International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

A Novel QoS-Based Approach for Web Service Recommendation and Visualization

Pradip N. Shendage[1], Santosh A. Shinde[2]

[1] PG Student, Department of Computer Engineering, VPCOE, Baramati, Maharashtra, India

[2] Assistant Professor, Department of Computer Engineering, VPCOE, Baramati, Maharashtra, India

Abstract: ***Web Services are becoming increasingly popular due to their subtle advantages. A Web Service is a standards-based software entity that accepts particularly formatted requirements from other software entities on remote machines via vendor. In service recommendation; the QoS-based approach is becoming more important. A lots of research has been studied on service recommendation but the performance of previous one is not suitable: Existing approaches is unsuccessful to consider QoS discrepancy according to users locations; and existing recommender systems have provided only limited information on the performance of the service candidates. This paper proposes a novel collaborative filtering algorithm planned for large-scale web service recommendation. This approach provides work for the characteristics of QoS by clustering users into different regions. With the help of recommendation visualization technique, this system will present that how a recommendation is grouped with other choices***.

Keywords: Service recommendation, QoS, collaborative filtering, self-organizing map, visualization

1. Introduction

Web Service: A Web Service is a standards-based software entity that accepts particularly formatted requirements from other software entities on remote machines via vendor. It also provides the communication protocols for producing particular application response". All applications are based on web, this applications are open to use, XML-based standards. With the use of transport protocols and it replace the data with each other. The Web service system contains two participants:
1) A service producer (provider)
2) A service consumer (requester)

The provider provides the interface and implementation of the service to requester with the help of HTTP protocols. The requester uses the Web service content. Nowdays; the web based applications are used in businesses and industry.

Service Oriented application: The Service oriented applications system contains three participants:

A registry, it behaves like a broker for Web services. A provider, it distributes the services to the registry. Finally A consumer, it can find out the services in the registry. The service user normally request to the provider through the service broker for list of web services. After getting the list of services from service broker, there is a need to identify the optimal one from functional equivalent candidates. It is very hard to find out optimal one from the list because service user has limited knowledge about the performance of services.

Quality of service: Quality of service is corresponding to the non functional performance of web service. And it has issued service selection as key factor.QoS is defined as set of user perceives properties. The properties like response time, availability, and reputation etc sometimes the QoS properties are not easy to users to attains the QoS information by evaluating all the service candidates, since it taking real-world web service invocations but this invocations is to taking long time and resource-consuming. And some QoS properties are critical to find out like reputation and reliability etc because this property requires the long time observation and chant.

2. Literature Survey

A. Collaborative Filtering

Z. Zheng, H. Ma, M.R. Lyu, and I. King [1] some recommender system uses the collaborative filtering such as Amazon.com. For prediction, we can use the CF algorithm. It supports to predict and recommend the superior item for a specific user from the web service recommendations. Normally, CF has contained some attribute such as users list and list of item and users rating on items. The relationship between users and items are shown in User-item matrix. Based on rating score, these items are arranged in ascending order. The user selects the particular items based on these rating score. The rating score has fixed range, like 1-5.Hence users only share their preferences on small number of items but this matrix is very insufficient. Sometimes CF has worked on User-Item matrix [2].

J.S. Breese, D. Heckerman, and C. Kadie [2] worked on Collaborative filtering and recommender systems .For recommender system, they have used the database regarding to user preferences. Based on preferences, they have calculated some new subjects or finds the new user might similar. Here, they have included another task which is based on correlation coefficients, vector-based same calculations, and arithmetical Bayesian methods. Collaborative filtering algorithm is distributed into two different classes:

In Memory-based algorithms, it made the some expectation over the whole user database. Normally, a Memory-based algorithm is used to predict the votes of a specific user from a

database of user elections from a section. It s similar to user based KNN. It is very simple to implement. It requires little or small training cost. It can easily add the fresher user into their rating score account but it cannot cope well, when large numbers of users are coming into picture. Since online performance is taking long time. Otherwise model-based algorithms are similar to K-means clustering, Bayesian model etc. It can develop the recommendations. It achieves the good online performance.
It can cope well, when large numbers of users are coming into picture.

