IJCTA, 9(10), 2016, pp. 4591-4601

Web Services QoS Prediction Basedon Dynamic Non-Functional Quality Factors and WS-Security Policy Specification of Web Service

***Subbulakshmi S *Ramar K *Arya Krishnan R **Divya S**

Abstract : Web service are basically software components that support interoperable interaction between machines over a network. With booming number of WS, it's difficult for users to identify the best quality services. Existing researchers focus mainly on selection of WS based on the functional requirements. This paper proposes a system to select an optimal WS by predicting its QoS value based on dynamic non-functional quality factors response time, throughput and the static factor security of the WS. Prediction result can be used in recommendation systems to select services with optimal QoS performance among a large volume of service candidates. Users and web services are clustered to make prediction of response time and throughput of WS. Security specifications of the web service and their vulnerabilities are used for prediction of security factor of the WS. Finally, QoS of the WS is predicted by aggregation of the predicted quality values for security, response time and throughput. The QoS prediction of the system reveals optimal results as it takes both dynamic and static quality factors of WS**.**

Keywords : Web service; Dynamic quality factors; Static quality factor; QoS prediction;

1. INTRODUCTION

Web serviceis a promising technique that provides a way to access software functions through standard web protocols and ensures effective communication between systems. In general web serviceare selected based on the functional and non-functional requirement of the users. Functional requirements focus on web service functionality and the non-functional requirements are concerned about quality of service. Quality of Service (QoS) is a fundamental factor in web service selection. Availability, scalability, reliability, throughput, response time, security etc. are the important QoS of a web service. QoS valuesdiffer for each user due to different factors like locations, network status and other objective factors. Since several web services render similar functionalities, web service selection based on desired QoS is considered as an optimized technique for service selection. Most researchers focus on web service selection based on functional requirements. We propose the implementation of a prediction system which is based onthe non-functional requirements of a web service. Non-functional requirements are both dynamic and static. The system proposes the prediction of QoS for a given web service specifiedby the enrolled user.The prediction of the dynamic quality factors response time, throughput is done based on the location details of user and the quality specification given by the service providers. The static quality factor security is predicted based on the security specification given in the WS-Policy [8].

* Department of Computer Scienceand Applications, Amrita School of Engineering, Amritapuri, Amrita Vishwa Vidyapeetham, Amrita University, India.,

** Department of Computer Science and Engineering, Einstein College of Engineering, Tirunelveli, India. [1]subbulakshmis@am.amrita.edu,[2] kramar.einstein@gmail.com, [3]aryarrun16@gmail.com,[4]divya.sivan006@gmail.com

The factors response time and throughput are considered as dynamic QoS as it is highly influenced by the external factors like users location, network bandwidth, server overloading etc. The system proposes a QoS prediction of those factors by consuming the location details of the present user, past experienced user's QoS values for different web services and the details of web services given by the service provider. The architecture diagram of the proposed prediction system is shown in Figure 1. Prediction of QoS starts with grouping of users and web services. Users are grouped based on location information and web services are grouped based on QoS value provided by service Providers. QoS prediction is performed from both user perspective and web service perspective using PCC similarity method the values thus obtained are used for predicting final QoS value for the given web service.

QoS prediction of security is an important issue while deploying a web service. Building a secure web services demands the identification of threats faced, effective trade-off to handle those threats and proper methods to integrate them with the security of web service. To predict the QoS value for security, system identifies WS-Policy description of web service and categories policy implementation based on different security features such as authentication, authorization confidentiality and integrity. For each security algorithms implemented, security levels are assigned based on their efficiency. Security value for web service is calculated by taking into consideration the web service domain, security levels assigned and the vulnerabilities in the web service.

Finally the QoS values of response time, through put and security will be used to predict the overall QoS of Web service. The predicted QoS values can be employed in recommendation systems to facilitate the users to single out the best quality web service which suits their requirements.

