Integrating Collaborative Filtering Technique in Recommender Scheme for Iot Applications

B. J Lakshmi Narayana[a], K.Tejo Venkat[b], K.PraveenKumar[c], V.BalaChandar[d],T.VamsiKrishna[e]
[a]Assoc.professor, [b,c,d,e] B.Tech.,Scholars
Dept. of Computer Science & Engineering, QIS College of Egineering and Technology

Abstract-To ensure reliable data and device connection, IoT applications must continuously monitor and respond to service calls. However, as the information grows larger, it becomes more difficult to offer accurate and timely service. The recommendation system has been presented as a solution to the problem of information overload. There have been several advances in the development of collaborative filtering (CF)-based recommendation systems in recent years, however most of them suffer sparse difficulties and scalability issues.This study presents a personalised recommendation algorithm based on location and time data. The missing value of the user-service-time tensor is padded across the adjacent time period to relieve data sparsity. Our methodology is scalable because the users and services sets are split into many clusters based on geographical information, and comparable things are picked in smaller and very similar clusters.In order to enhance prediction accuracy, we use the time decay function and customised weight in our technique. Extensive studies using real-world data show that our technique can significantly increase prediction accuracy when compared to other methods.

Keywords--Recommender systems, collaborative filtering, clustering model, personalized recommendation

I. INTRODUCTION

Many technologies are being used to IoT situations as a result of the fast growth of network technology, and recommendation technology is one of them. The exponential expansion of information resources in the Internet of Things has become one of the most significant hurdles to consumers extracting valuable information from all of the accessible databases rapidly and effectively. Because there is so much data, efficient information filtering techniques are required. Traditional search services return the same sorting information to all users and are unable to customise the search results based on the interests of individual users.To address this issue, recommender systems have been created to give tailored suggestion services based on a user's unique preferences, and they are gaining popularity in both academic and industrial research.Recommendation technology is a key component of Internet of Things (IoT) services [1], since it may enable consumers obtains information at any time and from any location.Collaboration filtering (CF) is a well-known and extensively used recommendation technique among them [2], [3], [4], [5], [6], [7].

The Internet of Things (IoT) can be defined as a worldof interconnected objects that can perceive, execute, andcommunicate with each other and with the environment. IoThas the most popular line of products or software, coveringeverything from sensors and controllers to cloud computing.Its versatility covers a wide range of fields including smarthome, transportation, and environmental monitoring [1].Quality of service (QoS) is widely used to describe thenon-functional characteristics of service, such as responsetime, throughput, etc [4]. In fact, the number of IoT servicesis far greater than users, which leads to great data sparsity.

Therefore, the main issue of service prediction research ishow to accurately predict and calculate the missing QoSvalue through the available QoS value.Collaborative filtering (CF) is the most widely used andmost successful recommendation technology inpersonalized recommendation systems. Legion approachesbased on CF are proverbially utilized to make the predictionof QoS values [2]. The CF algorithm discovers the user'spreferences by mining the user's historical behavior data andrecommends similar items in light of different preferences.In addition, CF can recommend completely different items.Existing CF algorithms are often plagued by thescalability problem because the number of users andservices is usually very large. Under this circumstance,calculating similarity values for each pair of users orservices is time-consuming. Moreover, in a highly dynamicInternet environment, the QoS value of IoT service maychange at any time [3]. Therefore, in order to achieve thebest performance of the prediction algorithm, it is necessaryto take into account the time and the location factor. Mostexisting methods seldom apply these factors to theprediction algorithm.

In this paper, a hybrid collaborative filtering method isproposed for IoT service recommendation. The averageQoS value of adjacent time intervals is used to fill part ofthe data, and then use the filled data and the enhancedsimilarity method to make the prediction and producerecommendation results. The contribution of our paper isfour-fold:

(1) An enhanced measurement for computing QoSsimilarity between different users and betweendiverse services based on the personalization ofservice items and user items is proposed to improvethe accuracy of similarity calculation.
(2) In order to improve the prediction accuracy, the timedecay function is exploited. Since the QoSperformance of service is highly correlated withtime.
(3) Based on the above enhanced similaritymeasurement, users set and services set are dividedinto numerous clusters according to locationinformation and we try to seek similar users orservices in smaller and highly similar clusters. Thismakes our approach scalable.
(4) Substantial experiments are conducted based onreal-world data. It is indicated that our method isbetter than that of other CF-based models.

