A QoS and QoE based Integrated Model for Bidirectional Web Service Recommendation

Sneha Jhaveri[1] Pooja M Soundalgekar[2] Kevin George[3] Sowmya Kamath S.[4]

Department of Information Technology,
National Institute of Technology Karnataka, Surathkal, India 575025
Email: [1]snehajhaveri910@gmail.com, [2]pooja27ms@gmail.com
[3]14it117.kevin@nitk.edu.in, [4]sowmyakamath@nitk.edu.in

***Abstract*— For a given requirement, identifying relevant Web services and recommending the best ones is an important task in Service-oriented application development. In this paper, a composite model that leverages Quality of Service (QoS) and Quality of Experience (QoE) for bidirectional Web service recommendation (bi-WSR) is proposed. The QoS based recommendation model is built on degree of user satisfaction, calculated using a special normalization technique and user satisfaction functions like response time and throughput. The QoE model is trained on a dataset containing positive, negative and neutral textual reviews of web services for sentiment analysis and mapped to each service's QoS values using a clustering method. This further optimizes the recommendation of web services to consumers, as the sentiment score of reviews is integrated with the user satisfaction using weighted average scoring. To describe the relationship between both web services & consumers and providers & consumers, a cube model is built. For recommending services to consumers and recommending potential consumers to service providers, hybrid collaborative filtering based techniques were used. The results obtained when only QoS is used, and when QoS and sentiment analysis scores are integrated to form QoE showed significant improvement in the quality of recommendation.**

***Index Terms*—Web Service Recommendation, Quality of Service, Quality of Experience, Sentiment analysis, Bidirectional recommendation**

I. Introduction

The Web Service paradigm effectively enables interoperable, machine-to-machine interactions on the World Wide Web, facilitating successful development and management of large-scale Web applications. Such applications are essentially complex workflows incorporating several tasks, each handled by different web services from multiple service providers. A Web service recommendation (WSRec) system can thus help users in choosing services most suited to their task given constraints on time, geographical location and cost of service (CoS). Such WSRec systems generally employ user feedback and service evaluation mechanisms for collecting information about the Web Services. Web service recommendation using the functional characteristics of a service is however independent of user preference/need, while approaches based on service non-functional characteristics can effectively incorporate both Quality of Service (QoS) attributes and user constraints/requirements. Usage statistics and feedback from other users can also be modeled in the form of Quality of Experience (QoE) and leverages to further improve user satisfaction in recommendation.

Generally, product recommender systems capture the relationship between items (web services) and consumers (service users) in a two-dimensional space referred to as user-item matrix. Social networks like Facebook & LinkedIn and e-commerce sites like Amazon have utilized collaborative filtering (CF) algorithms for services like friend suggestion, job recommendations, product recommendation etc. Such recommendations are highly personalized, i.e., they are driven by the previous choices made by the user; the system also estimates the value of a potential item from the particular user's point of view. Content-based filtering algorithms drive such recommendations, due to which relevancy can further improve if customer reviews are also taken into consideration along with their ratings. For example, domain-specific factors such as price, cuisine, location and user reviews can help in improving the recommendation accuracy of a restaurant finder application. Similarly, a hybrid recommendation approach that uses both service QoS information and users' text reviews for optimizing recommendation can be significantly more helpful for users.

When finding similarity between different available web services, the variations of service QoS perceived by service users should also be taken into account. QoS is regarded as a set of user-perceived properties, that highly relates to users' physical locations. Collaborative Filtering (CF) based methods have been successfully applied to QoS-aware Web Service Recommendation (QWSRec) over the past few years [1,2,3]CF methods have been further refined to consider geographical location to avoid recommending irrelevant neighbors (users or services) for a target user [4,5,6]. A significant challenge in QWSRec is the problem of dealing with missing QoS values. CF approaches have used machine learning models for missing QoS value prediction, using which subjective data can be predicted. Prediction algorithms for unknown QoS values of web services typically use objective characteristics of the service, like, response time, throughput and reliability [7,8] . QoS prediction approaches that model multi-dimensional QoS data that considers time and location along with QoS have also been developed [9,10].

