International Journal of Scientific & Engineering Research, Volume 4, Issue 11, November-2013
ISSN 2229-5518

QSWSP: QoS Based Scheduled Web Service Positioning

Reena MalikNisha Malik
Department of Computer ScienceDepartment of Computer Science
Uttarakhand Technical university,Dehradun,IndiaTula's institute Dehradun, India

ABSTRACT-There has been a huge increase in the scale of distributed network service with the rising demand of service-oriented Architecture (SOA), more and more alternative web services offered by different service providers becomes available over the internet to provide equivalent or similar-function for service user. It is estimated that there would be large number of web services developed, deployed in the internet in next several years. Earlier approaches proposed that two major properties should be taken into consideration that are functional and non-functional (QoS) when user selects services. There are number of different service provider and we have to effectively select the optimal one from web service candidates. In our approach the requests send by the user are scheduled on the basis of Earliest Deadline First. Then by combining the advantages of network coordinate based approach, simple regression, clustering approach and WSP approach, the response time between user and web services can be accurately predicted.

Index Terms-Scheduling and Web Positioning of web service, Simple Regression and Clustering, Network Coordinate

1. INTRODUCTION

Web services are designed as computationalcomponents to build service-oriented distributed systems, such as e-commerce, automotive system, multimedia services etc[1]. With the emergent number of Web services, it has become an urgent task to make operative selection from the large number of functionally-equivalent Web service candidates. Since there is rise in the demand of web services and many web services are available, so we have to effectively select the optimal web service for the user. To understand the effective use of services , it is not only sufficient to comprehend the semantics of the integrated services or find matching for mandatory positioning of services but also to take their Quality of Service (QoS) properties into account for service positioning. A service that does not provide an adequate QoS might be as useless as a service not providing the desired efficient results.

In this paper we focus on response time out of different QoS properties. It is impractical for every user to measure QoS performance of all the web service as there is large number of service candidate on the web. Recently a number of QoS prediction approaches have been proposed [2-5]. A prediction technique in this context can be explained as a methodology that analyzes historical QoS data from previous users of particular services and uses it to predict the QoS that will be experienced by a current user on the particular services [6]. Simple regression and clustering technique is used in this paper for QoS prediction. Network Coordinates Systems(NCS) have been proposed [7]to allow hosts to estimate delays without performing direct measurements and thus, reduce the consumption of network resources. The key idea of an NCS is to model the Internet as a geometric space and characterize the position of any node in the Internet by a position (i.e., a coordinate) in this space. The network distance between any two nodes is then predicted as the geometric distance between their coordinates. Explicit measurements are, therefore, not anymore required. Network coordinate system approach is used to calculate the coordinate. When the users request more than one web service at a time then we first schedule them according to basis of earliest deadline first.

Inspired by the success of network coordinate based prediction approaches and to address the process with minimum deadline problem, we propose a QOS based scheduled web service positioning novel framework by combining the advantages of earliest deadline scheduling and network coordinate based approaches.

Further we calculate response time for the first web

International Journal of Scientific & Engineering Research, Volume 4, Issue 11, November-2013
ISSN 2229-5518

service in the schedule queue as QOS properties.

2. RELATED WORK

2.1 Earliest Deadline First Scheduling

Earliest deadline first (EDF) or least time to go is a dynamic scheduling algorithm used in real-time operating systems. Processes are placed in a priority queue manner. Whenever a scheduling event occurs the queue will be explored for the process closest to its deadline.

This process is the next to be scheduled for execution. EDF is an optimal scheduling algorithm on preemptive uniprocessors, in the following sense: if a collection of autonomous jobs, each categorized by an arrival time, an execution requirement, and a deadline, can be scheduled such that all the jobs complete by their deadlines, the EDF will schedule this group of jobs such that they all complete by their deadlines.When service user request more than one web-service at a time then we first schedule them according to the earliest deadline first basis. Response-time of earliest deadline web service is calculated, then second earliest deadline is calculated and so on. We are using non-preemptive earliest deadline first algorithm. A processor following this algorithm always executes the task whose absolute deadline is earliest.

