International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 10 (2015) pp. 27425-27431

http://www.ripublication.com

Optimized Service Selection In Web Service Discovery

Mrs. M. Akila Rani

Assistant Professor,
Department of Computer Science and Engineering,
NPR College of Engineering & Technology. Natham-624401, Tamil Nadu, India.
Phone No: 04544- 246 500
Email: akilakamalam@gmail.com
.
Dr. D. Shanthi

Professor and Head,
Department of Computer Science and Engineering,
PSNA College of Engineering & Technology,
Dindigul-624622, Tamil Nadu, India.
Email-id: dshan71@gmail.com,

Abstract

Web services give standardized way in an Internet protocol backbone for integrating Web-based applications using the XML, SOAP, WSDL and UDDI. Now a days many users on company side uses cloud servers to store and retrieve their data. They can be of any one the following type as private, public or hybrid. Our concept mainly focuses on how to efficiently retrieve the web services. For this purpose we introduce concept as DSGA (Discovering webService using Genetic Algorithm) based upon KDS. For service discovering process our DSGA follows the four different procedures such as the discovery, equivalence processing, clustering, categorization, filtering, and presentation. In discovery stage the services are all stored and that are extracted based upon the service name, operation name etc. For service selection we use the genetic algorithm. Our main goal of this paper is providing optimal solution in order to meet the complex needs of users with one set of single web services with high degree of matching. Still there are problems such as reduction of accuracy, increase of response time in several methods for web service composition. We proposed an algorithm according to the response time, accuracy with the help of genetic technic and K mean ++ clustering algorithm.

Keywords: XML, SOAP (Service oriented Architecture protocol), WSDL, UDDI, Web services.

1. INTRODUCTION

Web service is a promising technology that allows constructing and sharing independent and autonomous software. Service interface publication, service discovery and service invocation are performed using XML based standards, known as WSDL, UDDI and SOAP .The Web service model consists of three entities, the service provider, the service registry and the service consumer. Web services can be described, published, located and invoked and they are self-contained component applications over Internet. Descriptions of web services that is WSDL enable web services to be discovered and used by other web services [1]. These descriptions are described using a standard XML-based language WSDL (Web service Description Language) file may be of two types1)with *The functional features* are needed to invoke the execution of a web service 2) *nonfunctional features* are such as cost, response time, reliability etc [2, 3, 4]. In the domain of web service, composition is an important technology for reusing existing web services when more than one services are available to meet the users requirement. In web service composition method two types are available 1) syntactic composition 2) semantic composition. The syntactic and semantic composition is based on syntactic descriptions and semantic descriptions. In semantic web service composition semantic description are added instead of the parameter values from concepts of ontology. In this paper, we propose a novel technique to find an optimal web service which exactly matches the users need. The focus of this work is on the identification of exact web service required by the users and also giving importance to the special symbols, comma (,), and space typed by the user when giving request.

The rest of this paper is organized as follows: the related works of web service composition approach is presented in section 2. In section 3 is discussed the description of the system. Finally in Section 4 conclusion is stated.

2.RELATED WORK

David Arthur et.all., [5] proposed k-means method is a widely used clustering technique that seeks to minimize the average squared distance between points in the same cluster. Although it offers no accuracy guarantees, its simplicity and speed are very appealing in practice. By augmenting k-means with a simple, randomized seeding technique, we obtain an algorithm that is O (log k)-competitive with the optimal clustering. Experiments show our augmentation improves both the speed and the accuracy of k-means, often quite dramatically.

Chunlai Chai et.all., [6] proposed the use of fuzzy techniques has been considered to be one of the key components of data mining systems because of the affinity with human knowledge representation. A hybridization of fuzzy sets with genetic algorithms is described for Web mining in their paper.Genetic algorithms provide efficient search algorithms to select a model, from mixed media data, based on some preference criterion/objective function.

