service $s$ observed by users in
$U_{ci}$ and $\bar{r}\ (U, s)$ (sometimes abbreviated as $\bar{r}\ (s)$ ) is a vector of average QoS values of service $s$ observed by all users.

### A. Overview of Algorithm

Here we give a high-level description of our QoS predicting algorithm. Let u be an active user and s be a target service, of which missing QoS values need to be predicted for u. Our QoS predicting algorithm comprises the following three sub-procedures.

1) User similarity computing and user-based QoS predicting. The steps are:

*Step 1: Search all users u' located in the same AS as u, and compute the similarity between each u and u' regarding their historical QoS experiences. If u has similar neighbors, goto Step 4.Step* 2: Similarly, search all users *u'* located in the same country as *u* and compute their similarity between *u'* and *u*. If *u* has similar neighbors, goto Step 4.

*Step* 3: For all users *u'* (irrespective of locations), compute the QoS similarity between *u* and *u'*. If *u* has similar neighbors, goto Step 4.

*Step* 4: Let *U'* denote the similar neighbors of *u*. If *U'* is not empty, based on their QoS experiences on *s*, predict the missing QoS values, which is denoted by $\hat{r}\ (u, s)\ u$ .

Otherwise, set $\hat{r}\ (u, s)\ u$ to be empty.

2) Service similarity computing and service-based QoS predicting. The steps are similar to above, as shown below.

*Step* 1: Find all services *s'* located in the same AS as *s*, and measure the similarity between each *s* and *s'* regarding their QoS records.If *s* has similar neighbors, goto Step 4.

*Step* 2: Similar to Step 1, find all services *s'* located in the same country as *s* and compute the similarity between *s'* and *s*. If *s* has similar neighbors, goto Step 4.

*Step* 3: For all services *s'* (irrespective of locations), compute the similarity between *s'* and *s* regarding their QoS records. If *s* has similar neighbors, goto Step 4.

*Step* 4: Let *S'* denote the similar neighbors of *s*. If *S'* is not empty, based on their QoS records concerned with *u*, predict the missing QoS values of *s*, denoted by $\hat{r}\ (u, s)\ s$. Otherwise, set $\hat{r}\ (u, s)\ s$ to be empty.

3) Integrating user-based QoS predicting and service based QoS predicting. The final predicting QoS values are computed as follows:

*Case* 1: If $\hat{r}\ (u, s)\ u$ and $\hat{r}\ (u, s)\ s$ are both not empty, synthesize them .

*Case* 2: If either $\hat{r}\ (u, s)\ u$ or $\hat{r}\ (u, s)\ s$ is empty, choose the one that is not empty as the result.

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 4, April 2016**

*Case* 3: Both r^ (*u*, *s*) *u* and r^ (*u*, *s*) *s* are empty. This may occur when the user-service matrix is sufficiently sparse or the active user has invoked few services or the target service has few users. The last two are so-called coldstart problem.

### A. Similarity Computation

Pearson Correlation Coefficient (PCC) is used in many recommendation systems to compute the similarity. User based collaborative filtering adopts PCC to compute similarity between two users *w* and *u*. Its computational formula is as Formula (1):

$$sim(w,u)=\frac{\sum_{s\in I}\left(r(w,s)-\bar{r}(w)\right)\left(r(u,s)-\bar{r}(u)\right)}{\sqrt{\sum_{s\in I}\left(r(w,s)-\bar{r}(w)\right)^2}\sqrt{\sum_{s\in I}\left(r(u,s)-\bar{r}(u)\right)^2}} \qquad (1)$$

where *sim*(*w*,*u*)denotes degree of similarity between user *w* and user *u*, *I* is the set of Web service items that are commonly invoked by user *w* and user *u*. *r* (*w*, *s*) and *r*(*u*, *s*) denotes the vector of QoS values which were produced when user *w* and user *u* invoke service *s* respectively. *r* (*w*) and *r* (*u*) represent the vector of average QoS values of user *w* and user *u* respectively. It can be seen from the Formula (1) that *sim*(*a*,*u*) is in the interval of [-1, 1]. The larger a value is, the more similar two users are.

