

# QoS Prediction Based on Context-QoS Association Mining

Yang Hu<sup>1,2(✉)</sup>, Qibo Sun<sup>1</sup>, and Jinglin Li<sup>1</sup>

<sup>1</sup> State Key Laboratory of Networking and Switching Technology,  
Beijing University of Posts and Telecommunications, Beijing, China  
{hu.yang, qbsun, jlli}@bupt.edu.cn

<sup>2</sup> Science and Technology on Information Transmission and Dissemination  
in Communication Networks Laboratory, Beijing, China

**Abstract.** Both the functional properties and the non-functional properties (known as quality of service, QoS) should be considered when recommending services. In mobile environment, the services are referred to under different contexts. Context-aware collaborative filtering is often employed to predict the QoS of candidate services when a user submits a service request. Existing collaborative filtering based methods seldom analyze the actual impact of each context property on the QoS properties in context-similarity mining. To address the problem, we propose a QoS prediction method based on the context-QoS association mining in this paper. The method is composed of two steps. We firstly propose an algorithm to mine the association between the context and the QoS properties. Then, a QoS prediction approach is proposed by taking the association and the context similarity into consideration. To study the effectiveness of our approach, we experiment on real world dataset. Experimental results show that our proposed method can improve the accuracy of QoS prediction.

**Keywords:** Web services · Mobile service · QoS · Context

## 1 Introduction

With the increase of the Web service providers, more and more Web services continue to be published. Therefore, many services could meet the functional requirements of the user. As a result, the user tends to prefer the Web service with the best Quality of Service performance [1]. QoS has become the key factor in service selection and service recommendation [2].

Collaborative filtering is widely used in QoS prediction by many Web services recommendation methods [3, 4]. Collaborative filtering is a very effectiveness mechanism in selecting the service with the optimal QoS values. In early collaborative filtering-based approaches, the contexts when the service is invoked are rarely taken care of.

In mobile environment, the services are referred to under different context. The QoS performance varies when the mobile Web service is invoked under different context. The context properties should be considered in QoS prediction for a specific

user [5, 6]. Context-aware collaborative filtering-based methods are proposed to address the problem [7]. These web service QoS prediction methods work by collecting user observed QoS records and matching together users who invoke the web services under similar context. In others words, the users who refer to the service under similar context would have similar QoS experiences. Therefore, the QoS values are predicted based on the QoS data of the users with similar context.

However, current context-aware collaborative filtering-based methods assume that the QoS values are affected by many context properties [8]. Actually, some context properties may have strong correlation to the Web service QoS performance, while other context properties may have weak or no correlation to QoS performance. However, current methods are lack of accurate association analysis.

To address the problem, we propose a context-aware QoS prediction approach based on context-QoS association mining (QPAM). Our approach mainly consists of two steps. Firstly, we mine the association between the context properties and the QoS properties as a fundamental step for further work of QoS prediction. Secondly, we predict the QoS values of candidate service for current user based on the results of step 1. We filter out the irrelevant context properties so that the accuracy of prediction would not be affected. We experiment on real world dataset to study the effectiveness of our method. Experimental results show that our proposed method can improve the accuracy of QoS prediction.

The main contribution of this paper includes: (1) context-QoS mining to get the relevant context properties and filter out the irrelevant context properties. (2) a context-aware QoS prediction approach based on Context-QoS association mining to improve accuracy of QoS prediction.

The rest of paper is organized as follows. Section 2 introduces the related work. Section 3 defines some notations and describes the overview of our method. Section 4 describes the technical detail of our method. Section 5 shows the experimental results. Section 6 concludes the whole paper and presents future work finally.

## 2 Related Work

Considering that there are plenty of functionally equivalent candidates, the QoS performance plays a vital role in service selecting. A large number of QoS-aware methods has been employed in service recommendation [9]. Chen et al. [10] introduced a map containing visualized information regarding QoS among the recommended services, so that the service consumers can understand the service performance better.

The definitions of context and context-awareness are various. Dey [11] defined context as an information such as location, date, time, nearby things and nearby people and the situation of an entity is characterized by the information. The Web service QoS performance has strong correlation to some user contexts i.e. time interval, location. And they take the user contexts into account when predicting QoS values and selecting the optimal Web services. Chen et al. [12] proposed a LoRec (location-aware Web service recommender) system by employing the location information to improve the accuracy of the QoS prediction. Yu et al. [8] presented a time-aware Collaborative Filtering algorithm to predict the QoS values, by calculating the similarity between

service users and candidates at different time intervals. Tang et al. [13] proposed a collaborative filtering algorithm base on location awareness to recommend Web services to the current user. Yu presented that the picture conversion service has low QoS performance at the intervals during which a discount is provided [8]. Tang presented that weather forecast services has relationship with the location so it is helpful to recommend the current user services through users with similar locations [13].

