IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

An Approach for an Effective Web Service Selection using Filtering and Skyline Method

Shahidul Islam[1], S. Britto Ramesh Kumar[2], Ab Rashid Dar[3]

[1]Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli-620002, Tamil Nadu, India
[2]Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli-620002, Tamil Nadu, India
[3]Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli-620002, Tamil Nadu, India

Available online at: www.ijcseonline.org

Accepted: 20/July/2018, Published: 31/July/2018

***Abstract*--**In Modern World, an exponential growth of web services is been observed over the Internet, This offers a big challenge for the optimal selection of the best web service among a group of web services with similar functionalities. The selection process must be made in order to determine which relevant Web services would satisfy a user's needs. In this paper, a methodology for web service selection combines prefiltering followed by skyline selection. The K-Means clustering technique act as a prefiltering step for grouping similar web services based on the similar quality of service and filter out unrelated web services. From the set of filtered web services, Skyline technique will obtain a set of non-dominant web services and select the skyline services as the best candidate services. This Approach will reduce the searching space and increase the performance of the service selection. More precisely, we propose a model for the selection of Web services based on nonfunctional criteria of QoS. In order to show the feasibility performance study of the proposed architecture, the result of the test shows improvement in the web service selection.

***Keyword*--**Web Service Selection, Quality of Service (QoS), K-Means Clustering, Skyline

I. Introduction

Web Services Modular, comprehensive and reusable software Technology designed to support machine to machine interoperable over Distributed environments, based on open XML-based standards and communication protocols like HTTP, Simple Object Access Protocol (SOAP), Web Services Description Language (WSDL) and Universal Description, Discovery and Integration (UDDI). The selection of web service plays an essential role is SOA.Web services are considered as self-contained, self-describing, modular applications that can be published in UDDI, located, and then invoked by SOAP leads an important approach for the integration of distributed Heterogeneous and autonomous applications across the Web world. The figure-1, illustrates The Architecture of Web services includes service requestor, Registry and service provider. Commonly, a service provider publishes service descriptions and its specifications in the registry (UDDI).Where service requestor discovers the Web Service by the client for desirable service over the web.

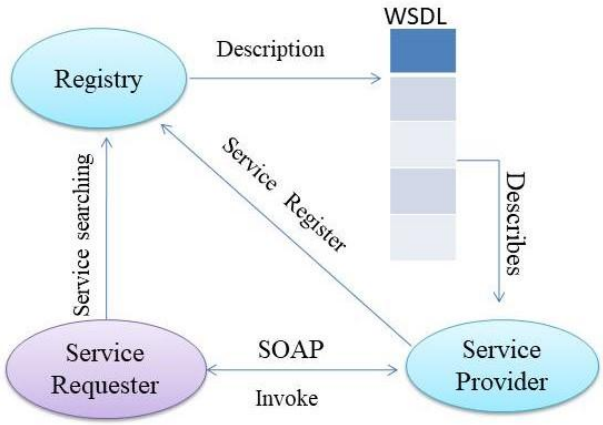

Figure 1. Web service Architecture

Web Services is the emerging field and even more in the pervasive environment, Currently due to a large number of services available in the Web, The most important challenges in Service-oriented Architecture is to selecting the appropriate services dynamically which meets the requirements of clients effectively. While selecting the effective web services the quality of service of web service must be considered also (non-functional).The description of web service contains the functional and non-functional

parameters attributes that add a specific functionality for web service candidates. Clustering technique act as pre filtrate mechanism which can be used based on QoS information to group similar web services, K-Means clustering is an unsupervised learning technique which considered only those web services which meet closely the end user specification, to cluster the WSs based on QoS information [2].More precisely the cluster, services are processes for the selection only those which is closest to the end user specification. Selection problem based on criteria can be solved by using MCDM (Multi-Criteria Decision Making) method [3] but skyline approach has been introduced to solve the problem by selecting skyline services as the best candidate service for the user.