In this paper, we propose an integrated approach for

QoS aware and QoE enhanced bidirectional Web Service recommendation. Services hosted on different sites may provide similar functionalities and our aim is to recommend the best Web service from the available pool. Our approach is built on a *Cube model*, that captures the relationship between service provider, service consumer and the service, for best-service-for-consumer and potential-consumer-to-provider recommendation. For supporting service recommendation, a user-item matrix that contains the QoS vector is interpreted by a user with regard to a specific service is used, without the involvement of service providers. We also consider service provider-specific information for generating good recommendations and improve prediction accuracy. Further, sentiment analysis is applied to natural language reviews of the service users, which are leveraged for measuring QoE for each web service. The rest of this paper is organized as follows: Section II presents a detailed discussion on the proposed methodology for building an integrated QoS and QoE model for bidirectional service recommendation. Experimental evaluation of the proposed approach carried out on standard QoS datasets is discussed in Section III, followed by concluding remarks with possible directions for future work and references.

II. Proposed Methodology

Fig. 1 depicts the important processes designed for building the proposed integrated model. Our approach is comprised of two parts primarily - processes dealing with QoS prediction and those performing QoE computation. The QoS ratings generated by the Consumer-WSRec algorithm is combined with generated QoE scores obtained after sentiment analysis of user reviews.

A. Web Service QoS Generation

For bidirectional service recommendation, it is essential to capture the latent relationships between the consumers, service providers and the services themselves. This is achieved using a cube model, which controls the Web Service Recommendation, effectively describing the relationship between consumers, providers and web services. The cube model is represented by three 2-dimensional matrices: provider-service binary matrix (A), consumers-provider frequency matrix (V) and consumer-service QoS matrix (Q). The consumer-service QoS matrix Q represents the QoS vectors perceived by users on services where 'a' gives the Response time, 'b' gives throughput values and 'c' represents the User Satisfaction Degree (USD) w.r.t QoS values. These QoS ratings are then combined with the sentiment scores. The WS-DREAM dataset [11]was used for the experimental evaluation of the proposed approach.

For accurate and reliable Web Service Recommendation, a Standard Deviation based Hybrid Collaborative Filtering (SD-HCF) algorithm [1] was used. Further, Inverse frequency based User Collaborative Filtering (IF-UCF) algorithm was employed for recommending potential consumers to service providers. The SD-HCF algorithm uses similarity of web services and consumers bias-from-mean weights from k nearest neighbors are modified to standard deviation. The service provider's impact is accounted for when determining the k nearest neighbors. Both IF-UCF and UCF are almost same when the weight of providers is adjusted using the inverse consumer frequency.

B. QoE Computation

The QoE for a web service is computed using both the sentiment score given to the text reviews of that web service