Table 1 Web Services in queue

WEB-SERVICES	DEADLINE
W1	3
W2	2
W3	10
W4	6

After scheduling

Table 2 Web Services in a queue after EDF scheduling

WEB-SERVICES	DEADLINE
W2	2
W1	3
W4	6
W3	10

Table 1 and 2 depicts the web services as a process in a queue before and after the earliest deadline first scheduling. First response time of W2 is calculated and so on.

2.2 Simple Regression and Clustering approach

It is a data mining approach where recorded observations are grouped into clusters on the basis of some similarity. Observations referred here is to the historical QoS related data. Proximity measure [8] or distance measure is used to measure the similarity. The Euclidean distance [9] is used as the distance measure in the research work[10-11] present in this section and is calculated between two QoS vectors q1=(x1,x2,x3) and q2=(y1,y2,y3) using the Pythagorean formula i.e D= the minimum distance[9] between two clusters A and B is given by: min{D(x, y): x in A, y in B}. Agglomerative or hierarchical[9] or Partitional clustering algorithms is used. With simple regression analysis technique[11] computed, predict QoS value by taking as input average value for QoS of the service for time instance in the past.

2.3 Network Coordinate System

The network coordinate system isproposed in [12] to estimate the network distances, i.e., round triptime (RTT), between pairwise Internet hosts. Triangulated heuristicand global network positioning (GNP) are two widely employedapproachesamongvarious network coordinate systems, due to their simplicity and generality.

Triangulated Heuristic [12] services a kind of relative coordinates based on the triangle inequality. A fixed set of landmarks are deployed in the network as locations. Then each ordinary host is apportioned an n-tuple relative coordinate, composed of the network distances between the ordinary host and the landmarks. Given the relative coordinate of each host, we can obtain the upper bound *U*and the lower bound *L*of the network distance between two hosts by triangle inequality. The network distance can be expected by the convex combination of *U*and (e.g. U+*L*2). It is stated in [12] that taking the upper bound *U*as the network distance prediction result can achieve better performance. The triangulated heuristic approach is broadly used in online shortest path distance calculation in large graphs.

GNP [12] is a emblematic landmark-based network coordinate system, which embed the Internet hosts into an Euclidean space for network distance approximation. After locating the coordinate of each host, the network distance between two Internet hosts can be well approximated by the corresponding Euclidean distance. Figure 1 illustrates a prototype of the network coordinate system. As we can see from the figure, the four Internet hosts can be embedded into a 2-dimensional Euclidean space by assigning each host a coordinate, and then the original network distances can obtain good estimation results using the corresponding Euclidean distances.
In this paper we use GNP (global network positioning). First we embed hosts into a Euclidean system and after that we calculate the coordinate of each host. When we have calculated the coordinate of hosts, we can easily calculate the distance between them by using Euclidean formula that is used to calculate Euclidean distance.

Example host A (x, y) Host B (a ,b) ; distance between them can be calculated by the formula

$$\sqrt{(x-a)^2+(y-b)^2}$$

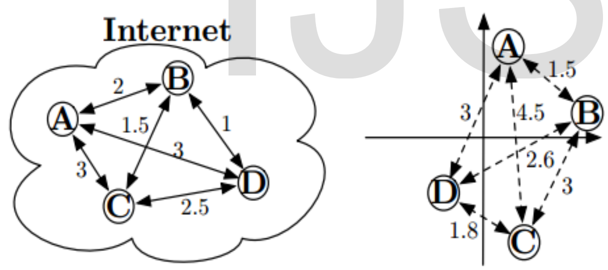

Figure 1 A prototype of Network Coordinate System [7]

2.4 Web Service Positioning Approach

This approach proposed a web service positioning framework to make response time prediction for web services.

In this framework, a small number of landmarks are deployed in the internet to construct the network coordinate system. This approach follows network coordinate based and CF- based approaches [11].The experimental results show that the earlier proposed WSP approach solves the data sparsity problem of CF-based approaches and significantlyenhances the prediction accuracy.