Present Work is structured as follows: Sections I, contains the introduction which describes the Web Service model and process related to web services selection. In Section II, the properties of Quality-of-Service (QoS) parameters are defined and their importance in the selection process. In section III provides a related work of the existing system. In Section IV, we define Web Service selection technique based on Functional and Non-Functional needs. Section V, explains the selection approaches Clustering, Skyline, Skyline Model and procedure. In Section VI, the Proposed Architecture and its Description. In Section VII, the result shows the comparison of service selection. In Section VIII, we conclude research work with future directions.

II. QoS PROPERTIES

Quality-of-Service (QoS), which is usually employed for describing the non-functional characteristics of Web services, has become an important critical point for differentiating different Web Services It is calculated based on the non - functional properties of the service. QoS of different Web Services can be divided into user-independent QoS properties like (Price, popularity, Accuracy) and user-dependent QoS properties like (Availability, response time, throughput). Quality of service has the ability to satisfy the importance of it by.

- Describe the operational metrics of web service.
- Differentiate the service providers and service.
- Web service ranking and filtering process.
- Select the most effective and suitable service that carry out the perfect client activity.

The most representative QoS properties of Web services are as follows.

- **Throughput:** The maximum number of web service requests successfully executed in a given particular unit time.
- **Accessibility:** First the Web service should be available and should have the capability of serving the Web Service request without worrying about the high volume of requests.
- **Execution Time:** The execution time measures the time duration between the moment when a user sends the request and the moment when the user received the desirable result.
- **Availability:** Availability of the web service is the probability of the service can be accessed anytime.
- **Response Time:** The time duration of a process between the web services receive a SOAP request and produce a SOAP response.ie time duration user send a request and receiving a response.
- **Cost:** The Economic condition of using the web service. The fee that a service user has to pay for invoking a Web service.
- **Reputation:** Reputation depends on the end users experience of the service by its trustworthiness, the value of the reputation is defined as the average ranking given to the service by the user and the average of the reputation of the service is shared by all the clients' feedback.
- **Accuracy:** Accuracy is defined as the number of errors compared to the total number of executions. It shows the Error rate produced by the service.
- **Security:** Comprises other non-functional properties (i.e. data encryption, authentication, authorization) in web services. It should enforce the security in the transaction who assures the ACID properties (Atomicity, Consistency, Isolation and Durability).[21]

III. RELATED WORK

In the work proposed, the author propose a methodology for web service selection that combines three techniques of MCDM methods, i.e. skyline, BWM (Best Worst Method) and Vikor .in which Skyline method is for reduce the research space and compute the weights of the QoS parameters and followed by the ranking of web services with vikor, Further provides the results which comparing the results with other techniques, so to find the performance of given procedure [3].

In the work proposed in Euclidean distance based method and the SASKS algorithm, then a spatial clustering technique called the Spherical Associated Keyword Space. Here QoS aware service clustering methodology is been followed to increase the performance of the service selection by functions a service cluster is been created as

the first factor, then QoS properties are being considered as a secondary factor [5].

In the work proposed in QoSBF (Quality of Service Bootstrapping Framework) is added so that it evaluates the QoS of the newly registered services at publishing time and plays an important role in the evolution of distributed paradigm used as the need for QoS qualification integration. Qualities like availability, Response time, Throughput, Reliability, Latency, Execution time and Accessibility are addressed in the paper [6].

IV. WEB SERVICE SELECTION

A web service selection must be made in order to determine which relevant Web services would satisfy a user's requirements and are becoming challenging and time-consuming tasks due to a large number of Web services available on the internet. A Web service selection technique consists of a set of instructions for selecting a web service at runtime, based on the information available to the system.

1. Selection Based On Functional Needs

In the case of selection based on functional need, the syntactic search is weaker because it is based on keywords that will not provide an effective result. Functional qualities contain information about the type of service, operation name, format and semantics of the input/output data.

2. Selection Based On Non-Functional Needs

In the case of selection based on non-functional needs, the search makesthe bestresponse for the user as the selection is based on several criteria and not on a single criterion. And these nonfunctional properties are called Quality of Service (QoS) and are generally expressed using Quality of Service criteria such as availability, response time, reliability etc. [7].

