IMMUNE-INSPIRED METHOD FOR SELECTING THE OPTIMAL SOLUTION IN SEMANTIC WEB SERVICE COMPOSITION

Merzoug Mohamed[1], Chikh Mohammed Amine[2] ,and Bekkouche Amina[1]

[1]Department of Computer Science, Abou Bekr Belkaid University, Faculty of Sciences, Tlemcen, Algeria
[2]Biomedical Engineering Laboratory, Abou Bekr Belkaid University, Faculty of Technology, Tlemcen, Algeria

ABSTRACT

The increasing interest in developing efficient and effective optimization techniques has conducted researchers to turn their attention towards biology. It has been noticed that biology offers many clues for designing novel optimization techniques, these approaches exhibit self-organizing capabilities and permit the reachability of promising solutions without the existence of a central coordinator. In this paper we handle the problem of dynamic web service composition, by using the clonal selection algorithm. In order to assess the optimality rate of a given composition, we use the QOS attributes of the services involved in the workflow as well as, the semantic similarity between these components. The experimental evaluation shows that the proposed approach has a better performance in comparison with other approaches such as the genetic algorithm.

KEYWORDS

Web Services ,Semantic Web Service Composition, Clonal Selection Algorithm, Semantic Similarity, QoS.

1. INTRODUCTION

Web services provide a promising approach for implementing Enterprise Application Integration (EAI), but their use is not restricted to this domain, in fact they might form the ideal technology for building large scale distributed applications. The elementary units of these applications can be owned by different stakeholders.(such as banking systems , healthcare systems, e-tourism…), and therefore it is not trivial to design and execute this kind of applications with using the service oriented computing principles.

In the real world, a single web service cannot fulfill a complex requirement of a given enterprise task, however a composition of different web services might meet the complex requirements.

Nevertheless, to discover and compose Web services, we need several informations that describe the service functionality. These informations are partially represented and supported by the signature of the operations and the message formats, which together form the Web service syntactical interface, captured in the WSDL document. The lack of any machine interpretable semantic requires human intervention in service composition, and consequently, we are not able to automate composition process. Semantic Web services [3] provide a solution to this problem

DOI : 10.5121/ijwest.2014.5402

by annotating services with semantic concepts, thus assuring a common understanding of the content, the functionality and the behavior of the software components.

It is worth noting that, the optimization of web service workflow is a crucial step in web service composition, in fact we have to select a set of services that satisfies several functional and non-functional requirements. The Selection of the optimal composition can be seen as a discrete optimization problem which requires specific kinds of optimization algorithms. Recent research studies demonstrated that principles inspired by the biological systems have encouraged the design of efficient optimization techniques. These biologically-inspired techniques are advantageous since they are capable of converging towards the optimal or a near-optimal solution in a short time without processing the entire search space. Such meta-heuristics include Ant Colony Optimization [6], Particle Swarm Optimization [13], Genetic Algorithm [9] or Artificial Immune Systems [5].

The objective of this paper is to solve the web service workflow optimization by using bio-inspired solutions, our main contributions are resumed as follows:

1-we present an optimization algorithm called ClonAlg [5], in order to find a near optimal composition. This later is able to explore intelligently a large space of solutions, in a short time and without processing the entire search space. We assume in this work that the user's request is described in terms of functional and non-functional properties, the functional aspect represents the service interface (the inputs, the outputs…), and the non-functional properties represent the QoS attributes such as response time, cost, availability, reputation, etc.

2- We provide an experimental comparison between the proposed approach and the genetic algorithm.

The reminder of the paper is organized as follows: Section 2 gives an overview of related work. The semantic model and QoS model for web service composition are described respectively in Section 3 and Section 4.In Section 5, we give the details of our immune-inspired technique, and experimental results are presented in Section 6. Finally, in Section 7, we present our conclusion and we give the directions for future work.