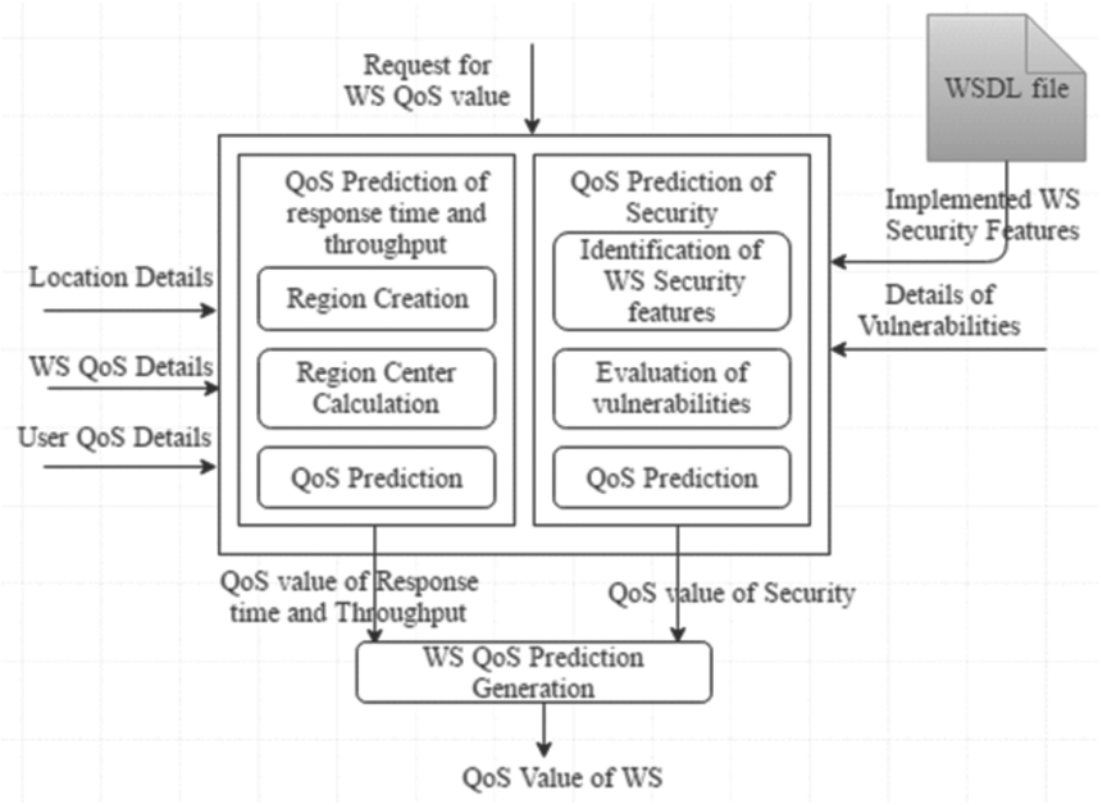

Fig. 1. System architecture.

The remaining section of this paper is organized as: Section II reviews of the related works. Section III describes the proposed system methodologies Section IV shows the results, and Section V concludes the paper with the scope for future enhancements.

2. RELATED WORKS

This section of the paper shows the existing work and research that have been done relating to QoS value prediction using different techniques. The main goal of these studies is to improve the prediction accuracy.

QoS is considered as a significant factor service selection. In reality most of the QoS values are unknown values. For predicting such unknown QoS values a neighborhood based collaborative filtering approach is proposed [1].To remove the impact of incomparable QoS scales cosine similarity calculation is performed and data smoothing process to enhance the prediction accuracy. To address the data sparsity problem a similarity fusion method is proposed. Scalability of the system is improved by using a two phase neighbor selection method by stimulating neighbor selection.

With the rapid increase of web services in the internet it is necessary to have an efficient QoS evaluation method which helps service users to select optimal services from the list of candidate services based on QoS. The problem is that most of the QoS values are missing values, so the prediction of those missing QoS values using the available QoS values are necessary. To overcome this problem a highlycredible QoS prediction approach namely Reputation based Matrix Factorization, RMF [11] is proposed. They find reputation on the basis of user's QoS values and the reputation is combined into a matrix factorization prediction approach to get more accurate predictions [2].

