UTM UNIVERSITI TEKNOLOGI MALAYSIA
RESEARCH UNIVERSITY

International Journal of Innovative Computing

Journal Homepage: http://se.cs.utm.my/ijic

# A Novel Decentralized Fuzzy Based Approach for Grid Resource Discovery

Saeed Javanmardi[1], Shahdad Shariatmadari[2], Mohammad Mosleh[3]

[1, 3] Department of Computer Engineering, Dezful branch, Islamic Azad University, Dezful, Iran
[2] Department of Computer Engineering, Shiraz branch, Islamic Azad University, Shiraz, Iran.
Email: [1] Saeedjavanmardi@gmail.com, [2] Shariatmadari@iaushiraz.ac.ir, [3]Mosleh@iaud.ac.ir

*Abstract*— **Grid computing is the next generation of distributed system. Its goal is to create a huge, independent and powerful virtual machine, and it has been created by gathering different machines with the aim of sharing them. Resource discovery is the most fundamental phase of resource management which has been considered in our work. Considering the advantages and disadvantages of P2P and grid and their architectural similarity; using some of P2P approaches as a solution in Grid can improve the performance of Grid. Semantic technology and fuzzy logic are the two intelligent approaches, which recently apply on P2P network and Grid environment. In P2P networks, nodes are grouping based on their interests. Moreover semantic can be very useful for grouping nodes. This paper follows the identification of resource discovery in grid environment with the help of semantic and fuzzy theory and seeking to present a resource discovery algorithm with respect to exiting obstacles. The proposed approach is assessed in simulated grid environment and the results are compared with other approaches in same conditions.**

**The goals of our approach are to obtain precious matching by clustering nodes which have services semantically related, improving search expressiveness by considering the distances between nodes (delay) in clustering phase, and decreasing response time and processing time by considering the nodes computational capabilities and free space of each node.**

**We evaluate the performance of our approach with some distributed semantic based grid resource discovery models. The results of the experiments show the efficiency of the proposed approach in term of scalability, precision, search expressiveness and response time.**

Keywords— **Grid, Resource discovery, ontology, Fuzzy theory, Semantic Overlay Network**

## I. INTRODUCTION

Grid provides an environment which all users can access the shared resources from their personal computers from anywhere without any time limitation [1]. According to the Foster and Kzlman definition, [2] grid is a hardware and software structural, which presents cheap, distributed and secure accessibility to the Powerful computing capabilities. Resource management consists of four stages Resource discovery, resource selection, and scheduling and resource allocation. Resource discovery is the most basic step of resource management. For utilization of distributed resources in the grid, efficient resource discovery approaches should be used. Resource discovery in a large scale environment such as grid can be very difficult because of the large number of resources. These resources have a varied, distributed and energetic nature [3].

To discover resources with regards to the scalability of the distributed environment, P2P approaches can be used in grid [4]. Usual P2P systems have a limitation in their search mechanism. They usually use keyword based search mechanism. To overcome this limitation semantic based resource discovery has been used in recent years [5]. Using ontology based techniques for exact resource discovery in P2P based grid environments is a proportionately novel research topic.

Service Oriented Architecture (SOA) is an architectural design whose major aim is to obtain loosely coupling among interacting entities. A service can be a part of more than one distributed system and different applications. SOA can be used in Service grid [10] and cloud computing [11] as well. The resources in SOA based architecture can be some web services which have an interface described with the usage of some XML language such as WSDL and UDDI [12].

Recently, a lot of attention has been paid to the web service annotation. We should consider that "semantic descriptions about a concept from a web service" is totally different from "semantic Web Services". Semantic Web service is a Web service that has semantic description about the service itself so that different Web services could understand each other. The semantic description here is about the Web service, not the data which the service returns. There are not any real standards for web service annotations, but there are a few projects that have addressed the problem in different ways. Amit Sheth has worked with a framework he calls METEOR-S that addresses some of the issues [13]. Sheila McIlraith has been working on a description for semantic web services based on the OWL Web Ontology Language. It's called OWL-S [14]. There is another standard WSDL-S [15] that can be used for semantically describing services in a SOA based network.