II. RELATED WORKS

CF is one of the widely used service recommendationtechnologies, which provides interesting information basedon the rating or behavior of others in the system [3].Collaborative filtering algorithms could fall into thefollowing three categories: user-based CF algorithm, itembasedCF algorithm, and model-based CF algorithm.The basic assumption of user-based CF algorithm is thatusers who like similar items may have the same preferences.Yu et al. [6] proposed a collaborative filtering algorithmbased on time perception and location awareness. Kennethet al. [7] developed a service recommendation method basedon CF. Scholars such as Kang et al. [8] presented a novelservice prediction approach incorporating the user'spotential QoS preferences and diverse feature of user'sinterests on services.Item-based CF algorithm utilizes all users' preferencesfor items and then recommends similar items to active usersbased on historical preference information.

Wang et al. [9]proposed a time-aware QoS prediction method. A zeromeanLaplace prior distribution hypothesis is proposed forthe residual of the QoS prediction. Liu et al. [10] raised alocation-aware QoS prediction method that took intoaccount the individualized deviations of services' QoSexperience and improved the accuracy of similaritycalculations. Unfortunately, it did not consider the impact oftime factor on prediction. Model-based CF algorithm aimsto train a prediction model on the basis of users' preference.Karim et al. [11] presented a predictive model forcalculating end-to-end QoS values in vertical services,which exploited historical QoS values and service and userinformation to predict unknown QoS values of compositeservices. Xu et al.

[12] proposed a user-based matrixfactorization model and a service-based matrix factorizationmodel.

Daniel et al. [13] utilized the Hidden Markov Modelto construct services' behavior models. The usage of HMMfor service recommendation is relatively innovative, but it'snot suitable for the larger dataset.Personalized prediction algorithm plays anindispensable role in mining users' potential interests andmitigating information overload. Over the past years, thepersonalized recommendation system has gradually becomeone of the research hotspots in academia. Jiang et al. [19]proposed a personalized hybrid collaborative filteringmethod that integrates personalized UPCC and IPCC. Liu etal. [20] studied the influence of user-item correlation on CF.They combined the relevance of users and items,significantly improving the accuracy of recommendations.Chen et al. [22] put forward a novel and scalable hybrid CFalgorithm called RegionKNN.Since service-oriented IoT applications typicallyintegrate many component services to satisfy certainapplication logic, applications need to maintain theflexibility of QoS changes for their component services toensure end-to-end QoS. In order to solve the aboveproblems, Zhu et al. [21] proposed an adaptive matrixdecomposition method to perform online QoS prediction ofcandidate services. This method can handle the problem ofQoS updating and can perform online prediction in real timeafter receiving the new QoS. Based on the previous research, our work focuses onenhancing prediction accuracy, improving the scalabilityproblem and solving data sparsity.Since IoT services are deployed on the environmentunpredictableInternet, the values of QoS attributes, such asresponse time and throughput, are highly correlated with thelocation of the target users and target services.