Most of the QoS predictionmethods make prediction based on the past QoS data submitted by identical users and services. The main problem in these methods is to identify the inaccurate data provided untrustworthy users.Reputation of the user is asignificant factor in QoSvalue prediction. To detect the entrusted users and to predict the QoS values, a credibility aware QoS prediction method, CAP [3] is proposed which take datacredibility into consideration and uses two-phase K-means clustering algorithm to identify the untrustworthy user. They perform index calculation and clusters users based on their indexand then predicts the unknown QoS values by clustering the dataprovided by trustworthy users.

Instead of employing the conventional methods, a spatial temporal QoS prediction approach is proposed for time aware web services where the sparse representation method is used to model QoS variations [4].To select the nearest web service for the sparse representation of QoS values effectively, the geo-location of web service is employed while improving prediction accuracy. Even though the method achieved promising performance, some limitations are addressed. The method requires retrieving QoS values at the current time slot, which cannot be applied to forecast future temporal QoS values.

QoS prediction of WS compositionthat implements business process,a graph reduction method is proposed where service composition is represented as a graph and the QoS of such graph is obtained from the composition of its nodes. They useda fast algorithm to predict the QoS values based on graph reduction [5]. The main limitation of the system is that it can't handle large service composition when the service count is high.

QoS is an essential factor for determining the usability of complex software systems. The QoS of allcomponents in complex system is identified with uncertainty and it varies for each invocation that is, composite system also exhibits QoS fluctuation. In [6] a method is suggested to predict the unknown QoS values of service oriented systems by employing the known probability distributions of the component QoS attributes and the composition structure. The accuracy of prediction result depends on the QoS data provided by other services.

Collaborative filtering (CF) [10] is a technique used by the recommender systems and to predict the QoS values of cloud services [12]. CF based QoS prediction approach is followed by many systems andwhich uses QoS values provided by similar users to make predictions.The concept is that predict QoS values and suggestthe optimal web service for users based on web service QoS values contributed by service users and services [7]. They employ hierarchical clustering algorithm to group users and services based on the location details and QoS values. The clustering output can be used in recommendation systems. Themain limitation is that the system is less scalable.

Different from existing work, this paper proposes a system to predict the QoS values of web service by considering purely the assessment of individual quality factors pertaining to that web service. First, the dynamic factor response time is calculated based on the users and providers details and interprets the QoS values from both

user's perspective and web service perspective. Throughput is predicted using the information provided by the service provider. Then the static factor security is calculated by considering the service specification of the web service and their vulnerabilities. The error correction method (RMSE) used to evaluate the accuracy of the predicted results. Thus, the proposed system predicts the optimized QoS of the web service.

3. METHODOLOGY

In the highly competitive information era, large numbers of similar web services are available and user feels it very difficult to find best web services. The web service QoS prediction is able to solve the problem, as it helps the user to identify the QoS value for the web services he/she is interested. Ultimately, this helps the user to find the web services with high QoS value

A. Qos Prediction Based on Dynamic Factors

For the implementation of the QoS Prediction based on dynamic quality factor response time, dataset with user details and web services details are maintained. User dataset includes the information of location details (latitude and longitude) of the users, and the details of QoS value experienced by the past users for different web services. Web service dataset includes the details of QoS values specified by the Service Providers for their services.

A.1. Region creation

First step in QoS prediction is to group users and web services based on thelocation information and QoS values of web service respectively. User regions are created to group the users who are closely located with each other according to their location details (latitude / longitude).Web service regions are created to group the services with similar QoS values specified by the service providers. The proposed system employs K-means clustering algorithm for creating user and service region.

K-means is a unsupervised learning algorithm which is used to cluster objects into K number of group.is a positive integer $K < n$ and where is the number of objects. Initially the objects are randomly placed into clusters. Then, a centroid is calculated for each cluster and each object's distance from the centroid is measured using Euclidean distance formula and evaluates the distance. If an object is found to be closer to another cluster then assign the object to that cluster. Recalculate the new cluster and re-compute the distance between each object and new obtained clusters. Repeat these until stability or convergence is achieved.

A.2. Region centers

Region center, which is an important feature in QoS prediction, is employed for both user region and service region. It is calculated for each user region and WS region separately and it is computed by consuming different web services QoS values specified by the experienced users pertaining to that region. A user region center is calculated using the median of all QoS values of the users in that region. Table 1 contains the details of five users and five services, which represents the response time of the five web service observed by the users in milliseconds. Suppose a user region consists of Ben, Charli and Darwin. The region center values for that region with respect to response time is (620, 2600, 600, 1700 and 2000).