2. RELATED WORK

Several bio-inspired methods have been proposed for selecting near optimal service compositions. A genetic-based algorithm for selecting the best solution in multi-path web service composition is proposed in [12]. A genetic chromosome is mapped to a service composition solution and each chromosome unit is represented as a triple containing the task context, workflow sign (specifies if the task is part of an AND/OR workflow) and the pointer towards the candidate service [12]. Initially, a set of service compositions (solutions) having different topologies (corresponding to different paths) are randomly generated.Then the actual selection process is performed by iteratively (i) selecting the best solutions based on QoS attributes, (ii)applying a crossover operator between solutions with different topologies, (iii) and randomly mutating a randomly chosen service composition solution.

In [10] a modified Genetic Algorithm is proposed to identify the best solution in web service composition, which was based on Ant Colony Optimization (ACO) and Genetic Algorithm. The approach aims to learn advantages of both GA and ACO algorithms and overcoming their shortcomings.

A hybrid method combining Particle Swarm Optimization (PSO)[13]with Simmulated Annealing is proposed in[7] for selecting the optimal service composition solution based on QoS attributes. A composition solution is considered as the position of a particle in PSO, while velocity is used to modify a composition. To avoid the problem of premature stagnation in a local optimal solution, a Simmulated Annealing-based strategy is introduced which produces new compositions (solutions) by randomly perturbing an initial solution.

Immune-inspired approaches have gained ground in the context of selecting optimal compositions. In [8] and [18], two immune-inspired selection algorithms are proposed in the context of Web service composition. Both approaches use an abstract composition plan which is mapped to concrete services, the service composition is modeled as a graph structure.

In comparison with the aforementioned approaches, we notice that our contribution uses a fitness function having two types of criteria: the semantic quality of the desired composition and the QOS attributes, however the precedent approaches use only the QOS attributes.

3. SEMANTIC MODEL FOR WEB SERVICE COMPOSITION

In this section, we formalize the semantic web service composition problem. Initially we give some definitions about semantic web services, composite service and user request , then we describe the computation of the semantic similarity between a candidate solution (composite service) and the user's request.

3.1. Conceptual Representation

Definition 1 (Semantic Web Service) A semantic web service is described by functional and non-functional properties. Formally: $S = \{ID, IC, OC, q\}$
Where ID is the service identifier, IC and OC are the input and the output concepts respectively, q is the QoS scores list (e.g., cost, response time, etc.).

Definition 2 (Composite Service) A composite web service (CS) is represented by a set of semantic web services. Formally: $CS = \{S_1, S_2, \dots, S_n\}$
Where N is the number of web services. In this work we consider only the sequential compositions case.

Definition 3 (Request) The user requirements, are defined as follow: $R = \{IC, OC, G\}$
Where IC and OC represent respectively the inputs and the outputs concepts of R, G is the global QoS constraints list. (See section 4.4 for more details)

3.2. Computing The Degree Of Composite Service Similarity

The matching between a composite service (candidate solution) and a request is determined in different ways, depending on the semantic descriptions of the elements to match. There are three main approaches to matching : IO-matching [15] ,PE-matching [16] and IOPE-matching [11].

In our approach, we have adopted an IO-matching based on subsumption reasoning; the subsumption test allows the computation of the semantic similarity between the user request and a candidate composite service. Since the adopted benchmark, ie SAWSDL [17] , doesn't contain the preconditions or the post conditions, we use only IO_matching. For this purpose, we use the four kinds of matching Exact, Plug-in, Subsume and Fail introduced in [15]. In what follows we consider a set of hypothesis that permit the formalization of the similarity functions.

First of all, we assume the existence of 02 consecutive services A and B, the outputs of the first service A are used as inputs for the second service B. in addition to that, we define four types of scores, for matching A and B [15] :

- Exact: two parameters are called similar if the concept of the output parameter and the concept of input parameter are equivalent.
- Plug-in: two parameters are called similar if the concept of the output parameter is more specific than the concept of input parameter.
- Subsume: two parameters are called similar if the concept of the output parameter is more general than the concept of input parameter.
- Fail: means that there is no subsumption relation between concepts.