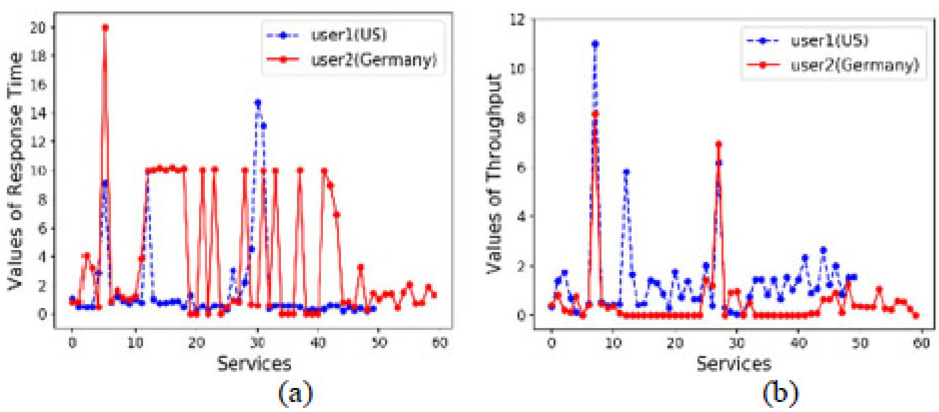


Fig.1. Two users' QoS values

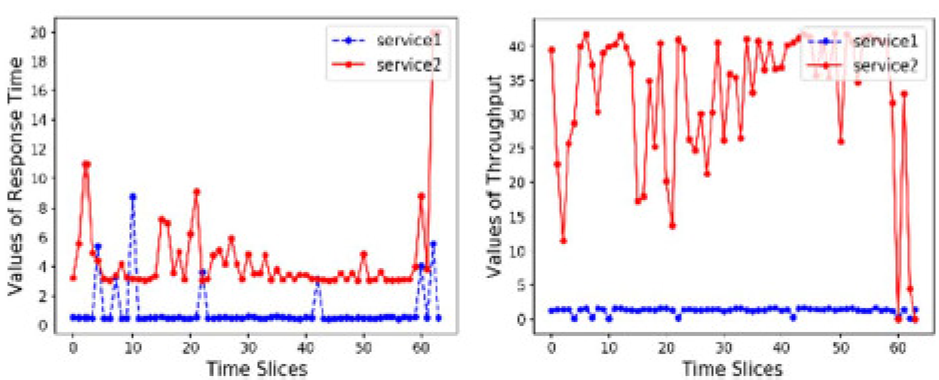


Fig.2. QoS values of two services within different time slices

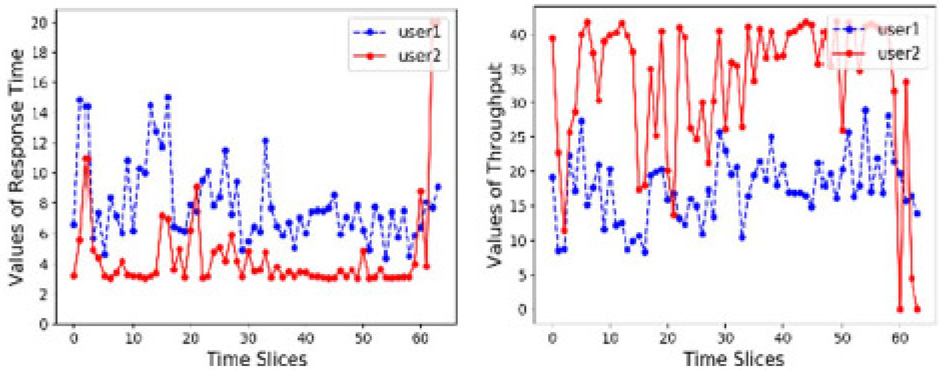


Fig.3. Two users' QoS values within different time slices

Fig.1 shows the response time and throughput values ofservices observed by two users in different locations (user1from the United States and user2 from Germany). It can beseen that although the same service is called, the responsetime and throughput values of the two users are completelydifferent. This experimental observation shows that thevalues of QoS attributes are constantly changing, influencedby factors such as the user's environment, networkconditions, and so on.Fig.2 shows the values of response time and throughputvalues for the same user observing two different serviceswithin 64-time slices. Figure 3 shows the changes inresponse time and throughput for two different usersobserving the same service over 64 time periods. As you cansee, the performance of services changes dynamically overtime. It means that time is an important factor.

III.PROPOSED METHOD

This section mainly introduces the framework of thepersonalized QoS prediction approach proposed in thispaper. The process flow of this model is shown in Fig.4. Ourmodel consists of two parts: data acquisition layer and dataprocessing layer. Users data and services data are collectedin the first layer from the IoT system such as Smart City,Intelligent Medicine, Industry 4.0 and so on. The secondlayer is composed of the data management module, theHPLT QoS prediction module, and IoT servicesrecommendation module. The data management module isin charge of information extraction, transformation, andload from multiple sources. In the IoT servicesrecommendation module, the predicted unknown QoS valueand the original QoS value are transmitted to the IoT servicerecommendation module to select a high-quality service. In our experiments, we focus on two QoS attributes:response time (RTT) and throughput (TP). The responsetime represents the time duration between the user submitting a request and receiving a response, while thethroughput stands for the data transfer rate of the user whoinvoked the service. The average values of response timeand throughput is 1.43 seconds and 10.8 kbps, respectively.Therefore, in order to facilitate data processing, it isnecessary to standardize the original data. The linearnormalization method is used in our experiments:

$$q_{u,i,t} = \begin{cases} 0, & if\ q_{u,i,t} = min(u) \\ 1, & if\ q_{u,i,t} = max(u) \\ \dfrac{q_{u,i,t} - min(u)}{max(u) - min(u)}, & if\ min(u) < q_{u,i,t} < max(u) \end{cases} \quad (1)$$

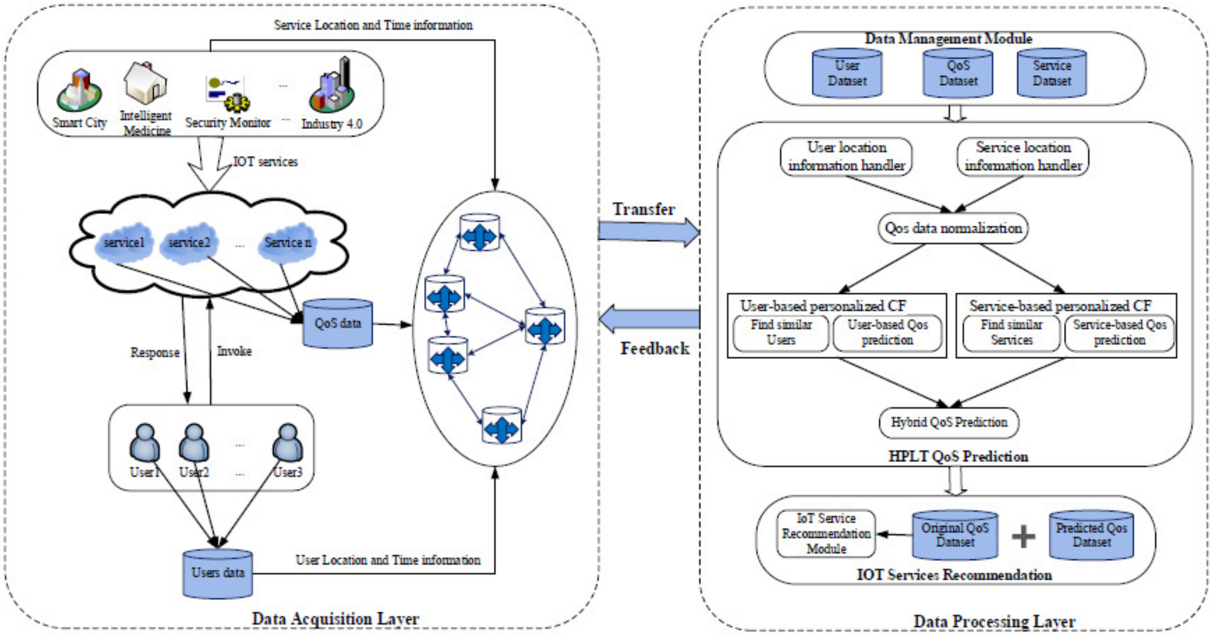


Fig.4. Recommendation model framework for IoT service

Similarity calculation is one of the cardinal steps inCF-based QoS prediction. Pearson correlation coefficient(PCC) has been applied to a great deal of CF-basedprediction algorithms. The selection of similar neighbor is an important stepin predicting the missing values because the choice ofdifferent neighbors will influence the prediction accuracy[25]. The traditional Top-K neighbor selection method isusually applied to CF. However, some entries in the userservice-time tensor may have similar neighbors that are farless than k, or even no neighbors. To solve this problem, wedivide users set and services set into multiple clusters basedon location information and try to find similar users orservices in smaller and similar clusters. When new services enter, similar users can be found in similar clusters, greatlyreduce the computational complexity. After predicting missing values of QoS, we utilizethem to select a high-quality IoT service for the user in theIoT service recommendation module. When an active userrequests a service with specific functionality, the predictedQoS value will help the user identify a set of high-qualityservices and recommend the optimal service to the activeuser.