B. Web Service Recommendation

Dong et al. [4] researched on key word-based web service search. For web service belongs to Woggle search engine, they delivered the similarity algorithm but it is insufficient for web services.

Liu et al. [5] have researched similarity measurement of web services. It made the graph based search model. This model helps to search the web services with related operations.

Mehta et al. [6] have researched on semantics and syntax based web services. It is not sufficient to search the services and number of times it does not meets the user's requirement. Hence, it used two more dimensions for service description such as quality and usage pattern. Based on service description, they delivered service mediation model.

Blake and Nolan [7] have designed the recommendation score for web. They worked on operational sessions of web service. Based on score it will provide the best possible strings from the users operational sessions and the description of the web services. They can easily discover how many users are involved in the service based on score. Maamar et al. [8] have researched on context based web services. It represented the some resources on which the web service processed. It delivered more interaction with web services.

C. Self-organizing map (SOM)

Tasdemir and E. Merenyi[9]the self-organizing maps (SOM) method useful for visualization, cluster mining, and data mining. When high dimensional data comes that time we can use SOM. SOM is good for high dimensional data. It also includes the data structure and finds the cluster limitations from the SOM. Here, the best method is to differentiate the SOM's knowledge by using visualization methods. The previous methods have given the bad performance on SOM knowledge with considered the Data topology. They have researched on data topology that combined into the visualization of the SOM and can deliver an extra intricate judgment of the cluster structure than current arrangements. They have presented a weighted Delaunay triangulation.

3. Implementation Details

The proposed framework is organized into three stages Region creation, QoS value prediction, and Recommendation Visualization respectively. The system architecture is:

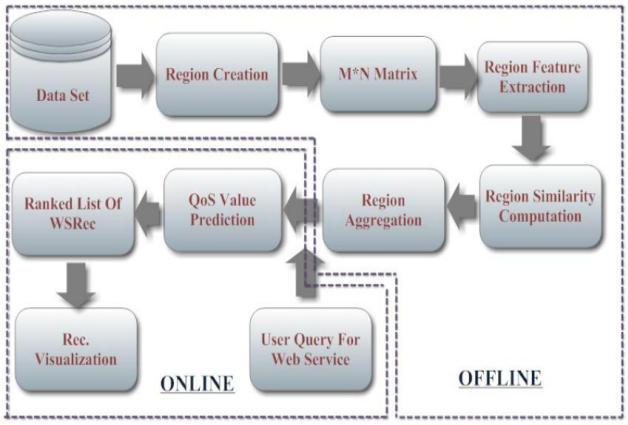


Figure 1: System Architecture

Phase 1: Region creation:

A region is defined as set of users who are closely situated with each other with their similar QoS profiles. For region, every users have their at least one region.In web service recommender system, users normally deliver the QoS values on a small number of web services. A usual memory-based CF algorithm travels from the some problem such as users contributed data set. Existing memory-based CF algorithm has hard to discover similar users without knowledge of their service experience. To address of above solution, we take the relationship between users' physical locations and QoS properties. Here we concentrate on the QoS properties those changes rapidly. It can be easily calculated by individual users, such as response time and availability.RTT means the time period between users sends the request and receive the request from server. Here we assume that there are number of users 'N' and number of services'M'.The relationship between users and services represented in the N*M matrix R. Every entry of Ri,j represent the RTT values of service j noticed by user i.Each user i associated with row vector Ri denotes the RTT values noticed on the different web services. The user a is called the active user or current user if he/she has delivered some RTT records and wants service recommendations.

There are three steps needs to create a region are as following:
1. To extract the feature from the region.
2. To find out the similarity between regions.
3. To aggregate the highly correlated region from small regions.