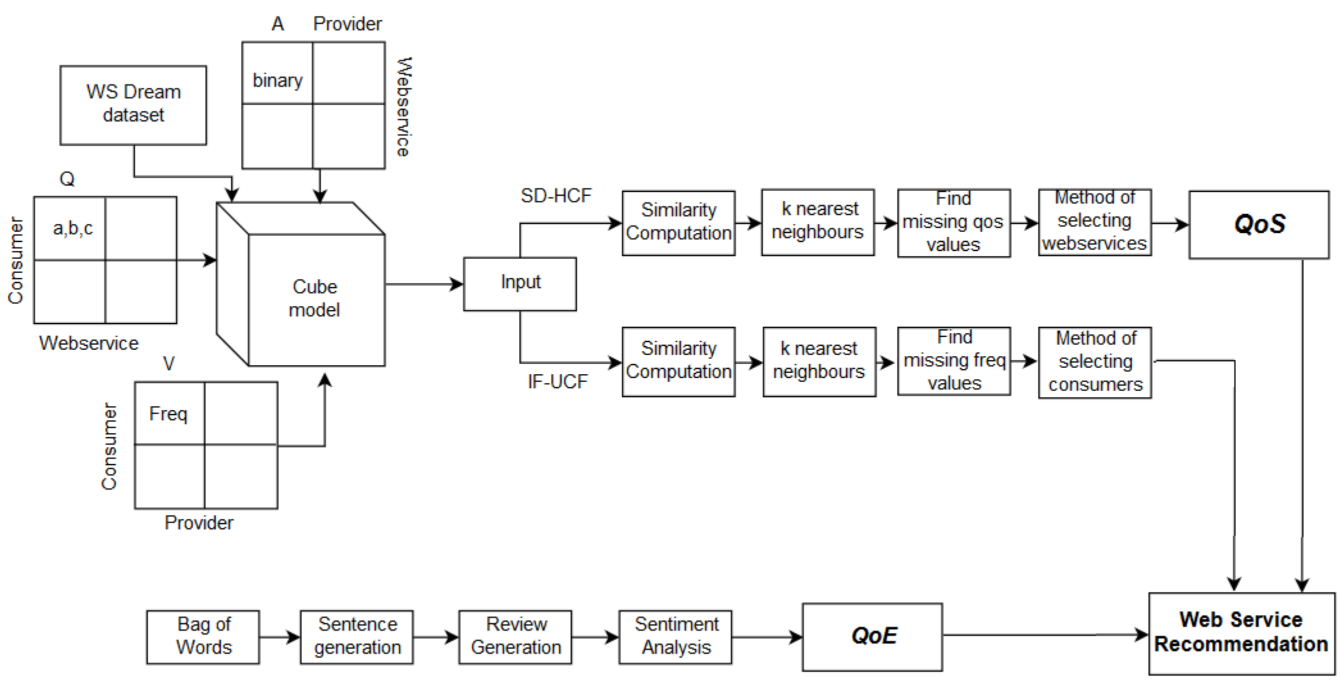


Fig. 1: Proposed Methodology

and the QoS value for the web service. For review sentiment analysis, a dataset consisting of reviews of each Web service was created. The process of creating and using this dataset for QoE computation is described next.

1) Review Dataset Generation: A major hurdle in using reviews for service recommendation is that, standard datasets are not available. Also, actual reviews of real-world Web services that can be collected from service portals like ProgrammableWeb.com, do not represent the services in WS-DREAM dataset. In order to take care of this problem, we created a synthetic dataset that contained actual user reviews of services published on service portals and mapped these reviews to the services in WS-DREAM dataset. Thus, the generated QoS values and computed QoE values can be used to rank a Web service for well-rounded recommendation. As part of future work, we intend to make this dataset available to the research community for experimental purposes.

For the reviews collected, a WordNet of extremely positive, positive, neutral, negative, extremely negative adjectives was designed. This was used for objectively comparing QoS parameters like throughput and latency. Negation phrases were also considered for dealing with specific comments like, "*not that bad*", "*really not good*" etc., which imply positive and negative meaning respectively. We also considered special words which could mean different in a web service terminology, that is domain based words like '*update*', for example- "*The Web service needs to be updated*". Here, '*update*' generally means a good thing but this sentence is a not positive sentence.

2) Sentence Generation: The WordNet adjectives defined earlier were used for creating the synthetic dataset. The sentences in the service reviews have 1 or 2 blanks which can be meaningfully filled in random order by these WordNet adjectives, QoS parameters and negations. With N_1 extremely positive adjectives, N_2 QoS parameters, S_1 adjective sentences and S_2 QoS sentences, the number of unique sentences which can be formed are $S_1 * N_1 * (N_1 - 1) + S_2 * N_2 * (N_2 - 1)$ and number of unique paragraphs which can be formed using one QoS filled sentence and two adjective filled sentences in random sequence in the paragraph are $S_1 * N_1 * (N_1 - 1) * (S_1 - 1) * N_1 * (N_1 - 1) * S_2 * N_2 * (N_2 - 1)$. So, with larger set of WordNet and blank sentences, a large dataset of reviews were obtained.

3) Review Generation: The reviews dataset is generated using an algorithm which incorporates the fact that reviews can be positive, negative or neutral. Algorithm 1 depicts this process and as it is very generic, it can be used to generate any other dataset as well. This is mapped to a scale of 1-5 where, a value of 1 represents *very bad*, 2 for *bad*, 3 for *neutral*, 4 for *good* and 5 for *excellent*. The dataset is then grouped into five set of paragraphs, each with a certain structure, with 2-3 sentences meaningfully filled with two adjectives and one QoS parameter, using repeated randomization technique.