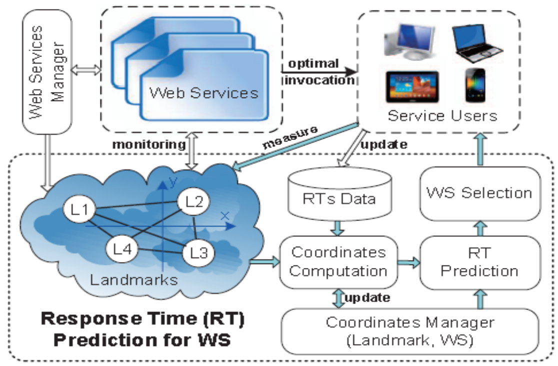

Figure 2: WSP Framework [11]

3. QOS BASED SCHEDULED WEB SERVICE POSITIONING (QSWSP) FRAMEWORK

We proposed a novel approach of web service positioning using earliest deadline scheduling in web services and calculating its response time. Our QSWSP framework combines the advantages of network coordinate based, simple regression and clustering approaches in EDF scheduled queue.

The response time is calculated using WSP based QOS algorithm [11] which includes Offline coordinate updating and online web services selection.

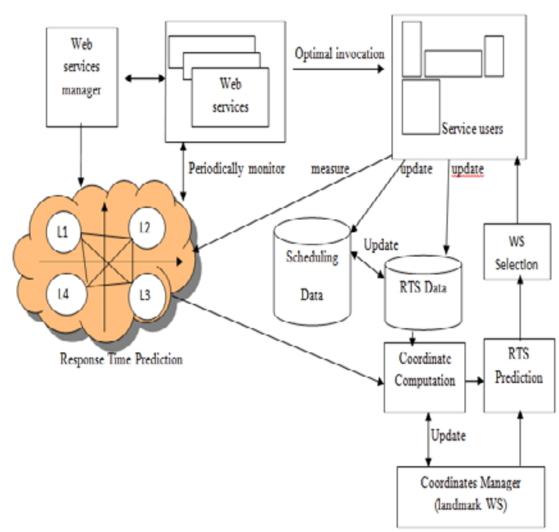

Figure 3 QSWSP Framework

Figure 3 depicts the framework of QOS based schedules web service positioning. The proposed novel approach will follow the following steps shown in figure 4.

After web services are scheduled in a queue using earliest deadline first scheduling, the WSP based QOS prediction algorithm [11] is followed in data gathered through simple regression and clustering method.

