Web Service QoS Prediction Approach in Mobile Internet Environments

Lubao Wang, Qibo Sun, Shangguang Wang*, You Ma, Jinliang Xu, Jinglin Li
State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Haidian, Beijing, 10086 China
imbarainbao@gmail.com; {qbsun; sgwang; youma; jlxu; jlin}@bupt.edu.cn

Abstract—**Existing many Web service QoS prediction approaches are very accurate in Internet environments, however they cannot provide accurate prediction values in Mobile Internet environments since QoS values of Web services have great volatility. In this paper, we propose an accurate Web service QoS prediction approach by weakening the volatility of QoS data from Web services in Mobile Internet environments. This approach contains three process, i.e., QoS preprocessing, user similarity computing, and QoS predicting. We have implemented our proposed approach with experiment based on real world and synthetic datasets. The results show that our approach outperforms other approaches in Mobile Internet environments.**

Keywords- *W**eb services; QoS; collaborative filtering; correlation coefficient***

I. INTRODUCTION

With the rapid development of Mobile Internet, a large number of Web services had emerged. Then it is very difficult to select suitable Web services from these services with the same function but different Quality of Services (QoS). Users want to know which services is better, especially QoS. Hence, how to accurately predict the QoS values of each Web services before users use these services is a very important issue.[1]

If you want to accurately choose the most appropriate Web service to meet user demand. You need to make an accurate prediction for QoS, and choose the best one. But, the existing approaches will have large prediction error when using these approaches to predict QoS values in Mobile Internet environments. The main reason is that compared to the traditional Internet, QoS values of Web services from Mobile Internet have greater volatility.

II. OUR APPROACH

In order to avoid the volatility of QoS values, we proposed a Web service QoS prediction approach (called WSQP) by weakening the volatility of QoS data from Web services in Mobile Internet environments. This approach first uses a preprocessing strategy to reduce the volatility of QoS values. Then we adopt Pearson Correlation Coefficient (PCC) to find similar users. Finally, we predict the QoS values by using normal QoS data that can be obtain from all similar users history QoS data.

A. QoS preprocessing strategy

When Web services runs in Mobile Internet, Packet loss, delay, retransmission phenomenon is a common occurrence. So QoS values of the Mobile Internet have greater volatility, and there are many normal values and abnormal values in history QoS data. Although the number of abnormal values is very small, they often degrade the accuracy of QoS prediction for Web services. Hence, it is very essential to weaken the abnormal QoS values of Web services.

In this paper, we assume $q^t_{a,j}$ represents the history QoS value of user *a* repeatedly invokes Web service *j*(*j*=1,2,3,...) for the *t*(*t*=1,2,3,...) time .Then the history QoS data set of user *a* repeatedly *t* times invokes Web service *j* is $Q^{history}_{a,j} = \{q^1_{a,j}, q^2_{a,j}, q^3_{a,j}, \cdots, q^t_{a,j}\}$. In order to weak the volatility of QoS data from Web services in Mobile Internet environments, a strategy is used as follows:

$$p^t_{a,j} = \ln(q^t_{a,j}) \tag{1}$$

Where $p^t_{a,j}$ represents the preprocessed result of the history QoS data $q^t_{a,j}$; *t* represents the number of repeated times. The history QoS data set of user *a* repeatedly *t* times invokes Web service *j* can be transformed a set weakens the volatility, i.e., $Q_{a,j} = \{p^1_{a,j}, p^2_{a,j}, p^3_{a,j}, \cdots, p^t_{a,j}\}$.

B. User similarity computation

Pearson Correlation Coefficient (PCC) has been introduced in a number of recommender systems for similarity computation, since it can be easily implemented and can achieve high accuracy. PCC is employed to calculate the similarity between two service users *a* and *u* using the following equation:

$$sim_{a,u} = \frac{1}{n}\sum_{l=1}^{n}\frac{\frac{1}{t}\sum_{k=1}^{t}(p^k{}_{a,l} - E_{a,l})(p^k{}_{u,l} - E_{u,l})}{\sqrt{D_{a,l}}\sqrt{D_{u,l}}} \tag{2}$$

Where $sim_{a,u}$ represents the similarity between two service users *a* and *u* ; $E_{a,l}$($E_{a,l} = \frac{1}{t}\sum_{k=1}^{t} p^k{}_{a,l}$) represents the average value of $Q_{a,l}$; $E_{u,l}$($E_{u,l} = \frac{1}{t}\sum_{k=1}^{t} p^k{}_{a,l}$) represents the average value of $Q_{u,l}$; $D_{a,l}$($D_{a,l} = \frac{1}{t}\sum_{k=1}^{t}(p^k{}_{a,l} - E_{a,l})^2$) represents the variance of $Q_{a,l}$; $D_{u,l}$($D_{u,l} = \frac{1}{t}\sum_{k=1}^{t}(p^k{}_{u,l} - E_{u,l})^2$) represents the variance of $Q_{u,l}$.