Yufei Tao et.all., [7] gave the survey on three principal categories, namely partitioning clustering approaches [10], [11], [12], density-based clustering approaches [13], [14], and possible world approaches . The first two are along the line of the categorization of clustering methods for certain data , the possible world approaches are specific for uncertain data following the popular possible world semantics for uncertain data .

Indira K et.all., [8] used Genetic algorithm. It can dispose large-scale data gathered in a lot. It is widely applied in mining association rules. Genetic algorithms are typically implemented using computer simulations where optimization is the main criteria for solving the problem.

Shang-Chia Liu et, all [9] apply genetic algorithm to quickly find the best-fitting service composition. They score and sort each service composition based on the service requesters' preferences towards QoS.

3.DESCRIPTION OF THE SYSTEM

3.1 SYSTEM ARCHITECTURE OF THE PROPOSED METHOD

Our concept focuses on how to efficiently retrieve the web services. For this purpose we introduce concept as DSGA (Discovering webService using Genetic Algorithm) based upon KDS. For service discovering process our DSGA follows the five different steps such as the equivalence processing for service selection, clustering, categorization, filtering, and presentation. In discovery stage the services are all stored and that are extracted based upon the service name, operation name etc., present in the WSDL (Web service Description Language) file. For service selection we use the genetic algorithm. Fig 1.shows system architecture.

The genetic algorithm consists of four different steps such as initial population, fitness function, mutation and cross over to discover the services. The initial population takes the input value as the integer but our given output is string which means that our given input is any sentence or words etc. And also our main aim is to discover the potentially value added service for this purpose first of all we are going to identify the similarity between the given input. For calculating the similarity, weight and also distance we apply the similarity functions. Our obtain output is given as input for the genetic algorithm.

During this similarity identification we perform clustering operations also so our DSGA offer the value added services. After that based upon the initial population the fitness function is calculated. Next step is cross over, for this purpose we use the two point cross over operator. From the result move to the mutation, here we use simple swap operator to find out the result. These processes are repeated until the end of iterations.

Next stage of our Web mining is clustering. For our proposed work we apply the k-means++ clustering algorithm for grouping the discovered services based upon their characteristics which offer efficient result when compared to the others. Most of the user uses the k-means clustering algorithm for the clustering process. It offers the interesting result but one of the drawback of that approach is it makes the clustering based upon the centre point. But it chooses the centre point in random manner. So it does not provide efficient result to overcome this problem we use the K-means++ algorithm. Next step is that clustered results are categorized.

In categorization stage we classify the services. After categorization we move to the filtering process here the unwanted things are filtered after that the final result is displayed such as the web services are discovered. If the discovered services contains more than one services means we composite the services after that display the result. The tour related data sets are considered for our web mining process. Based upon the user requirements the services are discovered. We maintain these details in cloud.

3.2 SIMILARITY AND DISTANCE FUNCTION CALCULATION PROCEDURE:

For similarity we use the wordnet tool. Our distance function is defined based on a given convex function. Our distance function is defined as,

$$d_B(x_a, x_b) = (\varphi(x_a) - \varphi(x_b))^T (x_a - x_b)$$

distance function is performs the distance calculation by using the kernel of reproducing kernel Hilbert space, which is denoted by $H_{\kappa,}$ author aimed to search for convex function $\varphi(x) \in H_\kappa$ such that the induced our distance function, denoted as $d_\varphi(x_a, x_b)$, minimizes the overall training error. It defines the convex function of $\varphi(x)$: $R^D \rightarrow R$. Here $\varphi(x)$ is the strict convex function.