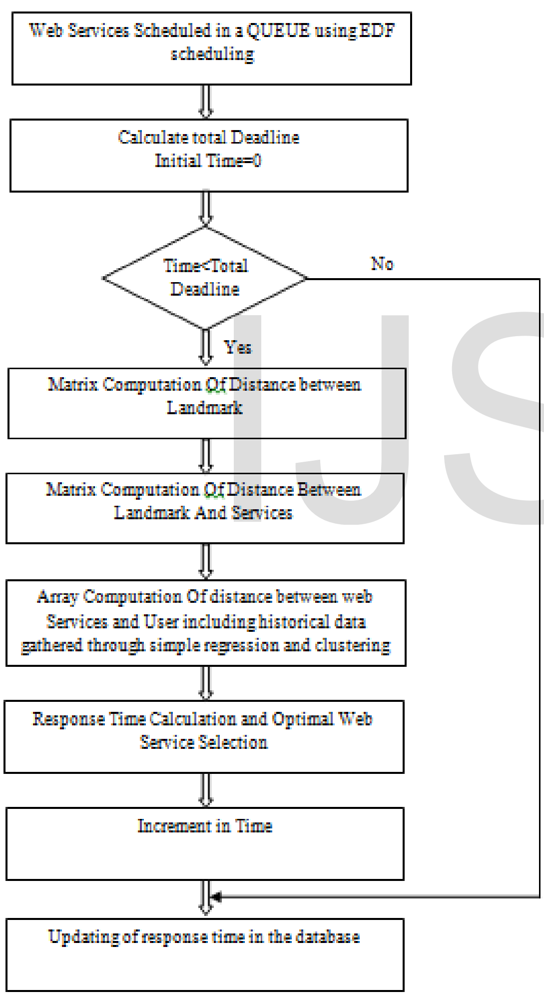

Figure 4 Flowchart depicting the working of QSWSP

1. Matrix Computation Of Distance Between Landmark

For n-dimensional Euclidean space, there should be at least n+1Landmarks. Problem that arises is that how landmarks should be deployed. We are using spectral clustering based approach [15] in this paper, m landmarks denoted by $L_e=\{ L_I , I = 1,2 \ldots m\}$ are deployed in the internet. Ping message is used to measure distance between landmarks and then transmitted to central node.

- *m*m landmarks distance matrix*

$$LD=\begin{vmatrix} 0 & d(l_1,l_2) & \ldots & d(l_1,l_m) \\ d(l_2,l_1) & 0 & \ldots & d(l_2,l_m) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ d(l_m,l_1) & d(l_m,l_2) & \ldots & 0 \end{vmatrix} \longrightarrow 1$$

Where entry $d(l_a, l_b)$ define the distance between landmarks l_a and l_b. matrix is to be assume symmetric along diagonal. Then we embed m landmarks' into an n-dimensional Euclidean space R^n and then each landmark obtain coordinate in Euclidean space denoted as $x^k_{la} = (x^1_{la}, x^2_{la}, \ldots x^n_{la})$ where $x^k_{la} \in R$, $1<=K<=n$

Function is define as the sum of error square i.e.

$$F_{l_e}(x_{l1},\ldots\ldots x_{lm}) = \sum_{k=1}^{m} [d'(l_a , l_b)-d(l_a , l_b)]^2 \text{ ----------2}$$

$$l_a,l_b \in L_e, a>b \qquad d'(l_a , l_b) = \| x_{la}-x_{lb}\|_2 = \sqrt{\sum (x^k_{la}- x^k_{la})^2} \text{ ---------------------3}$$

Regularization term is added to address unknown response time between users and web services.

$$F'_{L_e}(x_{l1} , \ldots\ldots x_{lm} , \mu_L) = \sum_{l_a,l_b \in L_e, a>b} [d'(l_a , l_b)-d(l_a , l_b)]^2 + \mu_L \sum_{k=1}^{n} \|x_{lk}\|^2_2 \text{ ---------------------4}$$

Simplex downhill algorithm [10] is used to solve minimization problem. Landmark should keep updated periodically to track the changes of network condition [11].

2. Matrix Computation Of Distance Between Landmark And Services

A small number of landmarks monitor the available web services by periodically invoking them[11]. Assume there are z available web services denoted by W={wi,i=1,2,…z}

Then m*z matrix is computed which include network distance between m landmarks and z web services.

$$LS=\begin{vmatrix} d(l_1,w_1) & d(l_1,w_2) & \dots & d(l_1,w_z) \\ d(l_2,w_1) & d(l_2,w_2) & \dots & d(l_2,w_z) \\ . & . & \dots & . \\ . & . & \dots & . \\ d(l_m,w_1) & d(l_m,w_2) & \dots & d(l_m,w_z) \end{vmatrix} \longrightarrow 5$$

Where entry d(la,wb) define the distance between landmarks la and web service wb. It is then transmitted to the central node.

3. Array Computation Of Service User Coordinate

Our QSWSP provide optimal web service selection to any web service requester. In the starting, the user measures the network distance between him and the landmarks using ping message. The result is transmitted to the central node for coordinate computation.

$D_{userl} = [d(su,l_1) \quad d(su,l_2) \quad \dots \quad d(su, l_m)]$. Where entry $d(su,l_a)$ define the distance between landmarks l_a and service user su.