IV. RESULTS AND DISCUSSION

A set of experiments were conducted to evaluate theperformance of our QoS prediction method. All of theexperiments in this paper are running on the Mac OS Sierraoperating system. The configuration is: Intel Core i52.7GHz processor and 8G memory. The experimentprogram was written in Python and the version was 3.6.4. This paper adopts a real-world service QoS data setWSDream1, allowing users to carry out service reliabilityand quality assessment in a collaborative manner. Itincludes the response time and throughput values of 4,500services from 142 users within 64 time intervals [4]. Thedata set is transformed into two tensors: the user-servicetimeresponse time tensor and the user-service-timethroughput tensor. Then, the 142 users are divided into twoparts: 20 users are randomly selected as test users and therest as training users. Each group of experiments isperformed 100 times and average values are taken as results.In our experiments, we compare the prediction accuracyof UPLT, IPLT and HPLT with other IoT servicerecommendation methods, as shown below:

(1) UPCC: UPCC [17] is an user-based CF method andit employs the average QoS value from the users onthe other service to predict the missing value. UPCCis a well-known prediction method.

(2)IPCC:IPCC [18] is an item-based CF method thatpredicts missing values from the average QoS valueof service items observed by other users. And IPCCis also a notable prediction method.

(3)WSRec: WSRec [16] is a hybrid recommendationmethod combing UPCC and IPCC. But WSRecsuffers from serious scalability problems.

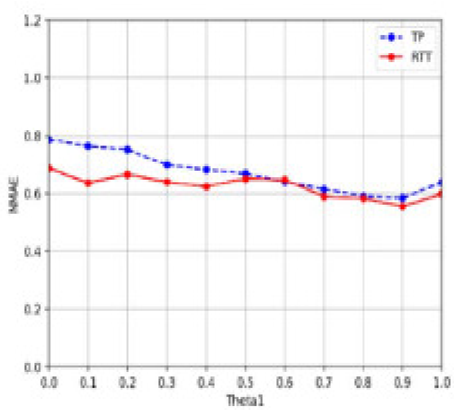
(4)HLACF: HLACF [10] divides users set andservices set into multiple clusters based on locationinformation to solve scalability issues, and it alsoconsiders the personalized impact of services andusers. But it can’t incorporate time factor into similarneighbor selection.

The experiments are carried out with the matrix densityof 30%, Top-K=20, λ=0.8, α=0.2 , and θ_1 = θ_2 =0.9.Besides, in order to reduce the scope depth, we set d to 1.MAE and NMAE performance comparisons betweenUPLT, IPLT, HPLT and other methods are shown inTable1.From Table 1, it can be seen that although theperformance of MAE and NMAE of IPLT and UPLT is notalways better than HLACF, the performance of the hybridalgorithm HPLT is lower than both IPLT and UPLT, andalso lower than other methods. It indicates that the HPLTcan achieve better accurate prediction.

TABLE 1. MAE and NMAE Performance Comparisons

		UPCC	IPCC	WSRec	HLACF	IPLT	UPLT	HPLT
RTT	MAE	0.697	0.903	0.622	0.543	0.662	0.521	0.508
RTT	NMAE	0.986	0.850	0.807	0.741	0.631	0.716	0.562
TP	MAE	50.630	45.372	36.024	27.659	26.328	27.014	26.945
TP	NMAE	0.973	0.891	0.819	0.611	0.640	0.703	0.581

Fig. 5 shows the results produced by θ_1. The influenceof θ_2 is similar to θ_1, as shown in Fig. 6. It indicates that $\theta_1$and θ_2 have a significant impact on prediction accuracy.