Although PCC can provide accurate similarity computation, it will overestimate the similarities of users who are actually not similar but happen to have similar QoS experience on a few common invoked Web services [2]. To address this problem, we employ a significance weight to reduce the influence of a small number of similar common invoked services. An enhanced PCC for the similarity computation between different users is defined as

$$sim'_{a,u} = \frac{|US_a \cap US_u|}{\sqrt{|US_a| \times |US_u|}} sim_{a,u} \tag{3}$$

Where $sim'_{a,u}$ represents the new similarity value; represents the number of Web service items that are employed by both the two users; $|US_a|$ and are the number of Web services invoked by user a and user u respectively.

C. QoS prediction

Based on the similarity between every two users, if we want to predict the QoS value of user a invoke Web service j, we need to find the most similar users with user a and select a normal values interval based on the characteristics of the QoS data. Then using the following equations:

$$F(u_c, j) = \frac{1}{v}\sum_{e=1}^{v} q^{e}_{u_c,j} \tag{4}$$

$$F_{wight} \frac{\sum_{c=1}^{K} sim'_{a,u_c}(F(u_c, j) - \frac{1}{K}\sum_{c=1}^{K} F(u_c, j))}{\sum_{c=1}^{n} sim'_{a,u_c}} \tag{5}$$

$$Forecast = \frac{1}{K}\sum_{c=1}^{K} F(u_c, j) + F_{wight} \tag{6}$$

Where Forecast represents the prediction QoS value of user a access Web service j; $q^{e}_{u_c,j}$ represents a QoS value from $Q_{u_c,j}$ which is in this normal values interval; v represents the number of QoS values which are in the normal values interval; K represents the number of most similar users with user a; $u_c (1 \le c \le K)$ represents a user who is similar with user a.

The normal values interval which is selected by the characteristics of QoS data only conclude the normal data. For example, the data of response time tends to become large when affected, and the data of throughput tends to become smaller when affected. So for the response time data, the normal values interval should be $(p_{\min}, E_{u_c,j})$ ($E_{u_c,j}$ represents the average value of $Q_{u_c,j}$,i.e., $E_{u_c,j} = \frac{1}{t}\sum_{k=1}^{t} p^{k}_{u_c,j}$), for the throughput, the normal values interval should be... $p_{\min}$ is the smallest value in $Q_{u_c,j}$ which is the result QoS set of preprocessed the history QoS set of user u_c repeatedly t times access Web service j. $p_{\max}$ is the largest value in $Q_{u_c,j}$.

III. EXPERIMENTS

We implement our approach [1] and conduct experiments using a datasets [2] named WSDream. It contains nearly 1million service response time records. We conduct experiments to compare WSQP (our approach) against WSRec[3] and AVG (It`s based on the average of history QoS data) in terms of response time.

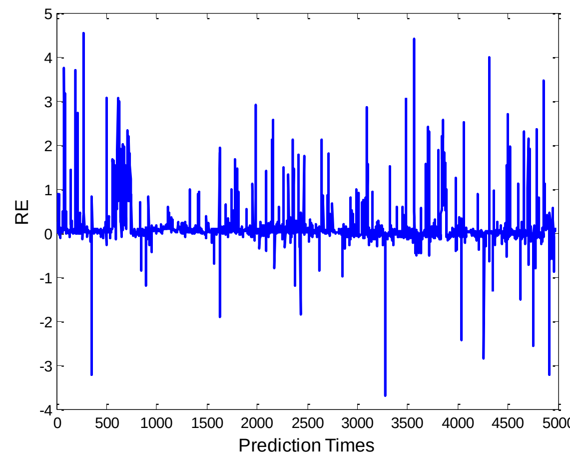
A. Comparison results on relative error

In this section, we perform experiments to compare WSQP with other approaches in terms of relative error (RE) where RE can be calculated as follows:

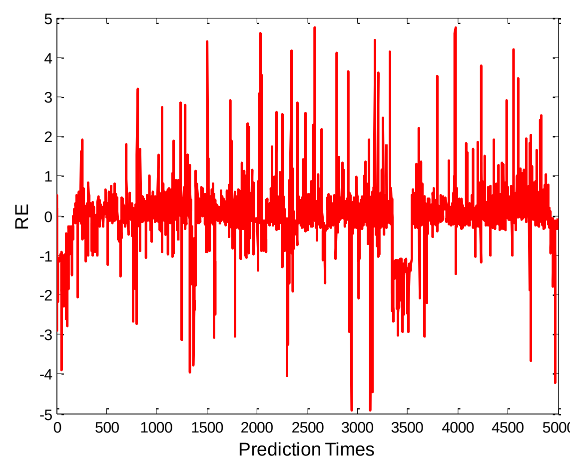
$$RE = F - F_{real} \tag{7}$$

Where F is the predicted QoS value; F_{real} is the real value which we had deleted before predicting QoS value; the smaller the absolute value of RE is, the more accurate the prediction is.