Definition 4 (degree of elementary matching) Given a composite service

$CS = \{S_1, S_2, \dots, S_n\}$ and a request R=$\{IC, OC, G\}$, a matching is called elementary, if it is applied on a pair of I/O parameters, we have 03 cases:
$EM_S(S_i.OC, S_{i+1}.IC)$: means the degree of elementary matching between the output concepts of service S_i and the input concepts of service S $_{i+1}$.
$EM_{RIN}(R.IC, S_1.IC)$: means the degree of elementary matching between the input concepts of request and the input concepts of service S_1.
$EM_{ROUT}(R.OC, S_N.OC)$: means the degree of elementary matching between the output concepts of request and the output concepts of service S_N.
The functions EM_S,EM_{RIN} and EM_{OUT} return a values in [0,1]. In Table 1,we summarize the semantic matching functions with their scores.

Table 1. Semantic matching functions

Match Type	**Exact**	**Plug-in**	**Subsume**	**Fail**
$EM_S(S_i.OC, S_{i+1}.IC)$	1	2/3	1/3	0
Logic meaning	S_i.OC and S $_{i+1}$.IC are equivalent	S_i.OC is subsumed by S $_{i+1}$.IC	S_i.OC subsume S $_{i+1}$.IC	Otherwise
$EM_{RIN}(R.IC, S_1.IC)$	1	2/3	1/3	0
Logic meaning	R.IC and S_1.IC are equivalent	R.IC is subsumed by S_1.IC	R.IC subsume S_1.IC	Otherwise
$EM_{ROUT}(R.OC, S_N.OC)$	1	2/3	1/3	0
Logic meaning	R.OC and S_N.OC are equivalent	R.OC is subsumed by S_N.OC	R.OC subsume S_N.OC	Otherwise

Definition 5 (degree of global matching) A matching is called global, if it aggregates all the elementary matching functions, formally: Given a composite service $CS = \{S_1, S_2, \dots, S_n\}$ and a request $R = \{IC, OC, G\}$, the degree of global matching or degree of composite service similarity can be computed as follows:

$$GM(CS) = ((EM_{RIN}(R.IC, S_1.IC) + \sum_{i=1}^{N-1} EM_S(S_i.OC, S_{i+1}.IC) + (EM_{ROUT}(R.OC, S_N.OC))/((N-1)+2) \quad (1)$$

Where N represents the number of component services of the sequential composition.

4. QOS MODEL

In this section, we first present the quality criteria in the context of elementary services, before turning our attention to composite services.

4.1. QoS of Web Service

The QoS of a Web Service is a value representing the non-functional property of the Web Service, such as response time, throughput, cost and so on. The set of QoS attributes can be divided into subsets: positive and negative QoS attributes. The values of positive attributes need to be maximized (e.g., availability), however the value of negative attributes need to be minimized (e.g., response time) [1].

We use the vector $q = \{q_1(S), q_2(S), \dots\dots\dots, q_l(S)\}$ to represent the QoS attributes of a service S, where $q_i(S)$ determines the value of the ith quality attribute of S.

4.2. QoS of Composite Service

Definition 6 (QoS of composite service) The QoS for a composite service can be defined by:

$$Q(CS) = \{Q_1(CS), \dots\dots\dots, Q_l(CS)\} \quad (2)$$

Where $Q_i(CS)$ represents the estimated value of the ith QoS attribute of CS.

4.3. QoS Aggregation

The QoS of a composite service (CS) is decided by the QoS values of its component as well as the composition model used (e.g. sequential, parallel,etc). In this work, we focus on the sequential composition model. So, the estimated value of the QoS attribute of CS can be aggregated from the QoS values of its component services [4]. Table 2 shows examples of aggregation functions according to some QoS attributes.

Table 2. QoS Aggregation functions

Aggregation type	**QoS attributes**	**Function**
Summation	Response Time	$Q(CS) = \sum_{i=1}^{N} S_i.q$
	Price	

	Reputation	$Q(CS) = {}^{1}/_{N} \sum_{i=1}^{N} S_i \cdot q$
Multiplication	Availability	$Q(CS) = \prod_{i=1}^{N} S_i \cdot q$
	Reliability	
Minimum	Throughput	$Q(CS) = \min_{i=1}^{N} S_i \cdot q$

4.4. Global QoS Constraints

Global QoS constraints represent the user requirements in terms of QoS. They can be expressed as a lower bound (minimum values) and/or upper bounds (maximum values) values of the overall composite service QoS.