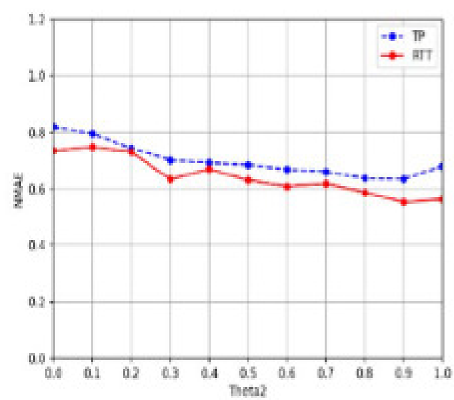
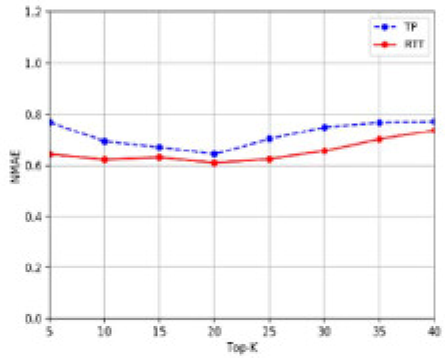


Fig.5. The influence of parameter θ_1Fig.6. The influence of parameter θ_2

Top-K needs to be determined when selecting similarneighbors. In order to study the impact of Top-K onprediction results, we increase the value of Top-K from 5to 40. The experimental results are shown in Fig. 7.It can be seen from Fig. 7 that with the increases ofTop-K, the value of NMAE decreases slowly. When Top-K reaches to 20, NMAE achieves a minimum value andthen increases rapidly. The situation is reasonable becausewhen Top-K is not large enough, only a few similarneighbors are selected for prediction, and the valuableinformation of similar neighbors is not fully utilized. When Top-K is too large, some users with low similarity will beused and lead to low prediction accuracy.The parameter α is a settable attenuation coefficientemployed in the Newton's law of cooling. The greater thevalue of α, the faster the user's interest in the service decays.In order to study the effect of the parameter α onprediction accuracy, the parameter α is set from 0 to 1 insteps of 0.1. Results is shown in Fig. 8. As can be seen fromFigure 8, with the value of α increases, NMAE decreasesfirst and then increases. When α is 0.2, the prediction getsthe highest accuracy. It is reasonable, because if the initialinterest of users in a service is 100 points at the beginning,his/her interest may be "cooled" to 1 point after 24 days.



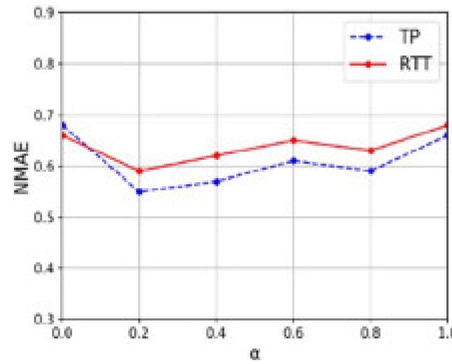


Fig.7. The influence of parameter top-k Fig.8.The influence of parameter α

V. FUTURE SCOPE AND CONCLUSION

Based on the temporal correlation coefficient (TCC) and CSK-means, this study suggested an unique collaborative filtering technique (TCCF). The CSK-means method is used by TCCF to break down a large data issue into smaller, more manageable chunks. The clustering approach is a pre-processing step that groups users who are similar in order to provide more accurate and timely recommendations. To begin, we used Cuckoo search, an unique intelligent optimization method, to increase the clustering impact of the K-means algorithm.Then, to account for interest drift over time, we devised a time factor. Finally, in order to increase the quality of TCCF, we created an effective and personalized recommendation model based on preference pattern (PTCCF). By studying the user's activity, it can provide a better suggestion. Our suggested models TCCF and PTCCF are effective and efficient for a rapid and accurate suggestion, as demonstrated by systematic experimental findings on MovieLens and Douban.

REFERENCES

[1] Qusay F. Hassan, "Introduction to the Internet of Things," in Internetof Things A to Z: Technologies and Applications , 2018.

[2] Arunachalam, N., Amuthan, A., Shanmughapriya, M., Kavya, C., &Gandhi, L. D. (2017). A compendium on web service predictiontechniques. International Conference on Computation of Power,Energy Information and Commuincation (pp.290-294).

[3] Ekstrand, Michael D., John T. Riedl, and Joseph A. Konstan."Collaborative filtering recommender systems." Foundations andTrends in Human-Computer Interaction 4.2 (2011): 81-173.

[4] Zheng, Z., Zhang, Y., &Lyu, M. R. (2010, July). Distributed QoSevaluation for real-world web services. In Web Services (ICWS),2010 IEEE International Conference on (pp. 83-90). IEEE.