V. WEB SERVICE SELECTION APPROACH

Web Service selection is still a difficult task in the web-based system. The overall performance of the Web Service is determined by the selection mechanism. Web Service is associated with description and function offered by the non-functional requirement of the Web Services. Web Service is defined by three ways, WS= {I, O, QoS}, where I stand for input(s), need by web services to process. O is the output(s) produce by the web service and QoS are the numerous quality of service provided by the Web Service. The problem of Web Service Selection is to select the WSc, if client QoS is the end user QoS Specification of the desirable Web service, than WSc{QoS}=User QoS. The web service selection task of satisfying the QoS requirements for the user is difficult and involves comparison and computation in large number when the new Web service is added task becomes more complex and time-consuming. So we must consider only those web services which have the capacity of satisfying the users' requirements and reject the other web services. However, if we cluster the services based on QoS values, then we can reduce the search space and computations. First, we can identify the cluster with better QoS value and then can limit the searching process within the clu$_2$ster.

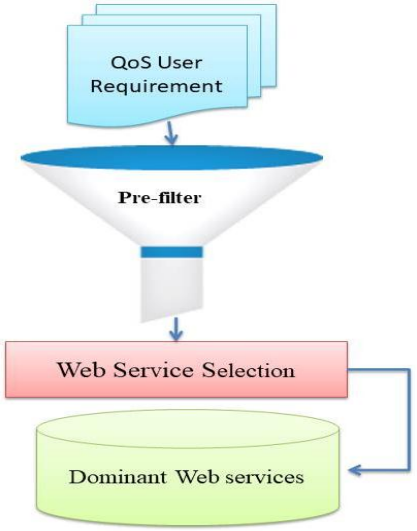

Figure 2.Web Service Selection Model

The Fig 2 illustrates, the cluster layer is added before the process of Web service selection is carried out. The process of a cluster layer is to identify those web services having similar QoS needs. So for generating a cluster of Web Services, we have employed K-Means clustering.

1. Clustering

K-Means is a clustering technique based on centroid. In which a new entity is added to the cluster, from the centroid the distance of an entity is found to be minimum. The centroid of each cluster is a point in n-dimensional vector space, where 'n' is the number of properties with the entity[19].Web Service contains QoS parameters are used to make the centroid,The cluster in web services are formed in a way that the intra-cluster QoS similarity of Web Services is high and inter-cluster similarity will be low. So for that, we can use the k-Means clustering Technique to generate clusters in web services. In this technique, it takes J as inputs and generates J cluster.

When the process starts it consider individual WSs as cluster and cluster centre as its QoS information. During each step of the process, merging takes place between two smallest distances of clusters. The cluster centre is re-evaluated as a mean of QoS of WSs in that cluster and process continue till function obtains optimal values. The criterion function is the mean square function represented in equation [19].

$$E= \sum_{i-1}^{k} \sum_{wseci} |ws - Mi|^2$$

Where, E represents the mean-square-error of all the web services, as is the web-service in the set and M_i is representing the mean of cluster C_i The process of the k-Mean is to generate optimalmal number of clusters in all candidate websites. Web services in this cluster are identified as WS = {WS_1,WS_2,.WS_n} and a set of QoS parameters Q = {Q_1,Q_2,……Q_n}, and the resulting web services are the four tuples, WS_{clust} = {I, O, QoS, C_n}, where, I stands for input(s) need by web services to process. O is the output(s) produce by the web service and QoS are the numerous web services provided by the Web Service and Cn is the cluster number. This cluster procedure will help enough for the user QoS requirements.

2. Skyline

The skyline is based on the relation of dominance defined as follows, Given a set of functionally similar Web services, A web service WS_i is said to be dominating web service WS_j, if all QoS of WS_i is better or equal than that of WSj at all dimensions and strictly one QoS parameter is better at least one dimension. The Skyline Procedure returns all web services wsi, such that wsi is not dominated by any other web services wsj. The skyline concept was initially introduced for the database domain with (i) Block Nested Loops (BNL) algorithm (ii) Divide and Conquer (D&C) algorithm.[2] The BNL paradigm is used to implement the skyline concept.