Table 1. User details

User	*Location*	*S1*	*S2*	*S3*	*S4*	*S5*
Antony	UK	20000	Null	2000	Null	Null
Ben	US	200	3000	Null	3000	2000
Charlie	US	250	2600	200	Null	Null
Darwin	US	220	2300	1000	400	Null
Edwin	UK	1000	2000	1000	4000	Null

A service region center represents the average QoS value of a set of web service in that region, experienced by differentusersand it is calculated as the median QoS value of all those web services. Suppose the web service region consists of three services, the service region center with respect to response time is (20000, 3000, 1625, 620, and 2000). The region center values for that region reveals that the average response time of for(Antony, Ben, Charli, Darwin and Edwin) is (20000, 3000, 1625,620 and 2000) respectively

A.3. Sensitive web service

It is observed that the response time value vary for each user region (QoS fluctuation) and certain services have unexpected response time. It is necessary to identify such web services having unstable performance. We propose a method to identify these services and regarded such services as sensitive web services.If a user realized QoS greatly deviates from other QoS values that web service is considered as a sensitive web service and is given more attention. The user response time {600, 620, 650, 1000, 20000}. The values for service 1 from the Table 1 have the values. The values reveals that first user Antony's response time (20000) is greatly differ from 650 (Median value). In order to find sensitive web services we use two measures: (Median *m*) and Standard deviation (s).If QoS > m + 3s If then that service is sensitive to that user region. Here we find that $m = 650$ = and s = 50. It is clear that 20000 > 650 + 3* 50 and service s1 and service s1 is sensitive to the Antony's region.

A.4. QoS Prediction

Web service QoS value prediction is an important factor as quality of the whole system is dependent on the services provided by the system. Our system predicts QoS values from both users and service point of view. Then, it combines both the prediction results to conclude the response time quality value for that web service.

A.4.1. Prediction from user perspective

For QoS value prediction of web service from user's perspective first, identify the cluster in which user belongs. Then check whether the service is sensitive or not to the user's region. QoS of sensitive services are predicted from the region center. That is if services is sensitive to the region of user e then the QoS value ($q_{e,s}$) for servicesis the region center value ($q_{c,s}$) of the region in which user *e* reside.

$$q_{e,s} = q_{c,s} \tag{1}$$

If the service is non-sensitive, then the similarity between other users in that region is considered. System evaluates the similarity between enrolled user and other users in the region. Enrolled user (*e*) is the user who currently invoked the web service.Different similarity measures are accessible but our system employs Pearson Correlation Coefficient (PCC). PCC measures the similarity between enrolled user and every other users in the same region.

$$\text{sim}(e,u) = \frac{\sum_{i \in I}((q_{e,i} - \bar{q}_e) - (q_{u,i} - \bar{q}_u))}{\sqrt{\sum_{i \in I}((q_{e,i} - \bar{q}_e)^2}\sqrt{\sum_{i \in I}((q_{u,i} - \bar{q}_u)^2}} \tag{2}$$

where e represents the enrolled user and u represents the user who belongs to the cluster of enrolled user and sim (*e*,*u*) measures the similarity between them. I represents the services accessed by both users. $q_{e,i}$ and $q_{u,i}$ are the QoS value observed by enrolled user and other users for the web service *i* respectively. $\bar{q}_e$ and $\bar{q}_u$ represents the average QoS values observed by both users *e* and *u* for different web services. User with the highest PCC is considered and the QoS value given by that user is assigned as the QoS value for the enrolled user to that non-sensitive web service.