A. Region Feature Extraction:
For each region, we use the region center. Region center plays the important role in region feature extraction step. Region center returns the average performance of services noticed by region users. Region center is defined as the median vector of all the RTT vectors associated with the region users. The performances of services are changed by region by region. From large number of QoS record region, the service performance will vary from region to region such as service response time and availability. Some services give the unexpected long response time to specific region. To solution of this problem, we use the Three sigma rule which will analyze to differentiate services with unstable performance to different regions. The region sensitive services is a another important feature besides the region center..The set of nonzero RTTs of service s, composed from

users of all regions. It is a sample from the population of service s response time. To estimation, the mean and the standard deviation of the population, we use two strong measures: median and median absolute deviation. The median of the absolute deviations from the sample's median called as MAD.

$MAD = median(|R_{i,s} - median_j(R_{j,s})|), i = 1,...,k, j = 1,...,k.$ (1)

Based on MAD, we get

$\mu = median\ (R_{i,S})\ i = 1,...,k,$

$\boldsymbol{\sigma} = MAD_i\ (R_{i,S}) i = 1,...,k.$

There are three important features for QoS predication and recommendations are as follows:

1) Region-Sensitive Service means same service used in two different region but their RTT values changed..
2) Region Sensitivity. The sensitivity of region defined as the fraction between the numbers of sensitive services in region over the total number of services.
3) Sensitive Region. Region is a sensitive region iff its region sensitivity exceeds the sensitivity threshold.

B. Region Similarity Computation:

In Region Similarity Computation, to find out the two similar regions is key factor. Before the aggregation step we need to find out the similar region and similar users are must. Here, This Recommender system uses the Pearson Correlation Coefficient (PCC).With help of PCC; we can find the similarity between two users. The similarity between two users is calculated by the similarity of their region centers. PCC will return the two values +1 and -1.The value +1 represents the two users have similar preferences, while negative PCC value means that the two user preferences are different.

$$Sim(m,n) = \frac{\sum_{s \in S(n) \cap S(m)} (R_{m,s} - \bar{R}_m) \cdot (R_{n,s} - \bar{R}_n)}{\sqrt{\sum_{s \in S(n) \cap S(m)} (R_{m,s} - \bar{R}_m)^2} \sqrt{\sum_{s \in S(n) \cap S(m)} (R_{n,s} - \bar{R}_n)^2}}$$

Where $S(n) \cap S(m)$ is the set of same services used by users from region M and N,.

$R_{m,s}$ and $R_{n,s}$ is the RTT value of service s delivered by region center m.

$\bar{R}_m$ and $\bar{R}_n$ denotes the average RTT value of all the services of center m and n, respectively.

PCC consider the RTT values of similar services are used in both regions n and m. PCC will not only consider the similar region but also consider the similar RTT valves of services.

C. Region Aggregation

After the Region Similarity Computation, we need to aggregates the highly correlated regions. With based on users' physical locations, we can form the each region at the QoS data set. Since users only make use of a small number of web services and delivered the limited QoS records. There are two difficulties:

1. Difficult to discover similar users.
2. To predict the QoS values of the unused web services for the active user.

To address of above problem, we use the region aggregation method .Which is based on the region features.

In region aggregation, we use the K-means recursive algorithm.

Steps for K-means recursive algorithm are as follows:

Step1. To select the K. .Assign values to K.

Step2.To finds out the vector median based on K values.

Step3.Based on vector median, we aggregate the similar region.

Step4. Stopping creation.

Phase 2: QoS Value Prediction:

After region aggregation, Based on users physical locations and historical QoS similarities, number of users is clustered into a certain number of regions. The region center is denoted in the service experience of users in a region .With the compacted QoS data, searching neighbors may be computed fastly. Building predictions for an active user can be computed quickly with the help of QoS data. The similarity between the active user and users of a region is calculated by the similarity between the active user and the region center. To predict the QoS value for active users based on their regions, for users in the same region to have similar QoS experience on the same web service, particularly on those region-sensitive ones. In Prediction, When active users want RTT record of services which has unused the before, for that we need to take following steps:

Step1. Find out the specific region for active users. If it is not found than active user will be treated as new member of region.