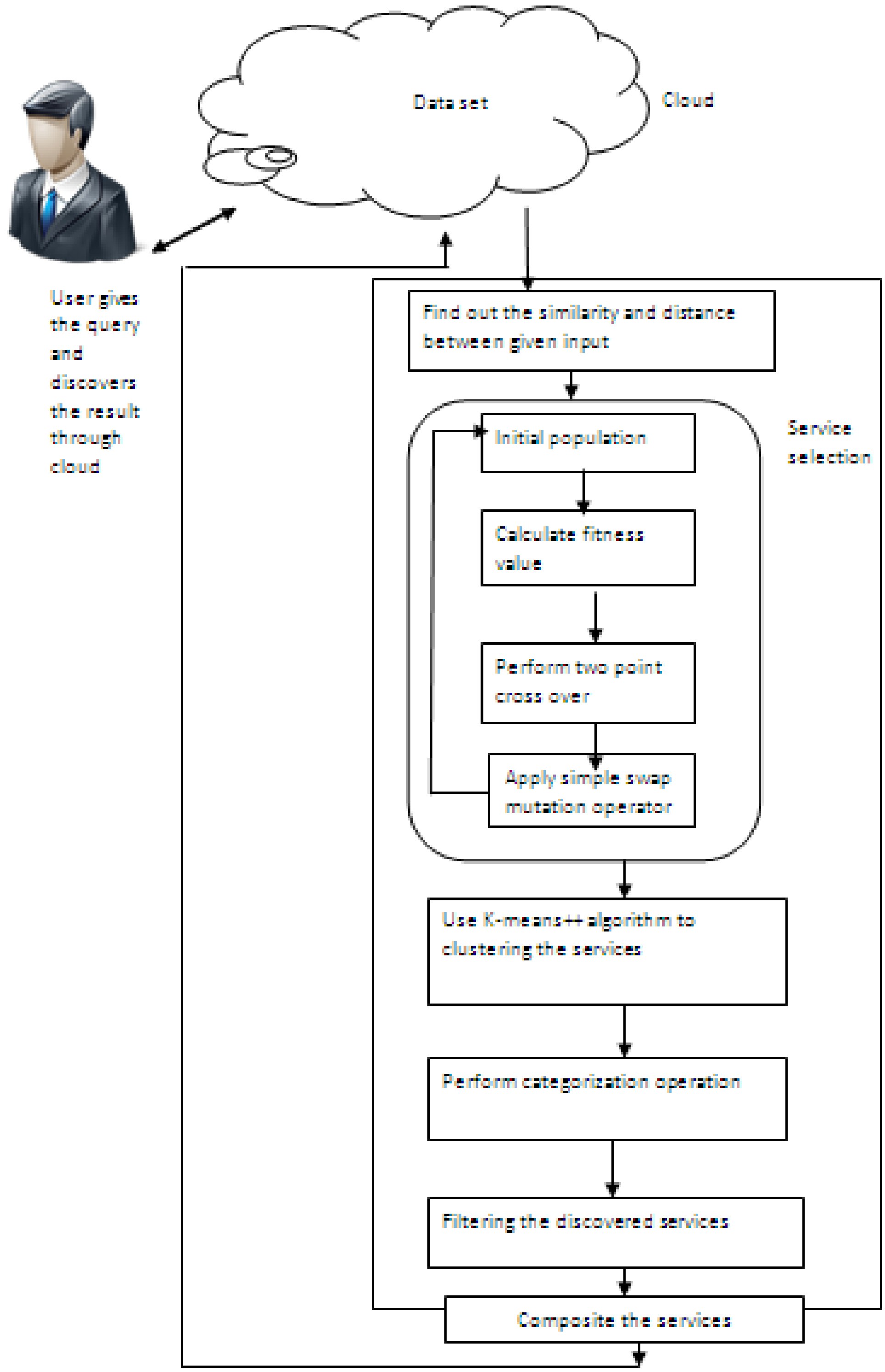

Fig.1.System Architecture

To perform the Bregman distance function, first process is to obtain the expressions for φ (*x*),

$$\varphi(x) = \sum_{i=1}^{N} \alpha_i \, h(x_i^T x)$$

and, as a result, $d_B(x_a, x_b)$ can be expressed as follows:

$$d_B(x_a, x_b) = \alpha_i \, h'(x_a^T x_i) - h'(x_b^T x_i) \, x_i^T (x_a - x_b).$$