Global QoS value constraints are proposed by users, for instance, a service cost is lesser than 100$, the service availability is greater than 99.9%.

5. The Immune-Inspired Selection Technique

In order to find, a near optimal composition we adopt the clonAlg approach[5]. We chose to adapt clonAlg [5] to the problem of web service composition because it better converges to the optimum by iteratively introducing new candidate solutions that enlarge the search space (compared to genetic algorithms which may stagnate on local optima). Our algorithm uses the functional and non-functional (QoS attributes) properties of the web services in order to find a near optimal solution. The desired composition must:

- Optimize a function that calculates the degree of semantic similarity between the user request and a candidate solution.
- Optimize the aggregated QoS of the composition solution.
- Satisfy the end to end user's global QoS constraints (e.g. the service user may have a limited budget and thus the cost is constrained).

5.1. The Clonal Selection Theory

Clonal selection is one of the most important processes of the immune system. It is triggered when a B-cell has high affinity to an invading pathogen (antigen presenting cell) and as a result, the B-Cell is stimulated to clone itself. Through cloning, identical copies of the B-cell are obtained. The number of copies is proportional to the affinity value. The clones are involved in an affinity maturation process which helps in improving their specificity to the invading pathogen by means of mutation. Mutation is inverse proportional to the affinity value, meaning that the clones having high affinity do not need to be mutated as much as the ones with low affinity. The affinity matured clones pass through new selection processes aiming at (i) keeping the clones having high affinity to the pathogen and (ii) eliminating the clones with low affinity. The selected clones are then differentiated into memory and effectors cells.

5.2. The Clonal Selection Algorithm

In our web service composition problem, clonal selection is mapped as follows: (i) a B-cell (or antibody) is represented by a service composition solution, (ii) a pathogen (or antigen) is represented by a function F that evaluates the set of composition solutions in order to find the optimal one.

In our approach, the antigen is represented by a multi-criteria function F(CS) (see Formula 3) that maximizes:

1. The degree of semantic similarity of a composition solution (CS) .
2. The aggregated QoS of a composition solution (CS)

In order to obtain a uniform objective function, we convert the negative QoS attributes (cost, response time) into positive attributes by multiplying their values by (-1).

$$F(CS) = W_s GM(CS) + Wq \sum_{i=1}^{l} W_i Q'_i(CS) \qquad (3)$$

Where Q'_i refer to Q_i (the ith QoS attribute of CS) normalized in the interval [0, 1], l is the number of QoS attributes.

W_s, W_q€ [0, 1] represent the weights established according to user preferences related to the relevance of QoS and semantic similarity during the evaluation of a composition solution .
W_i is the weight established according to user preferences related to the relevance of the QoS attributes .

In addition F must drive the evolution towards global QoS constraints satisfaction, e.g., 10< T (CS) <50 (10 and 50 are respectively the minimal and maximal value of the response time constraint).To this end, compositions that do not meet the constraints are penalized. Several penalty functions are proposed in the literature (static, dynamic, functions, etc.),we have chosen the penalty function adopted by [14](see formula 4).

$$F'(CS) = F(CS) - W_p \sum_{i=1}^{l} \left(\frac{\Delta Q'_i}{Q'^{max}_i(CS) - Q'^{min}_i(CS)}\right)^2 \qquad (4)$$

Where Q'^{max}_i, Q'^{min}_i are respectively the maximum and minimal values of the ith QoS constraint, l is the number of QoS attributes, W_p weights the penalty factor, and $\Delta Q'_i$ is defined as follows:

$$\Delta Q'_i = \begin{cases} Q'_i - Q'^{max}_i & if\ Q'_i > Q'^{max}_i \\ 0 & if\ Q'^{min}_i \leq Q'_i \leq Q'^{max}_i \\ Q'^{min}_i - Q'_i & if\ Q'_i < Q'^{min}_i \end{cases} \qquad (5)$$