The service recommendation methods based on context-awareness are usually combined with the CF (collaborative filtering)–based methods. CF finds similar users or items and recommend items by employing information from them to the current user [14]. Zheng et al. [3] presented a collaborative filtering method based on the past experiences of collaborative service users by employing data from similar service users who have similar experience with the current service user. Jiang et al. [15] described an algorithm combining personalized user-based and item-based collaborative filtering called PHCF (Personalized Hybrid Collaborative Filtering) to recommend the current service user the optimal Web services accurately.

The above methods only assume that the QoS values are affected by many context properties. Actually, some context properties may have strong correlation to the Web service QoS performance, while other context properties may have weak or no correlation to QoS performance. There, current methods are lack of accurate association analysis. We will address the problem in this paper.

### 3 Overview

#### 3.1 Problem Definition

**Definition 1: Service.**

*Service set*  $S = \{s_1, s_2, \dots, s_k\}$  is a set of functional equivalent mobile services that meets the need of the current user. The service set contains  $k$  services.

**Definition 2: Users.**

*User*  $U = \{u_1, u_2, \dots, u_l\}$  is a set of users that have invoked services at least once. The user set contains  $l$  users.

**Definition 3: QoS properties.**

*QoS properties*  $Q = \{q_1, q_2, \dots, q_m\}$  is a set of quality properties that is invoked by service users.  $Q$  contains  $m$  properties.

**Definition 4: Context properties.**

*Context*  $C = \{c_1, c_2, \dots, c_n\}$  is the set of context attributes of users that has invoked at least one services in  $S$ . The set contains  $n$  context properties. Context properties include location, time, season, temperature, network environment etc.

**Definition 5: Record.**

$R_{ij} = \{ \langle (q_1^{ij}, q_2^{ij}, \dots, q_m^{ij}), (c_1^{ij}, c_2^{ij}, \dots, c_n^{ij}) \rangle \}$  is the set of historical service invocation records.  $R_{ij}$  is the service invocation record of  $s_i$  is invoked by  $u_j$ .

### 3.2 Framework of Our Method

Figure 1 shows the framework of our method. Our method mainly consists of two phases. The first phase is context-QoS association mining. This phase consists of two sub-phases: QoS-based service clustering and association mining. The second phase is QoS prediction based on mining results. This phase also consists of two sub-phases: properties filtering and QoS prediction.

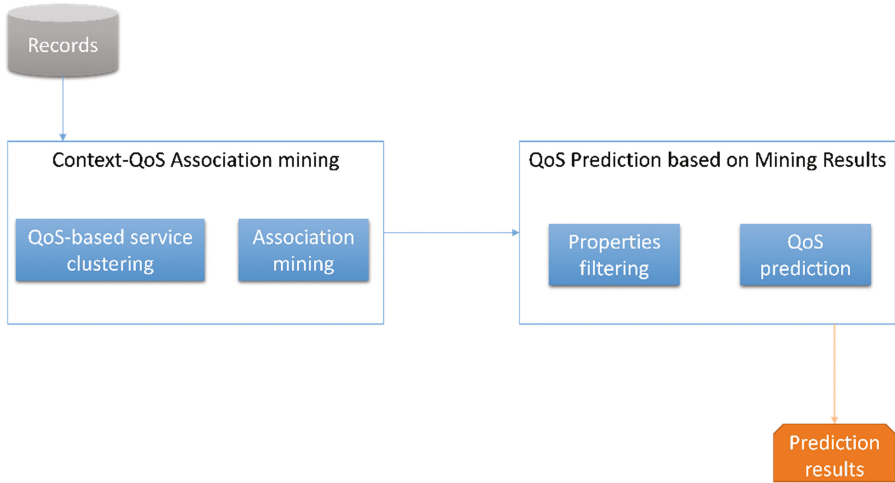


Fig. 1. Framework of our method

We will describe the technical detail of our proposed method in the next section.