Consider the example, illustrates in figure-3 where a user wants to find a hotel near to town and cheap in price among a set of hotels and each hotel is described by two parameters, namely cost and distance ,in 2D space hotels are projected and the coordinates of each point are corresponding to the values of the two parameters. It is seen that point P is not dominated by any other point, i.e. there is no other hotel has lower price and shorter distance than hotel P, so point P is in skyline and Q,R and S are also in skyline, point S dominated point T and U,so T and U are not in skyline. The result skyline could be understood as an anti-correlated set as it provides a set of trade-offs between the parameters. The points in this line region can only be dominated by point Q and is the only dominance region of point S

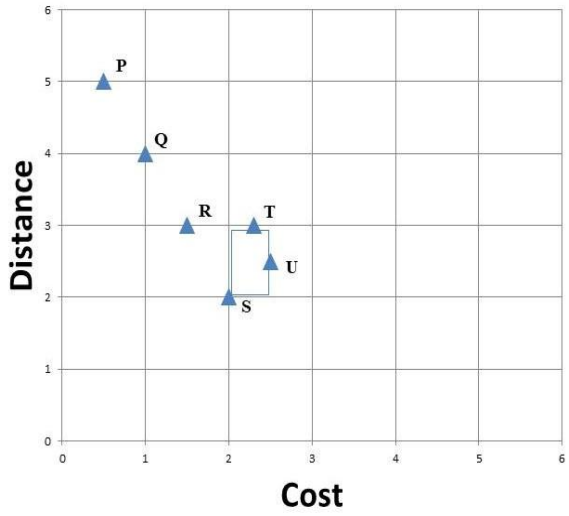

Figure 3. Skyline

3. Skyline Proposed Model

Given a set of functional n-similar Web Services WS= {*ws1, ws2…….ws*$_n$}, each Web Service contain n-set of QoS parameters (non-functional) attributes, Q= {q1, q2…….q_n}.The model of skyline should be projectin to dimension date space (figure-4). The dimension associated to an attribute of service, and the service coordinate in the data space dimension equalto the values of associated attributes. So we can change the QoS based service selection into the skyline pattern as given in the figure. We consider square dots as existing service and round dots as adding new service in environment .so original Web service is*{ws1,ws2,ws3,ws4,ws5,ws6,ws 9}*.Now new service *{ws7,ws8,ws9}*is adding to the existing environment. When the new web service ws9 is adding, it does not change the skyline service as ws9 is a dominated service, but in case of adding web service ws8,it will change the skyline service as no service can dominate it and skyline can get the changed to form the new skyline i.e.{*ws1,ws8,ws2,ws3,ws4,ws5,ws6,ws7*}.the further adding ws7,will also effect the skyline service as *s7* dominates the three existing services {*ws2,ws3,ws4*}therefore we get the final skyline will be {*ws1,ws8,ws7,ws5,ws6*}on the basic of two attributes. We find *ws7*has its response time equals to *s2* and its cost equal *s4*.So *ws7*fix best.

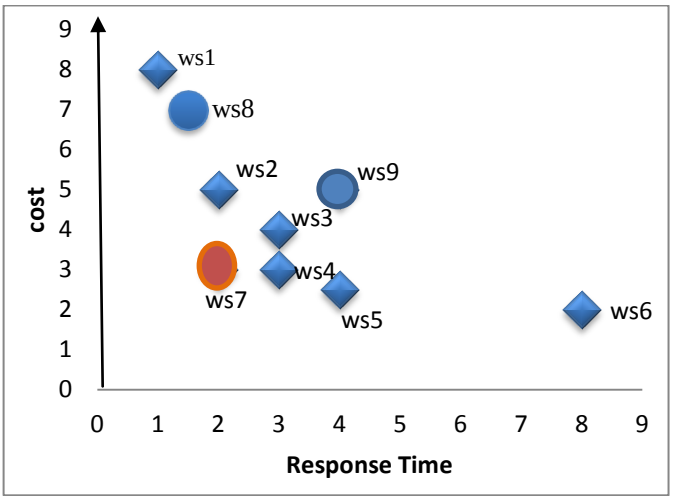

Figure 4.Skyline services

4. Skyline Procedure

i. First, compute the whole services to get the first skyline.
ii. Get value order of services in R-tree structure dimension.
iii. If value of S_r (i) + s_c (i) < S_r (j) + s_c (j) than WS (i) is dominated service.
iv. If there is any change, computing whole services again to get the skyline.