Algorithm 1 Web Service Review Generation

Input: Sentence files with blanks
WordNet of adjectives
QoS parameters
Domain-specific keywords.
Output: Five categories of Web service reviews

for all sentences with blanks for adjectives/QoS/domain words **do**
  if a blank is encountered **then**
    randomly add adjective, QoS or domain words
    add the sentences to respective sentence groups
  end if
end for
for all completed sentences **do**
  randomly combine any three sentences which represent the same group from (1-5) to generate review paragraphs
  add the paragraphs to respective review groups
end for
for all review paragraphs **do**
  if satisfy the random variable condition **then**
    add negation/domain based sentence to this paragraph
  end if
end for
return all five web service review categories

C. Sentiment Analysis of reviews

Sentiment Analysis plays a very important role in QoE based service recommendation. The user preferences are extracted from text reviews of the web services and these are mapped to the ratings generated from the consumer-service recommendation algorithm (SD-HCF). A numerical rating is extracted for these text reviews, which is fed to the algorithm, thus integrating QoE value to the web service representation. Sentiment analysis was performed using the python package, TextBlob, which internally uses NLTK (Natural Language Tool Kit) corpus for the sentiment computation. The additional WordNet used for sentiment analysis contained the domain specific words which were assigned by us. Firstly, textblobs are created for the input review, from which sentences are extracted. Based on the sentiment identified, the polarity of the sentence is found, which is used to compute the sentiment score of each review (as shown in Algorithm 2).

D. User Satisfaction Degree (USD)

User Satisfaction degree (USD) measures the performance of QoS attributes during the process of recommendation. The mapping of QOS parameters into USD is measured using the User Satisfaction Function (USF). USD has a value in the range [0,1], the higher its value, higher is the degree of user satisfaction. For USD computation, the two files *rtvalues* and *tpvalues* that contain the response time and throughput values respectively are used. The USF is computed by finding the

Algorithm 2 Sentiment Analysis of Web Service Reviews

Input: Web service review files
Output: Sentiment score for each review in the 5 files

```
Initialization:
for all the reviews do
    Convert the paragraph into a text blob
    for all the sentences in each text blob do
        Determine sentence polarity
        Map the sentiment score to respective reviews.
    end for
end for
```

average of all QoS values for each user and then find the overall average of QoS values for all users. This average value is considered as the median class and the rest of the USD values are divided into ten classes with the average being the middle or 5th class. After the USD values are calculated for both throughput and response time values, a user ratings file containing the weighted sum of USD of response time and throughput values of each web service is created. We assigned weights of 0.5 so that both QoS attributes contribute equally to the ratings.

E. Integrating Sentiment scores and QoS (USD)

The sentiment scores obtained for each review paragraph are now to be categorized into the five 1-5 scale categories described earlier. Each review is then mapped to a USD value belonging to the same category. While creating the synthetic dataset, we ensured that a sentiment score of *excellent* is integrated with a consumer-service pair for which the response time is low and throughput is high, thus obtaining a very high USD value. Higher the sentiment score, higher is the USD value with which it is integrated. For generating the 5 categories, K-means algorithm was used. To each cluster, a review which has the relevant sentiment score is assigned, i.e., the score (1-5) and QoS cluster (1-5) should match. During this integration, the weighted average of the sentiment score and the USD value is taken. If N is the number of QoS parameters involved in determining the USD value, then the weighted average is of ratio *1:N* to the USD value. As mentioned earlier, two QoS parameters, *response time* and *throughput* are considered, and the weighted average to determine QoE is computed as per Eq. (1).

$$Weighted\ average = \frac{(Sentiment\ score + 2 * USD)}{3} \tag{1}$$

F. Bidirectional Recommendation

The interaction between the *web service-provider*, *provider-consumer* and the *consumer-web service* defines the recommendation process as each of these are interdependent. This relationship can be exploited for supporting several applications like recommendation of web services to consumers and recommendation of consumers to providers. There may be many similar services with the same functionality, and several providers who offer such services, but they are differentiated using their computed USD values, then enabling effective ranking. Due to their different QoS values (response time, failure rate, throughput etc) and varied QoE values, it is easy to choose the best service to suit a specific purpose.