## 4 Methodology

### 4.1 Context-QoS Association Mining

In this section, we analyze the association between each QoS property and each context property. Suppose  $R_i$  denotes all records that are related to service  $s_i$ . We now cluster the records based on QoS. To cluster the records, the distance between two records  $R_{i1}$  and  $R_{i2}$  on QoS property  $q_j$  can be calculated by the following:

$$Sim_{q_j}(R_{i1}, R_{i2}) = |R_{i1}.q_j - R_{i2}.q_j|. \tag{1}$$

Based on the calculation of distance between two records, we now mine the invocation records to obtain the context-QoS association. The algorithm is illustrated in Algorithm 1. Firstly, we cluster all records of a specific service  $s_i$  in the candidate services by using K-means clustering. Then, for each context property  $c_k$ , we calculate the overlapping degree between any two clusters  $g_1$  and  $g_2$ . The values of the context property  $c_k$  are divided into a group of interval partitions  $P = \{p_1, p_2, p_3, \dots\}$ . We now construct a vector  $V_{g_n} = \{v_{g_n}^1, v_{g_n}^2, \dots\}$  for each cluster  $g_n$ . The value of  $v_{g_n}^1$  is the

number of records  $R_{g_n}$  in  $g_n$  that the value of  $R_{g_n.c_k}$  belongs to  $p_1$ . We define the overlapping degree between any two clusters to measure the similarity between them. The overlapping degree is based on calculating vector similarity of two clusters. The overlapping degree between any two clusters  $g_1$  and  $g_2$  is calculated by the following:

$$degree(g_1, g_2) = \frac{V_{g1} * V_{g2}}{|V_{g1}| * |V_{g2}|} * e^{\varphi-1}, \tag{2}$$

$$V_{g1} * V_{g2} = \sum_i v_{g1}^i * v_{g2}^i, \tag{3}$$

$$|V_{g1}| * |V_{g2}| = \sqrt{\sum_i v_{g1}^{i2}} * \sqrt{\sum_i v_{g2}^{i2}}, \tag{4}$$

$$\varphi = \frac{||V_{g1}| - |V_{g2}||}{||V_{g1}| + |V_{g2}||}, \tag{5}$$

$$|V_{g1}| \pm |V_{g2}| = \sqrt{\sum_i v_{g1}^{i2}} \pm \sqrt{\sum_i v_{g2}^{i2}}, \tag{6}$$

$$\bar{d} = \sum_{m,n} \varpi_{mn} * degree(g_m, g_n), \tag{7}$$

$$\varpi_{mn} = \frac{\sum_i v_{gm}^i * \sum_i v_{gn}^i}{\sum_{m,n} \sum_i v_{gm}^i * \sum_i v_{gn}^i}. \tag{8}$$

After we calculate the overlapping degree between any two clusters. Then, we calculate the average overlapping degree  $\bar{d}$  of service  $S_i$  based on Eq. (7). The weight  $\varpi_{mn}$  of the overlapping degree between any two clusters is calculated in Eq. (8). If the average overlapping degree is smaller than  $\sigma$ , we consider the context property has impact on the QoS. Otherwise, the context property has no impact on the QoS.

Through context-QoS association mining, we can obtain  $CI = \{c_{i1}, c_{i2}, \dots, c_{imax}\}$ . As we known,  $CI$  is a set of context properties that has impact on QoS property  $q_i$ .

### 4.2 QoS Prediction

After obtain the context-QoS association, the context properties that have impact on QoS are utilized for service recommendation. Euclidean distance is adopted by us to calculate The distance between two records  $r_1$  and  $r_2$  based on context properties. The distance between  $r_1$  and  $r_2$  in max context property dimensions is calculated by the following:

$$sim(r_1, r_2) = \sqrt{\sum_{ci \in CI} (r_1.ci - r_2.ci)^2} \tag{9}$$

Based on the calculation of distance between two services, we cluster all candidate services by using K-means clustering to predict the QoS values.

The algorithm consists of the following steps:

- Step 1:** Construct a virtual record  $R_v$ . The  $CI_{\psi}$  values of  $R_v$  are assigned according to the context environment of current user. Add  $R_v$  to  $WSR$ .
- Step 2:** Select  $k_{\psi}$  records from  $WSR_{\psi}$  as the centroids.
- Step 3:** Assign each record in the set  $WSR_{\psi}$  to the nearest centroid.
- Step 4:**  $k_{\psi}$  new centroids are calculated based on the clusters obtained from the previous step.
- Step 5:** Repeat steps 2 and 3 until all sub-clusters are stable. In other words, all centroids no longer move.
- Step 6:** Search for the sub-cluster  $C$  which  $R_v$  belongs to.
- Step 7:** The average values on  $q_i$  of the records in  $C$  is the prediction results.