The pseudo code of clonal selection Algorithm is given below:

```
1. Initialize the population (or antibodies) i.e. each antibody or composition is randomly initialized (n represents the number of web services involved in a composition solution).
2. Initialize the iteration number: t=1
3. while (t <=T_Max) do begin
    4. Compute the affinity value (formula 4) for each antibody CS ∈ Population.
    5. Compute the set of clones “Clones_set” for each antibody CS ∈ Population(according to
       the cloning rate)
    6. Hypermutate each antibody CS of “Clones_set” and compute the new affinity value of
       the result.
    7. Union = Population ∪ Clones_set.
    8. Sort Union according to the decreasing order of affinity.
    9. Build the new population by selecting the M first antibodies of Union (M= size of
       Population).
    10. Create L random antibodies, such as L=M*random insertion rate.
    11. Replace low affinity population members with the L random antibodies.
    12. t=t+1;
13. End while
14. Return the best antibody CS.
```

6. EXPERIMENTAL RESULTS

We have tested our clonal selection algorithm on SAWSDL-TC service test collection [17], a benchmark containing more than 400 semantically annotated services, SAWSDL-TC allows the evaluation of the performances of service matchmaking techniques. The services belong to the following domains: education,medical care,food and communication.

We have extended the SAWSD-TC standard collection with estimated values of QoS attributes (in our experiments we have considered two QoS attributes: cost and response time).

The clonal selection algorithm parameters are changed to get the best results (in terms of optimality):

The antibodies number M=50.
The size of the antibody (number of web services) n= 5.
The cloning rate =0.1.
The random insertion rate=0.01.
The maximum number of iterations T_max=100 .
and the weights are fixed as follow :
Ws=0.5 (semantic similarity weight), Wq=0.5 (QOS attributes weight), WT=0.5 (time response weight), WC=0.5 (cost weight).

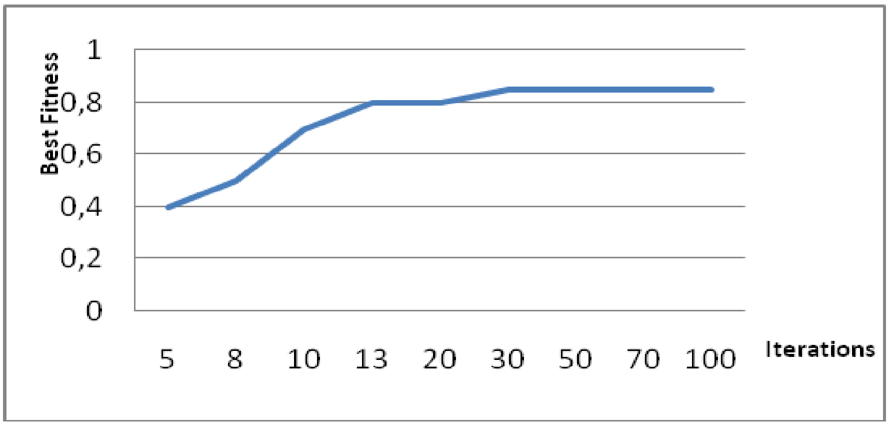

Figure. 1. The best fitness score of the clonAlg

Figure1 shows the best score of the objective function for each iteration. As depicted in this figure, the quick convergence to the optimal solution illustrates the effectiveness and efficiency of the proposed clonAlg.

Figure 2 shows a comparative study between the proposed approach, and the genetic algorithm (developed in [2]) in terms of fitness (or optimality).The results confirm the superiority of the clonal selection in comparison with the genetic algorithm (GA).

For instance, after 50 iterations, ClonAlg achieves a fitness of almost 84%, whereas GA only reaches a value of 75%.