We used the Standard Deviation-Hybrid Collaborative Filtering (SD-HCF) [1] for validating our approach's suitability for providing web service recommendations to consumers. SD-HCF is comprised of three main phases: (1) Calculation of Similarity between Consumer and Web Services (2) Computation of the k-Nearest Neighbors for Consumers and Web Services (3) Unknown QoS vector Prediction. The algorithm outputs a binary decision that indicates if a particular web service is recommended or not. The rating scale needs to be binary and any other scale needs to be mapped to binary level. For validating recommendation of potential consumers to service providers, we used the IF-UCF algorithm [12]. The 3 phases used in this algorithm is same as the above except that one more parameter is considered, i.e., Inverse Consumer Frequency is used to model the recommendation to the providers. After the completion of the phases, top N target objects with a high USD value for active providers and consumers are selected.

III. Experimental Results & Discussion

For validating the proposed methodology, we created the dataset of Web service sentiment reviews for services spanning five categories, each with 100,000 reviews. On this dataset, sentiment analysis was performed for understanding user satisfaction and computing its degree. It was found that, the rating predicted matched the rating assigned with almost 100% accuracy due to the effective domain modeling approach adopted by us. An analysis of the same is shown in Table I. The WS-DREAM dataset was used for the QoS computation, where 5825 web services with their response time and throughput values for 339 users were provided.

TABLE I: Results of Review Sentiment Analysis *(For 100,000 reviews)*

Sentiment Score #	*Correctly Classified*	*Mis-classified*	*Accuracy (%)*
Score 1	99,890	110	99.6
Score 2	99,384	616	98.9
Score 3	98,703	1297	98.5
Score 4	99,468	532	99.1
Score 5	99,895	105	99.8

For measuring the performance of the bidirectional recommendation framework, we used the metric Mean Absolute Error (MAE). MAE is the ratio of the total absolute difference between actual QoE (or actual QoS) denoted by Q and predicted QoE (or predicted QoS value) denoted by Q' to the number of predicted QoE or QoS values. It is computed as per Eq.(2), where, M is the number of predicted QoE or

QoS values, c represents the consumer and n represents the web service.

$$MAE = \sum_{i,n} \frac{|Q_{(C_i,S_n)} - Q'(C_i,S_n)|}{M} \quad (2)$$

For our experiments, we also considered the potential variation in the value MAE due to different data densities. In our case, *density* is defined as the percentage of data points considered for evaluation of MAE and accuracy computation. The density values were set at 25, 50, 75 and 100, which represent the percentage of data chosen randomly for the various testcases used. Here, the range $0-25$ means the first quarter, whereas, $50-100$ indicates that the second half of data is considered. Hence, this choice of which part of data should be taken is also randomized. During this process, MAE is computed for the predicted QoE (Sentiment score + USD value) and USD values individually. The results of these experiments are shown in Figure 2 and tabulated in Table IV.

From Table IV, some significant observations can be made. It can be seen that the value of MAE for SD-HCF remains stable with varying density for recommendations based on QoE value, whereas, it fluctuates for response time and throughput on changing density. The time take for computing missing values is lesser with QoE than USD individually as QoE involves calculation of weighted average of USD and sentiment score before predicting missing values. In contrast, when this is computed individually, the process is performed at the end. Thus, the computation time is reduced by a factor of N (where N is the number of QoS parameters contributing USF to the USD), and this can be seen in Table II. From Table III, it is clear that the recommendation results improved with QoE when compared to QoS for top 7 recommendations. A respectable improvement in recommendation relevancy, in the range of 5-6% was observed (here, *relevancy* captures the number of services recommended which were actually top recommendations for that particular customer).

TABLE II: Similarity computation time for recommendation for USFs and USD (Missing value prediction using kNN)

Parameter	*Computation Time (for 1 WS)*	*Computation Time (for 100 WS)*	*Computation Time (for 500 WS)*
USF - Response time & throughput (Before integration computation)	1.2	987	4878
USD (After integration computation)	0.88	653	3213

IV. Conclusion and Future Work

In this paper, a QoS and QoE based composite model for bidirectional Web service recommendation (bi-WSR) is proposed. The QoS based recommendation model was built on the computed degree of user satisfaction, calculated using a special normalization technique and user satisfaction functions. For training the QoE based recommendation model, a review dataset containing positive, negative and neutral text reviews of web services was created, to which sentiment analysis and clustering techniques were used to compute user satisfaction. This is mapped with QoS obtained from USD of web services to compute the Quality of Experience (QoE) values based on which the recommendation is provided. Experimental evaluation underscored the effectiveness of the proposed techniques in further optimizing service recommendation to consumers and potential consumer recommendation to service providers, as the sentiment score of reviews was integrated with the user satisfaction using weighted average scoring. The results obtained when only QoS is used and when QoS and QoE are integrated showed significant improvement in quality of recommendation. For further improvements, we intend to explore location-aware and time-aware bidirectional Web service recommendations. Also, suitable techniques for solving the inevitable cold-start problems associated with recommender systems also need to be addressed, to improve new users' experience.