Step2. Classify whether service s is sensitive to the specific region or not. If it is region sensitive, then the prediction is created from the region center, because users are noticed the service performance from this region which is considerably different from others.

$\hat{R}_{a,s} = R_{center,s}$

Step3.Otherwise, use PCC formulae to calculate the similarity between the active user and each region center that has estimated service s.d To discover k most similar centers.

Step 4: If the active user's region center has the RTT value of s, the prediction is calculated by using below equation:

$$\hat{R}_{a,s} = R_{center,s} + \frac{\sum_{j=1}^{k} (R_{cj,s} - \bar{R}_{cj}) Sim'(a,c_j)}{\sum_{j=1}^{k} Sim'(a,c_j)}$$

Where Rcj;s is the RTT of service s delivered by center cj, Rcj is the average RTT of center cj.

The prediction is calculated in two parts:

1)The RTT value of the region center of the active user Rcenter;s, which represents the average QoS noticed by this region users.
2)The normalized weighted sum of the deviations of the service s RTT from the average RTT noticed by the k most similar neighbors.

Step 5: Otherwise, we use the service s RTTs examined by the k neighbors to calculate the prediction.

3. Recommendation Visualization

Visualization technique is used to browse the some services on map. With the help of map ,user can easily understand the performance of services. Here we used the QoS space visualization of all web services on a map will behind the recommendations. QoS space visualization is method of computing. It covers the information of high dimensional QoS data into a visual form enabling service users to observe, browse, and understand the information.

There are two steps for QoS map:

1. Dimension reduction step
2. Map creation step.

In Dimension reduction step, we make a 2D representation of the high-dimensional QoS space by using self-organizing map (SOM), and each web services is represented to a unique 2D coordinates.

In Map creation step ,we generate a geographic-like QoS map depends on the SOM training results.

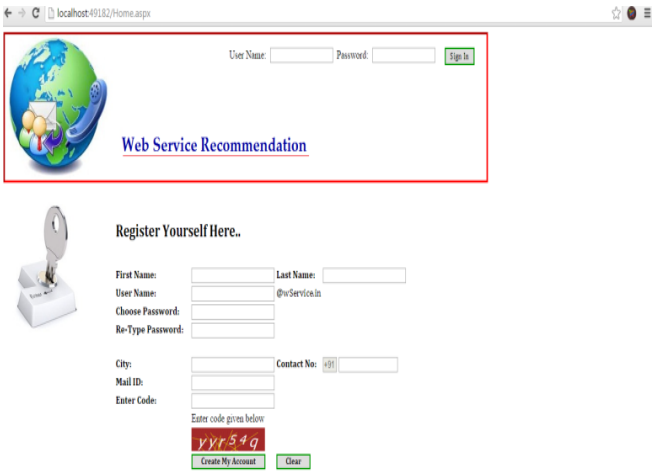
4. Results and Dataset

Web service QoS dataset was evaluated in the context of four different data sets collected from following web site: http://www.wsdream.net.

Dataset contain some information about the web service invocations records. It contains 100 Web services which have collected from the 20 regions. The dataset contain some attribute such as list of users, list of web services and RTT values of particular services etc. In users list, they have provided some attribute such as users location, country, latitude and longitude of specific region. In web service list, they have provided some attribute such as Address of web services, country, providers, Cost etc. Dataset contains 339 users and 5825 Web services.

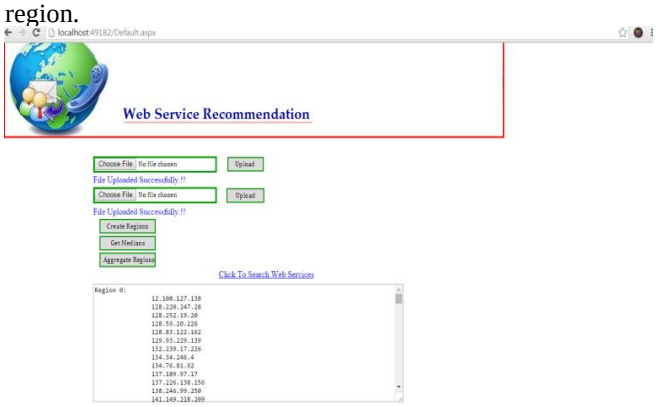
Results

1.Home page of System:

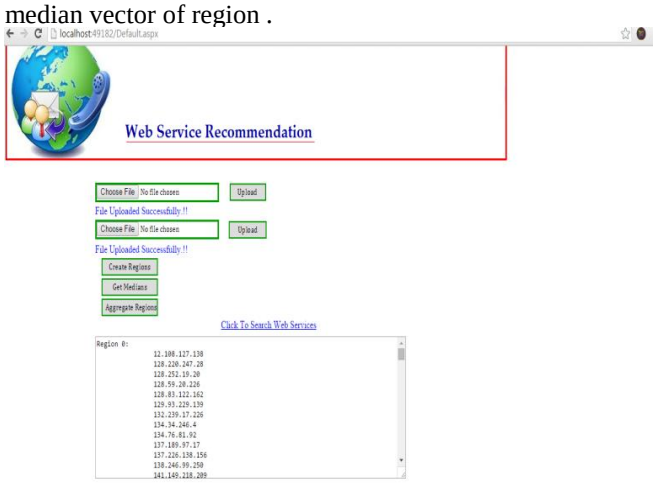


2.After login

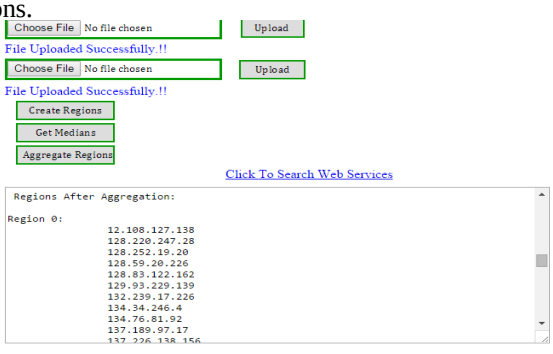
3. After clicking on create region button, we will get the region.



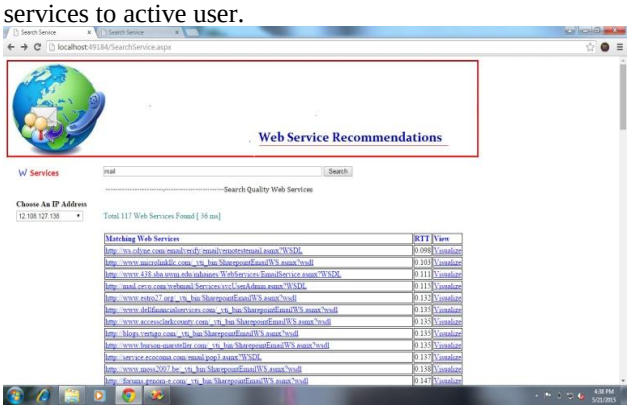
4. After clicking on get median button, we will get the median vector of region .



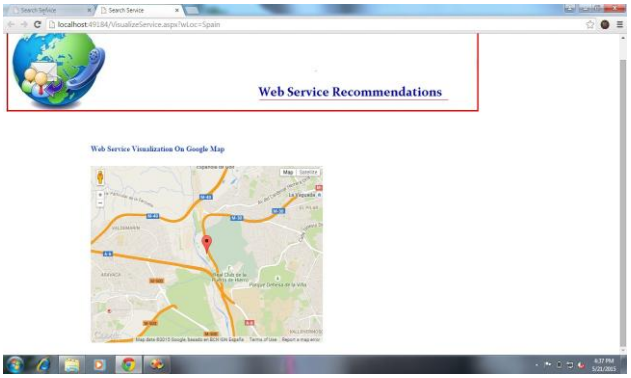
5.Region after aggregation.here we get the aggregated regions.