As similar as the user-based collaborative filtering, Item based collaborative filtering adopts PCC to compute similarity between Web service *i* and *j*. The computational formula is as Formula (2):

$$sim(i,j)=\frac{\sum_{u\in V}\left(r(u,i)-\bar{r}(i)\right)\left(r(u,j)-\bar{r}(j)\right)}{\sqrt{\sum_{u\in V}\left(r(u,i)-\bar{r}(i)\right)^2}\sqrt{\left(r(u,j)-\bar{r}(j)\right)^2}} \qquad (2)$$

where *sim*(*i*, *j*) denotes degree of similarity between service *i* and *j* , *V* is a set of users that commonly invoke service *i* and *j* , *r* (*u*, *i*) and *r*(*u*, *j*) denote the vector of QoS values which were produced when user *u* invokes service *i* and *j* respectively. *r* (*i*) and *r* ( *j*) represent the vector of average QoS values of service *i* and *j* respectively.

## VI. ANALYSIS OF WEB SERVICE RECOMMENDATIONS

Personalized QoS value prediction for service users by employing the available past user experiences of Web services from different users. Our approach requires no additional Web service invocations. Based on the predicted QoS values of Web services, personalized QoS-aware Web service recommendations can be produced to help users select the optimal service among the functionally equivalent ones. we find that the user-observed Web service QoS performance has strong correlation to the locations of users. To enhance the prediction accuracy, we propose a location-aware Web service recommender system which employs both Web service QoS values and user locations for making personalized QoS prediction.

1. Reduce risk and deliver high-quality business processes
2. Web service QoS performance has Strong correlation to the locations of users.
3. User Collaboration
4. Location information is also considered when clustering users and services.
5. Service region and User region

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 4, April 2016

Web applications such as social networking sites and self-publishing sites [13], encourage users to share their knowledge and learn from others. LoRec employs the idea of user collaboration and provides a platform for users to share observed Web service QoS values and search Web services. This system will generate personalized service recommendations based on user shared QoS values. The more QoS records users contribute, the more accurate the recommendations will be, since more information can be mined from the user-contributed QoS values. In this paper, we assume that users are trustworthy. Fig.1 shows the architecture of LoRec recommender system, which includes the following procedures:

- Web service users log on to LoRec system and share observed Web service QoS records with other users. Users who have submitted Web service QoS records to LoRec are called training users. If a training user requires Web service recommendation, then the user becomes an active user. QoS values of training users will be employed to make personalized recommendation for the active user.
- LoRec clusters training users into different regions according to their physical locations and past Web service usage experiences.
- LoRec clusters functionally similar Web services based on their QoS similarities.
- LoRec maps the active user to a user region based on historical QoS and user location.
- The recommender system predicts QoS values of candidate Web services for the active user and recommends the best one.
- The active user receives the predicted QoS values of Web services as well as the recommendation results, which can be employed to assist decision making (e.g., service selection, service composition, service ranking, etc.).

## VII. PROPOSED SYSTEM

We proposed an enhanced measurement for computing QoS similarity between different users and between different services. The measurement takes into account the personalized deviation of Web services' QoS and users' QoS experiences, in order to improve the accuracy of similarity computation.
Although several CF-based Web service QoS prediction methods have been proposed in recent years, the performance still needs significant improvement we propose a location-aware personalized CF method for Web service recommendation.

The proposed method leverages both locations of users and Web services when selecting similar neighbors for the target user or service.To evaluate the performance of our proposed method, we conduct a set of comprehensive experiments using a real-world Web service dataset.
Based on the above enhanced similarity measurement, we proposed a location-aware CF-based Web service QoS prediction method for service recommendation. We conducted a set of comprehensive experiments employing a real-world Web service dataset, which demonstrated that the proposed Web service QoS prediction method significantly outperforms previous well-known methods.