4.CONCLUSION

In this paper we have given great importance to special symbols, space given by the user at the time of their request. By using similarity and distance measure which given as input to service selection step, we can identify the similar services at the time of initial process itself. All the similar services only get into next steps. Thus here we have done preprocessing which is considerably reduce the processing time. All the semantics (meaning) of the string given by the user which including the special symbols, comma, space are considered here. Service selection using genetic algorithm will produce optimized output. As a result, we can get exact result queried by user rather than other methods with low processing time.

5. REFERENCES

[1] Narges Hesami Rostami and Esmaeil Kheirkhah and Mehrdad Jalali, ”An Optimized Semantic Web Service Composition Method Based on Clustering and Ant Colony Algorithm”, International Journal of Web & Semantic Technology (IJWesT) Vol.5, No.1, January 2014.

[2] Mansour H. Elfawal and T. Dillon, "Dependability and Rollback Recovery for Composite Web Services," IEEE TRANSACTIONS ON SERVICES COMPUTING, vol. 4, no. 4, pp. 328 - 339, OCTOBER-DECEMBER 2011.

[3] Pablo Rodrı´guez-Mier, Manuel Mucientes, Manuel Lama, and Miguel I. Couto, "Composition of web services through genetic programming, " vol. 3, no. 4, pp. 171-186, December 2010.

[4] Rajesh Karunamurthy, Ferhat Khendek, and Roch H. Glitho, "A novel architecture for Web service composition, " Journal of Network and Computer Applications, vol. 35, no. 2, pp. 787–802, March 2012.

[5] Zhou Huang, Yu Fang, Xuetong Xie, Mao Pan, “GEOBARN: A Practical grid geospatial database system”, Advances in Electrical and computer engineering, volume 9, Number 3, 2009.

[6] David Arthur and Sergei Vassilvitskii., 2006 “How slow is the k-means method? In SCG ’06: Proceedings of the twenty-second annual symposium on computational geometry. ACM Press.”

[7] Chunlai Chai, Biwei Li, 2010” A Novel Association Rules Method Based on Genetic Algorithm and Fuzzy Set Strategy for Web Mining”, JOURNAL OF COMPUTERS, VOL. 5, NO. 9, SEPTEMBER 2010, page no 1448- 1455

[8] Bin Jiang, Jian Pei, Yufei Tao, Xuemin Lin, 2013”Clustering Uncertain Data Based on Probability Distribution Similarity “ IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 4, APRIL 2013 pgno751-763.

[9] Indira, K., and Kanmani, S., 2012**, ”**Performance Analysis of Genetic Algorithm for Mining Association Rules” IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.ISSN (Online): 1694-0814.

[10] Shang-Chia Liu, Sung-Shun Weng, 2012, ” Applying Genetic Algorithm To Select Web Services Based On Workflow Quality Of Service “, Journal of Electronic Commerce Research, VOL 13, NO 2, 2012.pg. no-157

[11] W. K. Ngai, B. Kao, C. K. Chui, R. Cheng, M. Chau, and K. Y. Yip” Efficient clustering of uncertain data”. In ICDM, 2006.

[12] B. Kao, S. D. Lee, D. W. Cheung, W.-S. Ho, and K. F. Chan. “Clustering uncertain data using voronoi diagrams”. In ICDM, 2008.

[13] S. D. Lee, B. Kao, and R. Cheng.” Reducing uk-means to kmeans”. In ICDM Workshops, 2007.

[14] H.-P. Kriegel and M. Pfeifle. “Density-based clustering of uncertain data”. In KDD, 2005.

[15] H.-P. Kriegel and M. Pfeifle.” Hierarchical density-based clustering of uncertain data”. In ICDM, 2005.