6. After region aggregation ,here we predict RTT value of services to active user.



7. After predication step,we will get the visualize services in 2D co-ordinate.



5. Conclusion

The proposed system algorithm provides the characteristic of QoS by clustering users into different regions. A nearest-neighbor algorithm is proposed to produce QoS prediction established on region feature. In recommendation approach, it included the relationship between QoS records and users physical locations with IP addresses; it has obtained the good prediction performance as compare to existing methods.

6. Acknowledgments

I express great many thanks to Prof. S. A. Shinde for his great effort of supervising and leading me, to accomplish this fine work. I am thankful to our Head of the Department Prof. Mrs. G. J. Chhajed and Prof. Mrs. S. S. Nandgaonkar, (P. G. Coordinator) for their moral support and motivation. I would also like to thank our Principal Dr. S. B. Deosarkar who encouraged us and created a healthy environment for all of us to learn in best possible way.

References

[1] Xi Chen, Member, IEEE, Zibin Zheng, Member, IEEE, Xudong Liu, Zicheng Huang, and Hailong Sun, Member, IEEE "Personalized QoS-Aware Web Service Recommendation and Visualization,"JANUARY-MARCH 2013.

[2] Z. Zheng, H. Ma, M.R. Lyu, and I. King, "WSRec: A Collaborative Filtering Based Web Service Recommendation System," Proc. Intl Conf. Web Services, pp. 437-444, 2009.

[3] J.S. Breese, D. Heckerman, and C. Kadie, " Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Arti_cial Intelligence (UAI 98), pp. 43-52, 1998.

[4] X. Dong, A. Halevy, J. Madhavan, E.Nemes, and J. Zhang,"Similarity Search for Web Services," Proc. 30th Intl Conf. Very Large Data Bases, pp. 372-383,2004.

[5] X. Liu, G. Huang, and H. Mei, ―Discovering Homogeneous Web Service Community in the User-Centric Web Environment,‖ IEEETrans. Services Computing, vol. 2, no. 2, pp. 167-181, Apr.-June 2009.

[6] B. Mehta, C. Niederee, A. Stewart, C. Muscogiuri, and E.J. Neuhold, ―An Architecture for Recommendation Based Service Mediation,‖ Semantics of a Networked World, vol. 3226, pp. 250-262,2004.

[7] M.B. Blake and M.F. Nowlan,"A Web Service Recommender System Using Enhanced Syntactical Matching," Proc. Intl Conf.Web Services, pp. 575-582, 2007.

[8] Z. Maamar, S.K. Mostefaoui, and Q.H. Mahmoud,"Context for Personalized Web Services,"Proc. 38th Ann. Hawaii Intl Conf., pp. 166b-1620056b,

[9] K.Tasdemir and E. Merenyi, "Exploiting Data Topology in Visualization and Clustering of Self-Organizing Maps,"

Author Profile

Pradip Shendage received the B.E degree in Information Technology from K.G.C.E,karjat under the Mumbai university in 2012 and pursuing M.E degrees in Computer Engineering from Vidya Pratishthan's College of Engineering, Baramati under the Savitribai Phule Pune University.In B.E ,I have got 63% with first class.

Prof. Santosh Shinde received his B.E. degree in computer engineering (First Class with Distinction) in the year 2003 from Pune University and M. E. Degree (First Class with Distinction) in Computer Engineering in 2010 from Pune University. He has eleven years of teaching experience at undergraduate and postgraduate level. He has attended various National and International Conferences, Seminars and Workshops on various subjects like Multimedia Techniques, Multicore Computing and Software Engineering. He is a life member of ISTE (Indian Society for Technical Education) and IACSIT (International Association of Computer Science and Information Technology). He has worked as a review committee member for the 1st Internaional Conference on recent trends in Engineering and Technology (ICRTET'2012) held at SNJB's COE, Nashik. He has worked as a Judge for the State level paper presentation (REBEL) held at SVPM's COE, Malegaon. He has attended workshops on Robotics, Research paper writing and Latex conducted by IIT, Bombay remote center at VPCOE, Baramati. He has to his credit IBM RFT certification with 100% score. He has organized one day workshop on Software Engineering by Dr. S. A. Kelkar, adjunct Professor IIT, Powai.