As shown in Fig.1 (a), Fig.1 (b) and Fig.1 (c), we find that comparing with WSRec and AVG, our approach is the best.



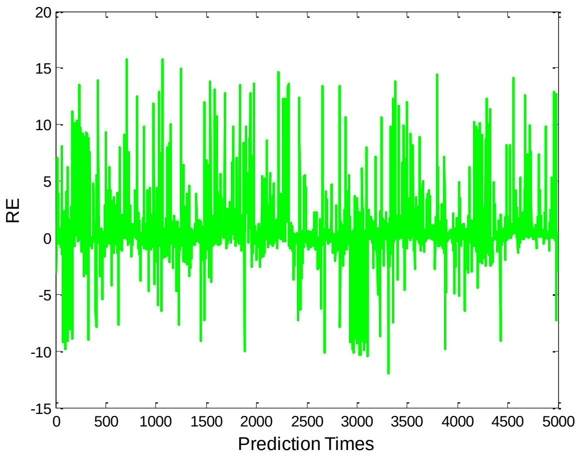
(a)WSQP



(b)WSRec

[1] http://www.sguangwang.com/Source code/wang-demo.avi

[2] http://www.wsdream.net



(c)AVG

Figure 1. Comparison WSQP with WSRec and AVG

B. *Comparision results on mean absolute error*

We use Mean Absolute Error (MAE) to evaluate the prediction quality .MAE can be calculated as follows:

$$MAE = \frac{\sum_{a,j} \left| F^{a,j} - F^{a,j}{}_{real} \right|}{N} \quad (8)$$

Where $F^{a,j}$ represents the predicted QoS value of user *a* invoked Web service *j*; $F^{a,j}{}_{real}$ is the real value of user a in voke Web service *j* which we had deleted before predicting QoS value; *N* is the number of predicted values. The smaller the absolute value of MAE is, the more accurate the predictio n is.

As shown in Fig.2 (N=100), we find that the MAE value of WSQP is the smallest in all approaches. These means our approach can significantly improve the accuracy of QoS pre diction for Web services in Mobile Internet environments.

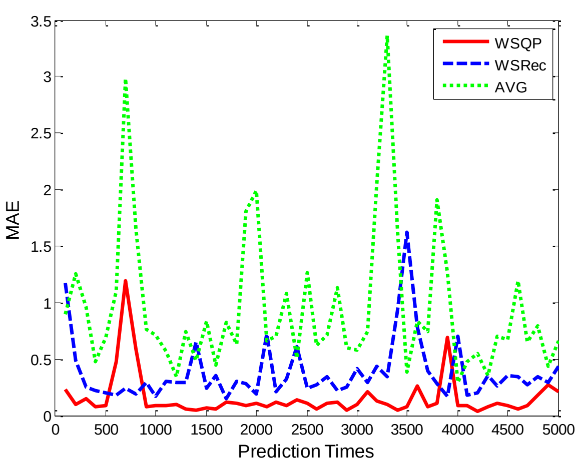


Figure 2. Comparision results on MAE

IV. Conclusions

In this paper, we presented an easy and accuracy QoS prediction approach for Web services in Mobile Internet environments. The experimental results show that our approach obtains more accuracy QoS prediction than other approaches.

Acknowledgements

The work presented in this study is supported by the Natural Science Foundation of Beijing under Grant No.4132048, NSFC(61472047), and NSFC(61202435).

References

[1] S. G. Wang, Q. B. Sun, and F. C. Yang, "Towards Web Service selection based on QoS estimation," International Journal of Web and Grid Services, 4, vol. 6, pp. 424-443, 2010.

[2] M.R. McLaughlin and J.L. Herlocker, “A Collaborative Filtering Algorithm and Evaluation Metric that Accurately Model the User Experience,” Proc. 27th Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR ’04), pp. 329-336, 2004.

[3] Z. Zheng, H. Ma, M.R. Lyu, and I. King, “QoS-Aware Web Service Recommendation by Collaborative Filtering,” IEEE Transactions Service Computing, vol. 4, no. 2, pp. 140-152, 2011.