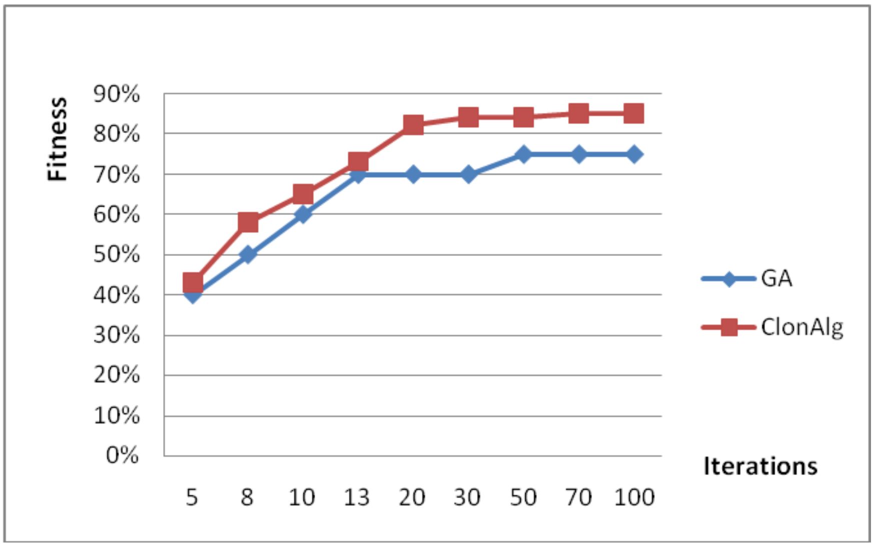

Figure. 2. Fitness vs. Iterations

In terms of execution time (see Figure 3), the GA spends more time for retrieving the best solution in comparison with ClonAlg. Concretely, ClonAlg takes only about 37 seconds to achieve the optimization with 100 generations, however GA needs more than 60 seconds to do the same task. Therefore, ClonAlg is more efficient and effective than GA.

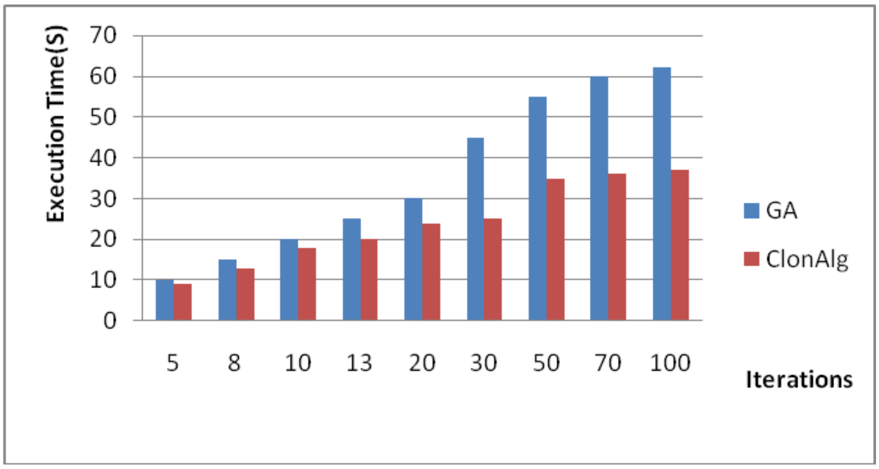

Figure. 3. Execution Time vs. Iterations

7. Conclusion

In this paper we have presented a immune-inspired technique for the selection of the best service compositions that meet the user requirements. The composition process take into account both functional and non-functional properties of web services. In comparison with GA [2], this approach increases the diversity of the population and avoids the trap of local optimality. Using this approach, the composition with better fitness value can be easily found.

As future work, we want to take into account, both the preconditions and effects of web services in the composition process. In addition to that, we intend to extend our approach to work with other composition patterns of workflow models (parallel, conditional,...) and finally we aim to compare our method with other selection approaches (such as, ant colony optimization algorithm, firefly algorithm,....) .

REFERENCES

[1] Alrifai, M. and Risse, T, (2009) " Combining global optimization with local selection for efficient QoS- aware service composition". In: WWW 2009: Proc. of the 18th International Conference on World Wide Web,pages 881- 890.

[2] Bekkouche, A., Benslimane, S.M.,Merzoug,M (2014) "A Genetic Algorithm Approach to QoS-based Semantic Web Service Composition". Book chapter In Business Intelligence and Mobile Technology Research: An Information Engineering Perspective.edt.by Sean Eom (Cambridge Scholars Publishing). http://www.c-s-p.org/.

[3] Berners-Lee,T.,Hendler,J.,Lassila,O, (2001) "The semantic web". Scientific American 284(5)34-43.