## VIII. CONCLUSION

This paper presents a location-aware collaborative filtering method for QoS value prediction and QoS-based Web service recommendation. Based on our observation that a strong relationship exists between users' location closeness of and users' QoS similarity, we propose a location-aware UPCC method to identify similar users for an active user. Likewise, we propose a location-aware IPCC method to identify similar services for a target service. We integrate location-aware UPCC and IPCC into a hybrid collaborative filtering method. Through experiments on a large-scale real-world Web services dataset, we show that the performance of our method outperforms existing collaborative filtering based recommendation methods by a significant improvement of both prediction effectiveness and efficiency. In future work, we will take relationships among QoS factors into consideration and study how to incorporate them into QoS prediction.

## ACKNOWLEDGMENT

My thanks to the Guide, Prof. R.R.Shelke, Head of Department Dr. A. B. Raut and Principal Dr.A.B.Marathe, who provided me constructive and positive feedback during the preparation of this paper.

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 4, April 2016**

## REFERENCES

[1] L.-J. Zhang, J. Zhang, and H. Cai, "Services computing", Springer and Tsinghua University Press, 2007.
[2] T. Yu, Y. Zhang, and K.-J. Lin, "Efficient algorithms for web services selection with end-to-end qos constraints," ACM Transactions on Web, 1(1):6, 2007.
[3] L. Shao, J. Zhang, Y. Wei, J. Zhao, B. Xie, and H. Mei, "Personalized QoS prediction for Web services via collaborative filtering," Proc. IEEE International Conference on Web Services (ICWS 07), 2007, pp. 439-446.
[4] Z. Zheng, H. Ma, M.R. Lyu, and I.King, "WSRec: a collaborative filtering based web service recommendation system," Proc. IEEE International Conference on Web Services (ICWS 09), 2009, pp. 437- 444.
[5] X. Chen, X. Liu, Z. Huang, and H. Sun, "RegionKNN: a scalable hybrid collaborative filtering algorithm for personalized Web service recommendation,"Proc. IEEE International Conference on Web Services (ICWS 10), 2010, pp. 9-16.
[6] E. Rich, "User modeling via stereotypes," Cognitive Science, Vol.3, No.4, 1979, pp. 329-354.
[7] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," IEEE Internet Computing, 2003, pp. 76 - 80.
[8] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," Proc. 14th Conference on Uncertainty in Artificial Intelligence (UAI 98), 1998,
pp. 43-52.
[9] Y. Jiang, J. Liu, M. Tang, and X. F. Liu, "An effective Web service recommendation method based on personalized collaborative filtering," Proc. IEEE International Conference on Web Services (ICWS 11), 2011, pp. 211 - 218.
[10] L. Zhang, B. Zhang, Y. Liu, Y. Gao, Z. Zhu, "A Web service QoS prediction approach based on collaborative filtering," Proc. IEEE Asia-Pacific Services Computing Conference, 2010, pp.725-731.
[11] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of net news," Proc. ACM Conference on Computer Supported Cooperative Work (CSCW 94), 1994, pp.175-186.
[12] R. R. Shelke, Dr. R. V. Dharaskar, Dr. V. M. Thakare, "Data Mining For Mobile Devices Using Web Services", International Conference on Industrial automation And Computing ( ICIAC - 12th & 13th April 2014) , Jhulelal Institute of Technology, Nagpur.
[13] R. R. Shelke ,Dr. V. M. Thakare, Dr. R . V. Dharaskar, **"Study** of Data Mining Approach for Mobile Computing Environment**",** International Journal on Computer Science and Engineering (IJCSE) , ISSN : 0975-3397, Vol. 4, 12 Dec 2012 , pp.1920-1923

## BIOGRAPHY

1) **Ekta K. Chainani** is currently pursuing Master's Degree in Computer Science and Engineering from H.V.P.M's College of Engineering And Technology, Amravati.
   She received the B.E. degree in Computer Science and Engineering from H.V.P.M's College Of Engineering And Technology, Amravati in 2014.
2) **Prof. Rajeshri R. Shelke** is an Associate Professor at H.V.P.M's college of Engineering and Technology, Amravati. She received the B.E. and M.E degree in Computer Science. Her field of specialization is Data Mining. She is currently pursuing PHD and working as