TABLE III: QoE and QoS based recommendation results

User	*Parameter*	*Services recommended*	*Relevancy (%)*
User 7	QoS	15,74,75,108,110,498,494	76.4
	QoE	56,64,74,96,104,108,110	82.1
User 20	QoS	76,77,126,174,442,446,451	67.8
	QoE	9,76,77,108,126,442,451	71.2

References

[1] Jie Cao, Zhiang Wu, Youquan Wang and Yi Zhuang : Hybrid Collaborative Filtering algorithm for bidirectional Web service recommendation : Knowl Inf Syst (2013) 36:607–627 DOI 10.1007/s10115-012-0562-1

[2] Zhu X, Zhang S, Jin Z, Zhang Z, Xu Z (2011), "Missing value estimation for mixed-attribute data sets", IEEE Trans Knowledge Data Eng 23(1):110–12

[3] Zheng, Zibin, Hao Ma, Michael R. Lyu, and Irwin King. "Qos-aware web service recommendation by collaborative filtering." IEEE Transactions on services computing 4, no. 2 (2011): 140-152.

[4] Jiang, Y., Liu, J., Tang, M., & Liu, X. (2011, July). An effective web service recommendation method based on personalized collaborative filtering. In Web Services (ICWS), 2011 IEEE International Conference on (pp. 211-218). IEEE.

[5] Cao, J., Wu, Z., Wang, Y., & Zhuang, Y. (2013). Hybrid collaborative filtering algorithm for bidirectional web service recommendation. Knowledge and information systems, 36(3), 607-627.

[6] Tang, M., Jiang, Y., Liu, J., & Liu, X. (2012, June). Location-aware collaborative filtering for QoS-based service recommendation. In Web Services (ICWS), 2012 IEEE 19th International Conference on (pp. 202-209). IEEE.

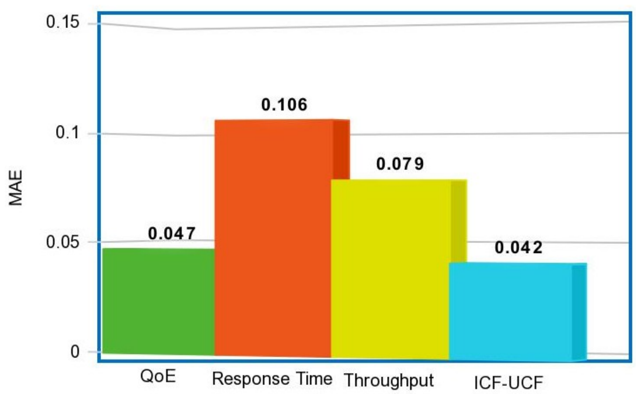
[7] Chen, Xi, Zibin Zheng, Qi Yu, and Michael R. Lyu. "Web service recommendation via exploiting location and QoS information." IEEE Transactions on Parallel and Distributed Systems 25, no. 7 (2014): 1913-1924.

[8] Yao, L., Sheng, Q. Z., Ngu, A. H., Yu, J., & Segev, A. (2015). "Unified collaborative and content-based web service recommendation". IEEE Transactions on Services Computing, 8(3), 453-466.