Through the above seven steps, we can predict the QoS values.

---

**Algorithm 1.** Association mining algorithm

---

**Input:** all the record list  $R$  on a specific service, a QoS property  $q_j$ , a context property  $c_k$

**Output:** the association between  $q_j$  and  $c_k$

---

```

1 For  $R$ , select init  $k$  records as centrioles and assign the records to  $tempCentrioList$ ;
2 while ! is stable do
3    $centrioList = tempCentrioList$ ;
4   for each  $r$  in  $R$  do
5      $bestCentri \rightarrow distance = \infty$ ;
6     for each  $centri$  in  $centrioList$  do
7        $distance = sim(r, centri)$ ;
8       if  $bestCentri \rightarrow distance > distance$  then
9          $bestCentri \rightarrow distance = distance$ ;
10         $bestCentri \rightarrow class = centri \rightarrow tag$ ;
11      end
12    end
13     $categoryList[bestCentri \rightarrow class].Add(r)$ ;
14  end
15  relocate centriole based on  $categoryList$ ,  $centrioList$ ;
16  assign the new centrioles to  $tempCentrioList$ ;
17 end
18 Calculate the overlapping degree between any two clusters;
19 Calculate the average degree;
20 If the  $average\ degree < \sigma$  then
21   return true;
22 end
23 else
24   return false;
25 end

```

---

## 5 Experiment

### 5.1 Experimental Setup

A set of experiments are conducted on the dataset WSDream published by Z. Zheng et al. Because we are the first to mine the context-QoS association, we compare our method with the RANDOM method. RADNOM refers to a set of methods that lack accurate association analysis when predicting the value of each QoS property based the context aware collaborative filtration, such as location-aware method [13] etc. Different from our method, RANDOM randomly decides whether a QoS property is affected by a context property. We use AE (the average error) to evaluate the QoS prediction accuracy. The definition of AE is demonstrated by the following:

$$AE = \frac{\sum_{i=0}^N |q_{prediction,i} - q_{real,i}|}{N}, \tag{10}$$

where  $q_{prediction}$  denotes the predicted value of the QoS property  $q_j$  and  $q_{real}$  denotes the actual value of the QoS attribute  $q_j$ .  $\sigma$  is set to 0.4 in our experiments.

### 5.2 Performance Evaluation

We randomly select 12 services from WSDream for response-time prediction and select 12 services from WSDream for throughput prediction. For each service, the related records store the response-time, the throughput, the time, and the interval ID etc. The records are divided into two sets: the training set and the test set.

Figure 2 shows the AE of the response-time prediction. Figure 3 shows the AE of the throughput prediction. It is evident that our method can reduce the prediction error based on mining the association between each QoS property and each context property.

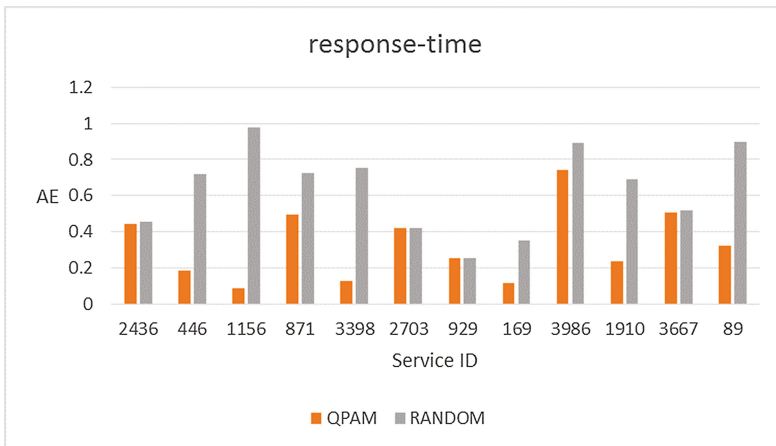
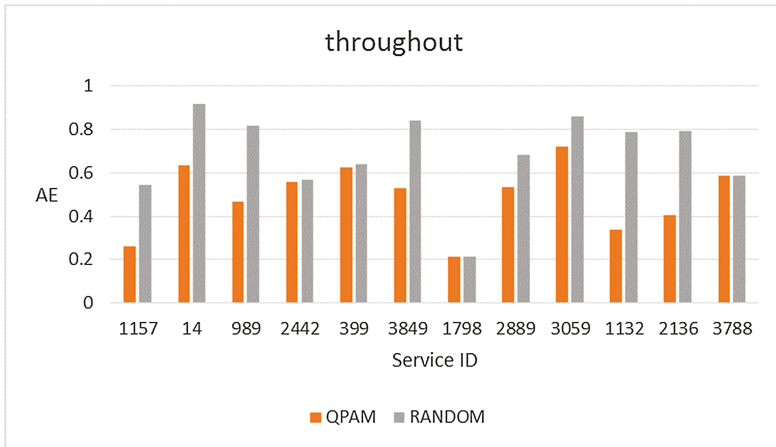