[4] Cardoso, J,(2002) " Quality of Service and Semantic Composition of Workflows". PhD thesis, Univ. of Georgia.

[5] Castro, L., von Zuben, F,(2002) " Learning and Optimization using the Clonal Selection Principle". Proc. Of the IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems, vol.6, n. 3, pp. 239-251.

[6] Dorigo, M., Birattari, M., Stützle, T ,(2006) "Ant Colony Optimization - Artificial Ants as a Computational Intelligence Technique", IEEE Computational Intelligence Magazine, 1(4), 28-39.

[7] Fan, X., Fang,X, (2010) "On Optimal Decision for QoS-Aware Composite Service Selection", Information Technology Journal, Volume 9, Issue 6, pp. 1207-1211.

[8] Gao, Y., Jun, N., Bin, Z., Lei, Y., Qiang, G. and Yu, D,(2006) " Immune Algorithm for Selecting Optimum Services in Web Service Composition". In Journal of Natural Sciences, Wuhan University Journals Press, Volume 11, Number 1, pp. 221-225.

[9] Holland,J.H, (1975) “Adaptation in Natural and Artificial Systems”, University of Michigan Press, Ann Arbor.

[10] Huan,L.,Farong,Z., Bang, O., Jiajie,W, (2010) "An Approach for QoS-Aware Web Service Composition Based on Improved Genetic Algorithm" , International Conference on Web Information Systems and Mining (WISM),vol.1,pp.123-128,23-24 Oct.

[11] Jaeger, M. C. , Rojec-Goldmann, G., Mühl, G., Liebetruth, C. , and Geihs, K,(2005) " Ranked Matching for Service Descriptions using OWL-S". pages 91–102, 2005. In Paul Müller, Reinhard Gotzhein, and Jens B. Schmitt, editors, Kommunikation in verteilten Systemen (KiVS 2005),Kaiserslautern, Germany, February, Springer.

[12] Jiang, H., Yang,X., Yin,K., Zhang,S., Cristoforo, J. A, (2011) " Multi-path QoS-aware Web Service Composition using Variable Length Chromosome Genetic Algorithm", Information Technology Journal, Volume 10, Issue 1, pp. 113-119.

[13] Kennedy, J., Eberhart, R.C, (1995) "Particle swarm optimization",The Proceedings of the IEEE International Conference on Neural Networks,Piscataway,NJ.1942-1948.

[14] Lecue,F, (2009) " Optimizing OoS-Aware Semantic Web Service Composition".In Proceedings of the 8th International Semantic Web Conference, ISWC 2009, Chantilly, VA, USA, October 25- 29.

[15] Massimo, P., Takahiro,K., Terry, R., Payne, and Katia, S, (2002) " Semantic matching of web services capabilities".

[16] Srinivasan, N., Paolucci,M., and Sycara,K, (2004) "Adding OWL-S to UDDI, implementation and throughput". In First International Workshop on Semantic Web Services and Web Process Composition (SWSWPC 2004), pages 6–9.

[17] SAWSDL-TC, http://projects.semwebcentral.org/projects/sawsdl-tc/

[18] Xu, J., Reiff-Marganiec,S, (2008) " Towards Heuristic Web Services Composition Using Immune Algorithm". In: Proceedings of the 2008 IEEE International Conference on Web Services, Beijing, China, pp. 238-245.

Merzoug Mohamed received his MS degree in computer science from the UABT University of Tlemcen in 2009.He is currently a Ph.D. candidate in the Department of Computer Science,University of Tlemcen, Algeria.His research interests include: web service discovery, service selection, ontology matching.

Chikh Mohammed Amine received his Ph.D degree in electrical engineering from the UABT University of Tlemcen in 2005.He is currently a professor in the department of electrical engineering, University Tlemcen, Algeria. His research interests include: machine learning, medical decision aiding.

Bekkouche Amina received her MS degree in computer science from the UABT University of Tlemcen in 2012.He is currently a Ph.D. candidate in the Department of Computer Science,University of Tlemcen, Algeria. Her research interests include: web service composition, web service discovery and optimization techniques.