In the pre-processing stage, we rank the services in the way of fast services response time and low cost.

VI. Proposed Architecture

In this section, here we will present the overall architecture (figure-5). In which we add a new component between the client and registry in such a way so that searching and selection process will be effective on the basic of the evaluation of the quality of service QoS criteria.

1. Description

The Architecture states the life cycle of a client who requests a service from Registry (UDDI) through discovery mechanism but the service provider should register the service with QoS information in the Registry before the client can invoke and utilize the service. When the client request for the service, the sending request contains its desired description, which must include the functional criterion and other non-functional criteria of the service.

When the Registry receives the request The K-Means clustering technique grouping the web services with similar Quality of Service (QoS) under a common umbrella(cluster) is explored so as to filter out unrelated web services which reduce the future searchingprocess.

A similar list of web services based on QoS is sent to discovery module which sends it for rankings procedure that operates on the reduced problem set and identifies the best web service among the cluster group dynamically skyline.

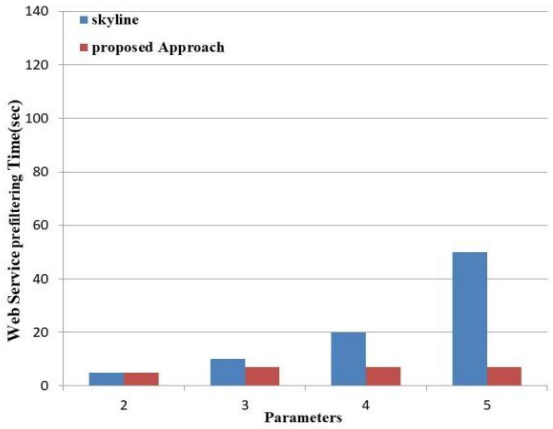

Figure 5. Web Service Selection Architecture

VII. RESULTS

To test the approach, parameters namely Reliability, Throughput, Availability, Cost, the Response time from QoS parameters dataset [23].The data set is generated by using web service broker frame work. The dataset includes a set of web services with QoS measurement of each parameter. K-Mean based clustering for pre-filtering the web services, followed by the skyline technique for web service selection approach and compared with skyline based PROMETHEE [22].Clustering based pre-filtering approach has the edge in Time efficient mechanism as compared to the Promethee algorithm(figure 6).

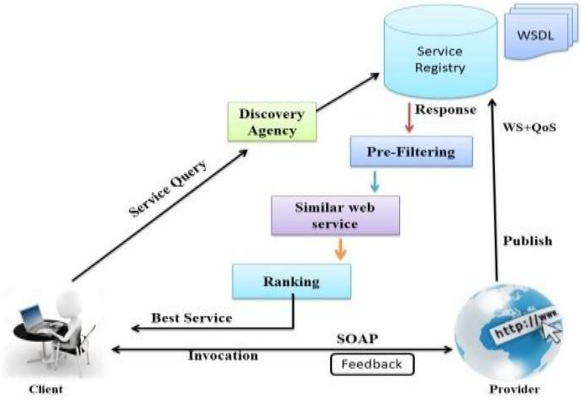

Figure 6. Prefiltering Time Comparison

K-Means based clustering uses Prefiltering of Web Service in the proposed approach, in previous works Skyline technique for prefiltering has less impact over The clustering based prefiltering approach for web Service selection.The proposed k-Mean based clustering prefiltering mechanism has more time efficient significance as compared to the skyline based prefiltering mechanism[22]. In clustering based prefiltering technique QoS parameters from 2 to 5 take almost similar time for each other parameters. The results from Prefiltering layer are given to selection algorithm. With the changing number of QoS parameters from 2 to 5 the selection time comparison of the mechanism [22] and the proposed approach is done (shown in figure 7).skyline algorithm will take less time as compare to Promethee algorithm when the limited number of inputs are provided to the skyline. Furthermore, it's observed that it can satisfy the end user specification and capability to search web service better than the existing method.