Fig. 2. AE of response-time prediction



**Fig. 3.** AE of throughput prediction

RANDOM randomly decides whether a QoS property is affected by a context property. There is a higher chance that an important context property is ignored or an irrelevant context property is taken into consideration. Therefore, our method show a higher accuracy.

## 6 Conclusion and Future Work

We proposed a mobile service recommendation approach based on context-QoS association mining. The approach employs a context-QoS association mining algorithm and a recommendation algorithm to improve the accuracy of context-aware QoS prediction. The experiment results on WS-Dream dataset show the proposed method improves predictive accuracy and outperforms the compared methods.

Our future work will include (1) introduce other factors (trust, social information) to achieve better accuracy of recommendation; and (2) employ user preference to provide personalized services recommendation.

**Acknowledgment.** This work was supported by Science and Technology on Information Transmission and Dissemination in Communication Networks Laboratory. The research was supported by NSFC (61571066).

## References

1. Ma, Y., Wang, S., Yang, F., Chang, R.N.: Predicting QoS values via multi-dimensional QoS data for web service recommendations. In: 2015 IEEE 22nd International Conference on Web Services, pp. 249–256 (2015)
2. Tao, C.H., Feng, Z.Y.: Novel QoS-aware web service recommendation model. *Appl. Res. Comput.* **27**, 3898–3902 (2010)



3. Zheng, Z., Ma, H., Lyu, M.R., King, I.: QoS-aware web service recommendation by collaborative filtering. *IEEE Trans. Serv. Comput.* **4**, 140–152 (2011)
4. Wu, X., Cheng, B., Chen, J.: Collaborative filtering service recommendation based on a novel similarity computation method. *IEEE Trans. Serv. Comput.* **5**, 1–11 (2015)
5. Xiao, H., Zou, Y., Ng, J., Nigul, L.: An approach for context-aware service discovery and recommendation. In: 2010 IEEE 17th International Conference on Web Services, pp. 163–170 (2010)
6. Liu, D., Meng, X.W., Chen, J.L.: A framework for context-aware service recommendation. In: 2008 IEEE 10th International Conference on Advanced Communication Technology, pp. 2131–2134 (2008)
7. Kuang, L., Xia, Y., Mao, Y.: Personalized services recommendation based on context-aware QoS prediction. In: 2012 IEEE 19th International Conference on Web Services, pp. 400–406 (2012)
8. Yu, C., Huang, L.: Time-aware collaborative filtering for QoS-based service recommendation. In: 2014 IEEE 21st International Conference on Web Services, pp. 265–272 (2014)
9. Moraru, A., Fortuna, C., Fortuna, B., Slavescu, R.R.: A hybrid approach to QoS-aware web service classification and recommendation. In: 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing, pp. 343–346 (2009)
10. Chen, X., Zheng, Z., Liu, X., Huang, Z., Sun, H.: Personalized QoS-aware web service recommendation and visualization. *IEEE Trans. Serv. Comput.* **6**, 35–47 (2013)
11. Dey, A.K.: Providing architectural support for building context-aware applications. *Elementary Education*, **25**, pp. 106–111 (2000)
12. Chen, X., Zheng, Z., Yu, Q., Lyu, M.R.: Web service recommendation via exploiting location and QoS information. *IEEE Trans. Parallel Distrib. Syst.* **25**, 1913–1924 (2014)
13. Tang, M., Jiang, Y., Liu, J., Liu, X.: Location-aware collaborative filtering for QoS-based service recommendation. In: 2012 IEEE 19th International Conference on Web Services, pp. 202–209 (2012)
14. Herlocker, J.L., Konstan, J.A., Borchers, A., Riedl, J.: An algorithmic framework for performing collaborative filtering. In: *International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 230–237 (1999)
15. Jiang, Y., Liu, J., Tang, M., Liu, X.: An effective web service recommendation method based on personalized collaborative filtering. In: 2013 IEEE 20th International Conference on Web Services, pp. 211–218 (2013)