A.4.2. Prediction from service perspective

Considerservice *s*,QoS value for service *s* is calculated by considering the web service region center value of user e and service with highest similar value.

$$q_{e,s} = \text{average} ({}_{qe,c} + q_{e,m}) \quad (3)$$

$q_{e,c}$ represents the region center value of the enrolled user and $q_{e,m}$ is quality value of the enrolled user of the web service *m*, where

$$m = \text{WS withmax} (\text{sim} (j, k)) \quad (4)$$

Here *j* and *k* represents two web services and sim (*j*, k)is the similarity between them and it is calculated as

$$\text{sim}(j,k) = \frac{\sum_{u \in U}(q_{u,j} - \bar{q}_j) - (q_{u,k} - \bar{q}_k)}{\sqrt{\sum_{u \in U}(q_{u,j} - \bar{q}_j)^2}\sqrt{\sum_{u \in U}(q_{u,k} - \bar{q}_k)^2}} \quad (5)$$

whereU represents the users who have accessedservices *j* and *k* $\bar{q}_j$ represents the average QoS values of service *j* submitted by all users. $q_{u,j}$ and $q_{u,k}$ are the QoS value of the user *u* for the web service *i* and *j* respectively. If the users have no commonly accessed servicesthen the similarity becomesnull. After predicting QoS value from user perspective and service perspective, our system combines the prediction result which helps to achieve better results. Average of the two values will be taken for predicting the QoS value of response time for the web services.

QoS value of throughput is predicted in the same way as response time prediction, based on the QoS values provided by the Service providers.

B. QOS Prediction Based on Security

When deploying a web service, security is considered as an important issue. QoS prediction of security is highly essential to identify thesecured service with high confidentiality, reliability, integrity and authenticity[9]. QoS prediction of securityforweb service along with the response time and throughput helps the recommendation system to list out the optimal web services to the users. The system predicts the QoS value of security by consuming the security features and vulnerabilities of the WS. Security features indicates the measures implemented by the web service in order to protect the services and information handled by the WSfrom fraudulent users.

Table 2. Vulnerabilities in Web Services

Vulnerability	*Finance*	*Govt.*	*Health*	*Manu.*	*Technology*	*Retail*
Crypt Issue	60%	66%	61%	51%	62%	63%
Infor Leakage	58%	62%	60%	49%	62%	55%
Insufficient Input validation	41%	15%	13%	33%	37%	44%
Credentials Management	25%	20%	26%	24%	28%	24%

Vulnerability assessment of WS is essential to identify highly risky web services. If the percentage of vulnerability is above normal, then those web services are considered as risky web service. The system is designed to calculate the security factor based on the risk factors of the WS and the counter measures taken by them to handle those risk factors. The WS which handles all the vulnerable attacks in the best possible manner is considered as an optimal secured WS.

For the Prediction of security for a particular web service the system collects the security features of the web service as given below: Description of web services are available in the WSDL and the details of security algorithms implemented in the web services are specified in the WS-Policy. The security features are thus taken from the WS-Security policy specification. It reveals the details of the algorithm and purpose for which it is implemented.The details of the vulnerabilities is identified as follows: The detailed study about the vulnerabilities of web services is carried out to ascertain the percentage of different vulnerabilities for various domains of web services. The sample data of the vulnerabilities of web service is shown in Table 2. The domain of the given web service is identified from the web service description. Then the percentage of vulnerability for the different security factors of the web service is identified.

Importance of the Algorithms implemented in web service is essential to assess the quality of the counter measures taken to handle security threats. It is identified by taking into consideration the efficiency of the algorithm. Each algorithm is given values High/ Medium/ Low based on their efficiency to handle vulnerabilities. Table 3 shows sample crypt algorithms and theirsecurity levels. Moreover, in order include the importance of the threat handled each security levels is assigned a minimum and maximum value as shown in Table 4.

Table 3. Classification of Security Algorithm

Algorithm	*Security levels*
DES	LOW
3 DES	HIGH
AES	HIGH
RSA	MEDIUM
BLOWFISH	HIGH
TWOFISH	HIGH
RC5	MEDIUM
ECC	HIGH

Table 4. Security Levels

Security levels	*Min value*	*Max value*
High	5	6
Medium	3	4
Low	1	1

Evaluation of the vulnerabilities is performed to find out the impact of threat for each security factor of web service and the counter measures taken to handle them. Generally web service is prone to different threats, it is essential to identify the most risky threats. The risky threats are assessed as follows: the average of all the vulnerabilities for that web service domain is calculated, if the percentage of vulnerabilities of that security factor is above the average it is assumed to bea risky threat. The security value for each quality factor is assigned either minimum value (if it is not risky) or maximum value (if it is a risky threat) with respect to the importance of algorithm implemented to handle them.