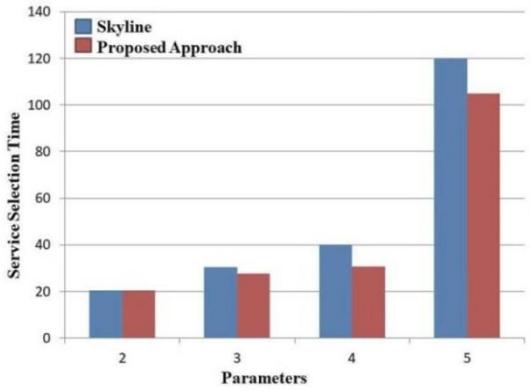

Figure 7. Web service selection time comparison

VIII. CONCLUSIONS

In Web Service based systems service selection is an important task, QoS based mechanism is an effective technique to govern the efficient web service selection, As the number of functionally equivalent Web Services are increasing, existing solutions are found to be insufficient, using K-Means clustering technique task of prefiltering is conducted.The task of the prefiltering layer is to identify Web Services with similar QoS need and reduce the overall time of Web Service selection than skyline determines the dominated set of web services. In future, we plan to conduct a thorough evaluation of QoS aware Web Service clustering approach dynamically over existing methods.

REFERENCES

[1]. Ronak R Kayastha, JwalantBaria, "A survey on Web Service Selection and Ranking Methods", International Journal of Computer Applications 117(16):5-8, May 2015.

[2]. LalitPurohit, Sandeep Kumar, "Exploring K-Means Clustering and skyline for Web Service Selection", 2016 11th international conference on industrial and information systems(ICIIS).

[3]. WalidSerrai ,AbdelkrimAbdelli , Lynda Mokdad , YoucefHammal, "An Efficient Approach for Web Service Selection", ISCC 2016, PEDISWESA.

[4]. NarimanAmmar, Zaki Malik, BrahimMedjahed, Mohammed Alodib, "K-Anonymity Based Approach For Privacy-Preserving Web Service Selection", 2015 IEEE International Conference on Web Services.

[5]. Banage T. G. S. Kumara, Incheon Paik, T. H. A. S. Siriweera,Koswatte R. C. Koswatte, "QoS Aware Service Clustering to Bootstrap the Web Service Selection", 2017 IEEE 14th International Conference on Services Computing.

[6]. Zainab Aljazzaf, "Bootstrapping quality of Web Services", in Journal of King Saud University Computer and Information Sciences (2015) pages 27, 323–333, 2015

[7]. HoucineBelouaar ,OkbaKazar, KhaledRezeg, "Web service selection based on TOPSIS algorithm", 2017 International Conference on Mathematics and information Technology, Adrar, Algeria – December 4 - 5, 2017.

[8]. Ronak R Kayastha, JwalantBaria, "A survey on Web Service Selection and Ranking Methods", International Journal of Computer Applications 117(16):5-8, May 2015.

[9]. Y. Xia, P. Chen, L. Bao, M. Wang, J. Yang, "A qos-aware web service selection algorithm based on clustering in Web Services", (ICWS), 2011 IEEE International Conference on, July 2011, pp. 428– 435.

[10]. K. Zheng, H. Xiong, Y. Cui, J. Chen, L. Han, "User clustering based web service discovery", 2012 Sixth International Conference on Internet Computing for Science and Engineering, April 2012, pp. 276–279.

[11]. A. Ouadah, K. Benouaret, A. Hadjali, F. Nader, "Combining skyline and multi-criteria decision methods to enhance web services selection," inProgrammingandSystems(ISPS),201512thInternational Symposium on, April 2015, pp. 1–8.

[12]. Jian Wu, Liang Chen, Tingting Liang," Selecting Dynamic Skyline Services for QoS-based Service Composition", College of computer science and technology, Zhejiang University, Hangzhou, China.