$\{d(su,w_b),\ su \varepsilon\ S\}$ is the available historical data between service user and web services. S is the web service set which include available historical data. To obtain the coordinate of the service user (x_{su})

$$F_{su}(x_{su}, \mu_{su}) = \sum_{l_a \varepsilon L_e} [d'(su, l_a) - d(su, l_a)]^2 + \sum_{w_a \varepsilon S} [d'(su, w_a) - d(su, w_a)]^2 + \mu_{su} \| x_{su} \|^2_2 \text{-----------------------------} \textbf{-6}$$

Where entry $d'(su, l_a)$ define the distance between landmarks l_a and service user su.

Where entry $d'(su, w_a)$ define the predicted response time value between web service $w_a$in S and service user su.

4. Response Time Calculation and Optimal Web Service Selection

Now we calculated the Euclidean distance between coordinates of service user and web services which is equal to the response time.

$d'(su, w_a) = \|x_{su} - x_{wa}\|_2$, $w_a \varepsilon$ S, $w_a\ \varepsilon\ S_a$, means set of web services with unknown response time[11].

5. CONCLUSION AND FUTURE WORK

In this paper we propose a novel approach for QOS based scheduling method for web service positioning when there are number of request at a time .By combining the advantage in network coordinates based approaches earliest deadline first scheduling and simple regression and clustering approaches, our QSWSP is constructed to support response time prediction for service users.It can also serve users without available of historical data .In this paper using scheduling technique we only focused onresponse time. In future we will extend QSWSP model to be implemented using various experiments to other QoS properties. We will further study to improve the working of QSWSP using different data mining techniques.

6.REFERENCES

[1]L. Zhang, J. Zhang, and H. Cai, "Services Computing: Core Enabling Technology of the Modern Services Industry". *Tsinghua University Press*, 2007.

[2] Z. Zheng, H. Ma, M. R. Lyu, and I. King, "QoS-aware web service recommendation by collaborative filtering," *IEEETrans.on Services Computing*, Vol. 4, no. 2, Pp. 140–152, 2011.

[3] L. Shao, J. Zhang, Y. Wei, J. Zhao, B. Xie, and H. Mei, "Personalized QoS prediction for web services via collaborative filtering," in *Proc. of the IEEE ICWS'07*, Pp. 439-446,2007.

[4] X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, "Personalized QoS-aware web service recommendation and visualization,"*IEEE Trans. on Services Computing*, 2011.

[5] Y. Jiang, J. Liu, M. Tang, and X. F. Liu,"An effective web service recommendation method based on personalized collaborative filtering," in *Proc. of the IEEE ICWS'11*, 2011.

[6]Deng, X. and Xing, C,"A QoS-oriented Optimization Model for Web Service Group", *proceedings of the 8th IEEE/ACIS, Pp.* 903 – 909.
[7] Donnet, B., Gueye, B. and Kaafar,"A Survey on Network Coordinates Systems, Design, and Security", *Proc. Of Communications Surveys & Tutorials, IEEE* , Vol. 12 , Pp. 488-503,2010.
[8] W3C (2003). "QoS for Web Service: Requirements and Possible Approaches". *W3C Working Group* Note 25 November 2003. Retrived from http://www.w3c.or.kr/kr-office/TR/2003/ws-qos/
[9] Dubes, R., Jain, A., K. "Clustering Methodologies in Exploratory Data Analysis. Advances in Computers"(Edited by: Yovits C. M.), Vol. 19, *Academic Press*,2009.
[10] Vu, L.-H., Hauswirth, M., Aberer, K., "QoS-based service selection and ranking with trust and reputation management",*Proc. of the 13th International Conference on Cooperative Information Systems*, *Springer: Berlin, Germany*,2005.
[11] Zhu J., Kang Y., ZhengZ.,and Michael R.,"WSP: A Network Coordinate based web service postioning framework for response time prediction", *proc. Of IEEE 19th International Conference on Web Services*, 2012.
[12] T. S. E. Ng and H. Zhang, "Predicting internet network distance with coordinates-based approaches," *in Proc. of the IEEE INFOCOM'02*, 2002.