Finally, the Prediction of Security for the web service is calculated as the average value of all the previously assigned value for different security factors. The web service with high predicted value (*i.e.*, above 4) is considered as high quality secured web service.

C. Prediction Generation

For predicting QoS value for a given web service the system takes into consideration both the dynamic and static quality factors. It predicts the quality values for response time, throughput and security. Average of above values will be the QoS of the given web service Q_w?

$$Q_w = \text{Avg}\left(\sum_{k=1}^{m} qk\right) \tag{6}$$

where, q_k is the quality value predicted for each dynamic and static quality factors, and m is the number of factors considered.

D. Prediction Accuracy

It is necessary to check the accuracy of proposed prediction approach. To assess the prediction accuracy, system employs Root Mean Square Error (RMSE) as the evaluation method. It makes an excellent general purpose metric for predictions.

$$\text{RMSE} = \sqrt{\frac{\sum_{w=1}^{n}(Q_w - \hat{Q}_w)^2}{N}} \quad (7)$$

where $\hat{Q}_w$ is the calculated QoS value based on the values given by the provider for the web service (w) and Q_w is the QoS value predicted by this system for web service w and N is the number of prediction.

4. EXPERIMENTAL RESULTS

The system is implemented with the dataset for users and web services. Dynamic quality factor prediction uses dataset with user details and web services details. User dataset includes location details (latitude and longitude) of the users and the details of the past users QoS values. Web service dataset includes the details of Service Providers QoS values. The response time and throughout are calculated for 10 different web services for a particular user. The security values for those web services are also calculated. Their results are shown in the Table 5.

Table 5. Results of Quality value Prediction.

Web services	*Security*	*Response time*	*Throughput*
1	5.25	0.799	1.392
2	4.75	0.222	1.592
3	4.5	0.204	1.506
4	3.5	0.203	1.316
5	4.25	0.396	1.536
6	4.5	0.712	1.676
7	3.5	0.605	1.363
8	2.5	0.4	1.758
9	5.5	0.591	11.239
10	2.5	0.427	1.402

The quality factor values in the above table are normalized and is it used for the analysis and display of the results for the system.

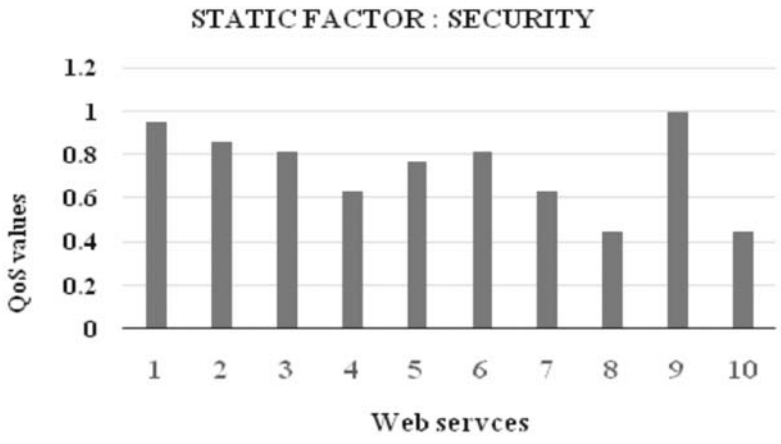

Fig. 2. QoS values of security for web services.

Figure 2 and Figure 3shows the predicted values for the static and dynamic non-functional quality factors for different web services.In Figure 2web service 9 is optimal in terms of security factor when compared to other services.