[13]. Yue Wang, You Song, Mingyang Liang," A Skyline-based Efficient Web Service Selection Method Supporting Frequent Requests "proceedings of the 2016 IEEE 20th international conference on computer supported cooperative work in design.

[14]. Mohammad Alrifai, DimitriosSkoutas , Thomas Risse," Selecting Skyline Services for QoS-based Web Service Composition" April 26-30 ,Raleigh ,Nc, USA.

[15]. A. Ouadah, K. Benouaret, A. Hadjali, F. Nader, "Combining skyline and multi-criteria decision methods to enhance web services selection,"inProgrammingandSystems(ISPS),201512thInternational Symposium on, April 2015, pp. 1–8.

[16]. K. Zheng, H. Xiong, Y. Cui, J. Chen, L. Han, "User clusteringbased web service discovery," in 2012 Sixth International Conference on Internet Computing for Science and Engineering, April 2012, pp. 276–279.
[17]. DongmeiLiu ,Zhiqing Shao, Caizhu Yu, Guisheng Fan "A Heuristic QoS-Aware Service Selection Approach to Web Service Composition", 2009 Eigth IEEE/ACIS International Conference on Computer and Information Science.
[18]. Narges Hesami Rostami ,Esmaeil Kheirkhah ,Mehrdad Jalali", An Optimized Semantic Web Service Composition Method Based On Clustering And Ant Colony Algorithm", Department of Computer Engineering, Islamic Azad University Branch Mashhad, Iran.
[19]. J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques", 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc, 2011
[20]. Ab Rashid Dar, Dr. D Ravindran, "A Comprehensive Study On Cloud Computing Paradigm", International Journal of Advance Research in Science and Engineering, Vol. 7, No.4, 2018, pp. 235-243
[21]. Shahidul Islam, Dr.S.Britto Ramesh Kumar, Ab Rashid Dar,"A COMPREHENSIVE STUDY ON WEB SERVICES BASICS", International Journal of Advance Research in Science and Engineering, Vol. 7, No.4,2018,pp.243-248
[22]. A. Ouadah, K. Benouaret, A. Hadjali, F. Nader, "Combining skyline and multi-criteria decision methods to enhance web services selection", Programming and systems(ISPS),2015,12thInternational Symposium on, April 2015, pp. 1–8.
[23]. Q. H. M. E. Al-Masri, "Qos-based discovery and ranking of web services", Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on, Aug 2007, pp. 529–534.
[24]. Deepti Sharma, Vijay B. Aggarwal, "Dynamic Load Balancing Algorithms for Heterogeneous Web Server Clusters" International Journal of Scientific Research in Computer Science and Engineering Vol.5, Issue.4, pp.56-59, August (2017).
[25]. Tusha Agarwal, Neeta Sharma, "Efficient Load Balancing Using Restful Web Services in Cloud Computing: A Review", International Journal of Scientific Research in Computer Science and Engineering Vol.6, Issue.3, pp.67-70 , June (2018).

Authors Profile

Shahidul Islam received his Bachelor's and Master's degrees in Computer Science from University of Kashmir and Baba Ghulam Shah Badshah Shah University respectively. He is presently pursuing Master of Philosophy in Computer Science at St. Joseph's College (Autonomous) an affiliated college of Bharathidasan University Tamil Nadu. His current research interests include, Web Services, Big Data Analytics and Internet of Things.

Dr. S. Britto Ramesh Kumar is working as Assistant Professor in the Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli, Tamil Nadu, India. He has published many research articles in the National/International conferences and Journals. His research interests are Cloud Computing, Data Mining, Web Services, and Security.

Ab Rashid Dar received his Bachelor's, Master's and Master's in Philosophy degrees in Computer Science from University of Kashmir, Baba Ghulam Shah Badshah Shah University, and Bharathidasan University respectively. He is presently pursuing Ph.D. in Computer Science at St. Joseph's College (Autonomous) an affiliated college of Bharathidasan University Tamil Nadu. His current research interests include, Scalability, Load Balancing in Cloud Computing, Fog & Edge Computing.