[5] Zheng Z, Zhang Y, Lyu M R. Investigating QoS of Real-World WebServices[J]. IEEE Transactions on Services Computing, 2014,7(1):32-39.
[6] Yu C, Huang L. A Web service QoS prediction approach based ontime-and locationaware collaborative filtering[J]. Service OrientedComputing and Applications, 2016, 10(2): 135-149.
[7] Fletcher K K, Liu X F. A collaborative filtering method forpersonalized preference-based service recommendation[C]//WebServices (ICWS), 2015 IEEE International Conference on. IEEE,2015: 400-407.
[8] Kang G, Tang M, Liu J, et al. Diversifying web servicerecommendation results via exploring service usage history[J]. IEEETransactions on Services Computing, 2016, 9(4): 566-579.
[9] Wang X, Zhu J, Zheng Z, et al. A spatial-temporal QoS predictionapproach for timeaware web service recommendation[J]. ACMTransactions on the Web (TWEB), 2016, 10(1): 7.
[10] Liu J, Tang M, Zheng Z, et al. Location-aware and personalizedcollaborative filtering for web service recommendation[J]. IEEETransactions on Services Computing, 2016, 9(5): 686-699.
[11] Karim R, Ding C, Miri A, et al. Incorporating service and userinformation and latent features to predict QoS for selecting andrecommending cloud service compositions[J]. Cluster Computing,2016, 19(3): 1227-1242.
[12] Xu Y, Yin J, Deng S, et al. Context-aware QoS prediction for webservice recommendation and selection[J]. Expert Systems withApplications, 2016, 53: 75-86.
[13] Canton-Puerto D G, Moo-Mena F, Uc-Cetina V. QoS-Based WebServices Selection Using a Hidden Markov Model[J]. JCP, 2017,12(1): 48-56.
[14] Ma H, Zhu H, Hu Z, et al. Multi-valued collaborative QoS predictionfor cloud service via time series analysis[J]. Future GenerationComputer Systems, 2017, 68: 275-288.
[15] Chen Z, Shen L, Li F. Exploiting Web service geographicalneighborhood for collaborative QoS prediction[J]. FutureGeneration Computer Systems, 2017, 68: 248-259.
[16] Zheng Z, Ma H, Lyu M R, et al. Wsrec: A collaborative filteringbased cloud service recommender system[C]//Cloud Services, 2009.ICWS 2009. IEEE International Conference on. IEEE, 2009: 437-444.
[17] Breese J S, Heckerman D, Kadie C. Empirical analysis of predictivealgorithms for collaborative filtering[C]//Proceedings of theFourteenth conference on Uncertainty in artificial intelligence.Morgan Kaufmann Publishers Inc., 1998: 43-52.
[18] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborativefiltering recommendation algorithms[C]//Proceedings of the 10thinternational conference on World Wide Cloud. ACM, 2001: 285-295.
[19] Jiang Y, Liu J, Tang M, et al. An Effective Web ServiceRecommendation Method Based on Personalized CollaborativeFiltering[C]// IEEE International Conference on Web Services.IEEE Computer Society, 2011:211-218.
[20] Liu R R, Jia C X, Zhou T, et al. Personal recommendation viamodified collaborative filtering[J]. Physica A Statistical Mechanics& Its Applications, 2008, 388(4):462-468.
[21] Zhu J, He P, Zheng Z, et al. Online QoS Prediction for RuntimeService Adaptation via Adaptive Matrix Factorization[J]. IEEETransactions on Parallel and Distributed Systems, 2017.
[22] Chen X, Liu X, Huang Z, et al. RegionKNN: A Scalable HybridCollaborative Filtering Algorithm for Personalized Web ServiceRecommendation[C]// IEEE International Conference on WebServices. IEEE Computer Society, 2010:9-16.

[23] Yu C, Huang L. A Web service QoS prediction approach based ontime- and location-aware collaborative filtering[J]. Service OrientedComputing and Applications, 2016, 10(2):135-149.

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

[25] Zheng, Z., Ma, H., Lyu, M. R., & King, I. (2011). QoS-aware webservice recommendation by collaborative filtering. IEEETransactions on services computing, 4(2), pp. 140-152.