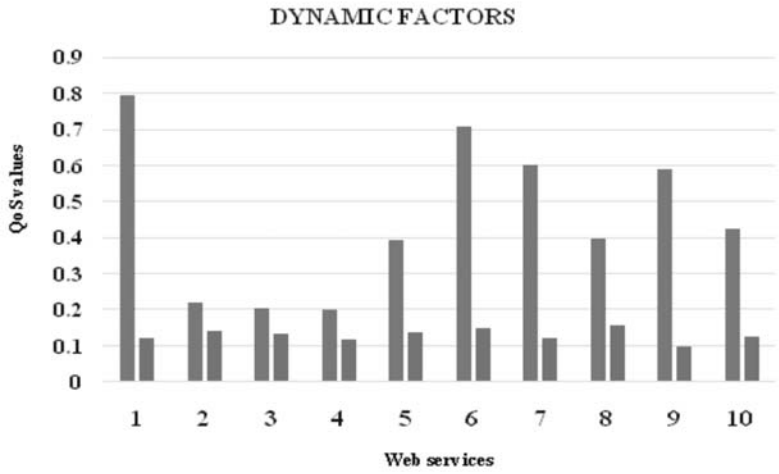

Fig. 3. QoS values of response time and throughput for web services.

In Figure 3 web service 1 and 8 are optimal in terms of response time and throughput respectively.

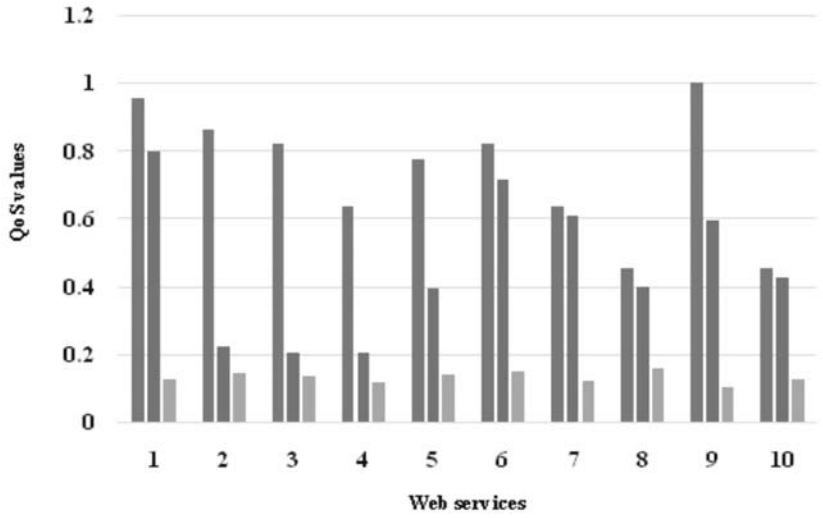

Fig. 4. QoS values of security, response time, and throughput for web services.

The above figure shows web services and the predicted QoS values of security, response time and throughput. Here the services 9, 1 and 6 are the optimal services in terms of response time, security and throughput respectively.

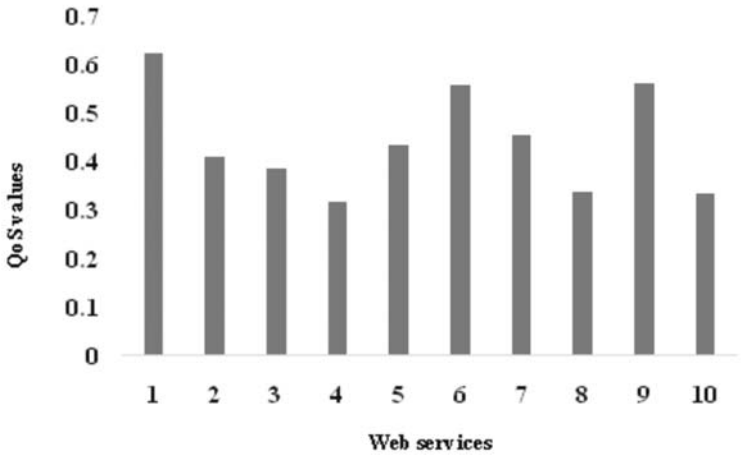

Fig. 5. Integrated QoS values of response time, security and throughput for web service.

The abovefigure shows QoS value of web service calculated by taking the average of the response time, security and throughput.The individual results of three factors are integrated to get enhanced result. Optimal web services are selected based on the predicted QoS values.Here web service 1 is the optimal web service.