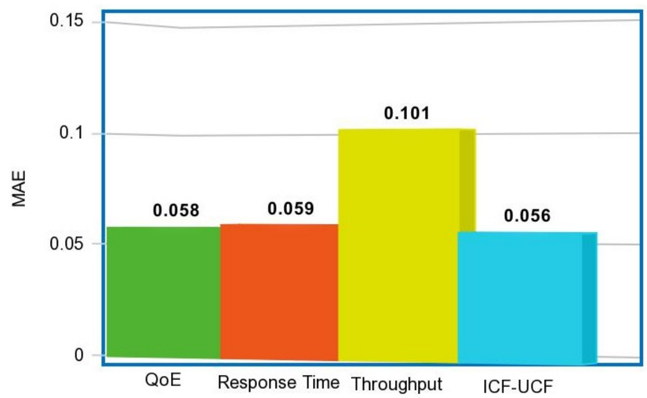
[9] You Ma,Shangguang Wang,Patrick C.K. Hung,Ching-Hsein Hsu,Qibo Sun, and Fangchun Yang (2016) : A Highly Accurate Prediction Algorithm for Unknown Web Service QoS Values: IEEE TRANSACTIONS ON SERVICES COMPUTING,VOL.9 NO. 4.

TABLE IV: Accuracy in terms of MAE for varying density values

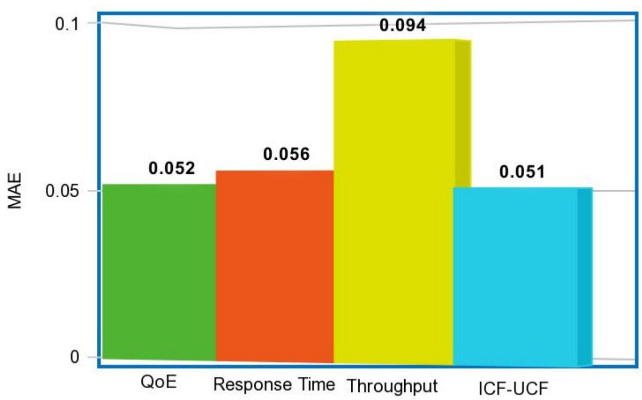
Recommendation	*Parameter*	*Density-25*	*Density-50*	*Density-75*	*Density-100*
Service-to-consumer (SD-HCF)	QoS (Response Time)	0.106 (25-50)	0.059 (25-75)	0.056 (0-75)	0.162(0-100)
Service-to-consumer (SD-HCF)	QoS (Throughput)	0.079 (25-50)	0.101 (25-75)	0.094 (25-100)	0.107(0-100)
Service-to-consumer (SD-HCF)	QoE	0.047 (25-50)	0.058 (0-50)	0.052 (25-100)	0.054(0-100)
Potential Consumer-to-Service	ICF-UCF	0.042 (0-85)	0.056 (85-255)	0.051 (0-255)	0.053 (0-339)



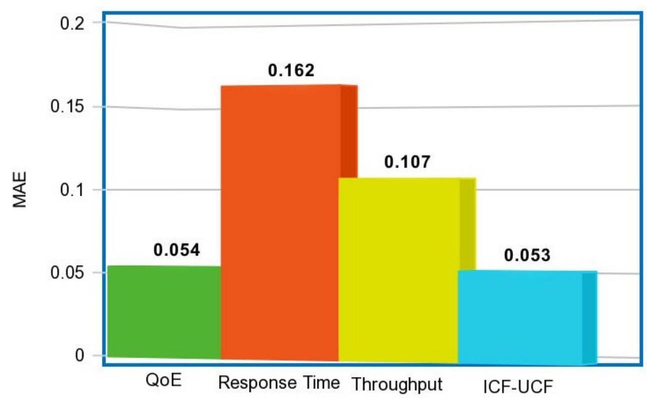
(a) MAE at density 25



(b) MAE at density 50



(c) MAE at density 75



(d) MAE at density 100

Fig. 2: MAE plot for QoE, QoS (Response time, throughput) and ICF-UCF

[10] Wang, Xinyu, Jianke Zhu, Zibin Zheng, Wenjie Song, Yuanhong Shen, and Michael R. Lyu. "A spatial-temporal QoS prediction approach for time-aware web service recommendation." ACM Transactions on the Web (TWEB) 10, no. 1 (2016): 7.

[11] Wang, Shangguang, You Ma, Bo Cheng, and Rong Chang. "Multi-dimensional QoS prediction for service recommendations." IEEE Transactions on Services Computing (2016).

[12] Wei Lo, Jianwei Yin, Shuiguang Deng, Ying Li and Zhaohui Wu (2016) : Collaborative Web Service QoS Prediction with Location-Based Regularization : 2012 IEEE 19th International Conference on Web Services.