5. CONCLUSION AND FUTURE WORK

In the present highly competitive world, there exists a huge number of web services which renders similar functionalities. It is more challenging for the user to select the best service. Our system helps the enrolled users to identify the best service by predicting the quality of WS. The system predicts the dynamic quality factors - response time, throughput and static quality factor - security value for the WS specified by the user. The dynamic QoS factor is calculated by i) clustering the users and web services based on location information, QoS values of the WS and calculates the cluster center values ii)finding the similaritybetween the users and web services in the cluster iii) predicting the quality factor value for response time with regard to users clustered data plus the web services clustered data and throughput based on the web services clustered data. The system predicts static QoS value for security based on the security methods implemented and vulnerabilities in the web service. Finally the prediction of QoS of WS is optimized by combing the quality value of both the dynamic plus static quality value of the given web service.

The present system predicts QoS values only for certain dynamic factors like response time and throughput in future other factors like availability, reputation would be calculated. In future singular value decomposition method will be used for filling the missing data present in the QoS values given by the providers and users so as improve the prediction result.

6. ACKNOWLEDGMENT

We are extremely thankful toall the faculty members of Department of Computer Science and Applications, Amrita Vishwa Vidyapeetham, Amritapuri for providing help and guidance. Our sincere thanks to Dr. M. R. Kaimal, Chairman, Computer Science Department, Amrita Vishwa Vidyapeetham, Amritapuri for his prompt support.

7. REFERENCES

1. Jian Wu, Liang Chen, Zibin Zheng."Predicting Quality of Service for Selection by Neighborhood-Based Collaborative Filtering"IEEE transactions on systems, man, and cybernetics: systems, Vol. 43, No. 2, March 2013
2. Jianlong Xu, Zibin Zheng, MichaelR. Lyu, "Web Service Personalized Quality of Service Prediction via Reputation-Based Matrix Factorization", IEEE Transactions on Reliability, vol. 65, No. 1, March 2016.
3. Chen Wu, WeiweiQiu, Zibin Zheng, XinyuWang and Xiaohu Yang "QoS Prediction of Web Services based on Two-Phase KMeans Clustering" 2015 IEEE International Conference on Web Services, pages 161 -168.
4. XinyuWang, Jianke Zhu, Zibin Zheng, Benjie song."A spatialtemporal QoS prediction approach for time aware service recommendation", ACM conference. Web Vol. 10, No. 1, Article 7, January 2016.
5. Alfredo Goldman, "On Graph Reduction for QoS Prediction of Very Large Web Service Compositions", Software Engineering, IEEE Transactions on, vol. 30, no. 5, Jul 11, 2012
6. Dragan Ivanovic, PeerachaiKaowichakorn, Manuel Carro, "Towards QoS Prediction Based on Composition Structure Analysis and Probabilistic Environment Models", ACM, vol. 54, no. 5, pp. 88-98, 2011.

7. Xi Chen, Zibin Zheng, Qi Yu, Michael R. Lyu, "Web service recommendation via exploiting Location and QoS information. IEEE transactions on parallel and distributed systems2010 Volume 25 Page 1913-1924,

8. Robert Warschofsky, Michael Menzel, Christoph Meinel. "Transformation and Aggregation of Web Service Security Requirements". EighthIEEE European Conference on Web Services Page No. 43-50 Dec. 2013.

9. MichaelMenzel, Robert Warschofsky, Christoph Mein., Pattern-driven Generation of Security Policies for Service oriented architecture" 2010 IEEE International Conference on Web Services. pp.243-250,

10. J.S.Breese, D. Heckerman, and Carl Kadie, "Empirical analysis of predictive algorithms for collaborative filtering", inProceedings of the 14th Annual Conference Uncertainty in Artificial Intelligence(UAI'98), pp. 43-52, 1998.

11. Dheeraj Kumar Bokde, SheetalGirase, Debajyoti "Role of Matrix Factorization Model in Collaborative Filtering Algorithm" International Journal of Advance Foundation and Research in Computer (IJAFRC) Volume 1, Issue 6, May 2014.

12. Mingdong Tang, Wei Liang, Buqing Cao and XiangyunLin, "Predicting Quality of Cloud Services for Selection", International Journal of Grid Distribution Computing Vol. 8, No.4, (2015), pp. 257-268.