Semantic small world (SSW) is a new paradigm which has some similarity with Semantic Overlay Network (SON). It leverages three issues; small world network, semantic clustering and dimension reduction. The major goal of SSW is to cluster nodes with semantically similar resources simile to each other in a semantic space and arrange the clusters into an overlay network. The second goal is to improve search expressiveness by reducing average path length (APL) and decreasing clustering coefficient (CC) [16, 17]. High clustering coefficient and short Average Path Length are two characteristics of small world [40]. An overlay network can be called small world if it has small average path length and a vast amount of cluster coefficient. In a small world environment, The APL of two randomly chosen nodes should be approximately six hops [18]. The proposed method has a good value in term of APL.

The goals of our approach are to obtain precious matching by clustering nodes which have services semantically related, improving search expressiveness by considering the distances between nodes (delay) in clustering phase, and decreasing response time and processing time by considering the nodes computational capabilities and free space of each node. Our scheme consists of two phases. Nodes grouping and resource (service) discovery. We use fuzzy theory for each phase as a backbone. In this work we use fuzzy theory for creating SON.

The rest of this paper is as follows: in the next section we provide preliminary in which we glimpse at semantic grid and fuzzy theory as a background. Related works are presented in section III; in this section, we take a brief look at some works which are about grid resource discovery. In section IV we propose our model. The performance evaluation and experimental results are presented in Section V. Finally, in section VI, we make a conclusion.

## II. PRELIMINARY

In this part we want to take a brief look at semantic grid and fuzzy theory as the background.

### *A. Semantic grid*

Semantic Grid is a novel ambition to reveal semantically rich information associated with Grid services to generate more intelligent Grid services. In fact In the Semantic Grid, related resources and services are provided a well-defined meaning, better enabling providers and consumers to work in cooperation [19].Ontology is the major building block for Semantic Grid. It explains and clarifies the concepts, services and the relationships between them, with the aid of ontological annotation languages such as OWL [20].

### *B. Fuzzy theory*

Fuzzy logic is a logic which is much less severe than the computation computers typically carry out. Fuzzy Logic tenders various singular characterizes that make it an especially good optional for many control problems. It is essentially strong since it does not require exact inputs and can be programmed in a safe and fault tolerant manner [21]. Since, the Fuzzy logic controller procedures rules determined by the user controlling the goal control system, it can be modified without difficulty to make better or intensely modify system performance. Fuzzy Logic handles the examination of knowledge by utilizing fuzzy sets, each of which can show a linguistic expression such as “low”, “adequate”, etc. [22].

## III. RELATED WORKS

Resource discovery schemes are divided into three types [6]; Centralized schemes, hierarchical schemes and distributed schemes. In this section we provide a brief overview of some works on resource discovery.

From the popular models which use Centralized scheme, we can point to Globus toolkit [7]. Amarnath Balachandar R., et al. [8] proposed a Globus based scheme for resource management in semantic grid whose focus is on the resource discovery. The semantic based component which has been used in this scheme is a centralized one. This semantic component has been used for service description and service discovery in this scheme. Gridbus broker does not know Monitoring and discovery system (MDS). Hence, semantic component sits in between them. The semantic component takes the information gathered by MDS and creates knowledge base. The discovery module in semantic component, queries the knowledge base to discover suitable resource that matches the requirements. Here, Algernon inference engine is used; a reasoner. A reasoner is needed to query knowledge base, just the same we use SQL to query database. This information is then given to job descriptor that creates application description file and resource description which is then used by Gridbus.

Habib Esmaeelzadeh R., et al. [9] proposed a scheme which uses resource grouping based on Quality of Services criteria. This resource grouping is based on delay, band width and semantics. Groups are arranged hierarchically. There are three geographical, bandwidth and semantic groups in the highest level of our system. The requests for resources are examined and allocated to one of the groups, and it is guided to the lower subgroups, until which it reaches the desired resource. The authors claim, that in their proposed approach, the delay, waiting time and response time are decreased.

To solve the problems of single point failure in centralized schemes, recently decentralized schemes most of which use P2P technologies have been used in grid environment. Shaikh A. K. et al. [23] proposed a semantic decentralized model which uses P2P chord protocol for grid resource discovery. As

Chord protocol has knowledge of a small amount of routing information of other peers with the aid of peer successor, it prevents overheads. A semantic method is used for identifying the relationship between resources. For obtaining semantic similarity between resources, this plan uses a semantic similarity equation which is explained in [24]. The equation is used for calculating semantic similarity between concepts.

Pirrò G. et al. [25] proposed a service discovery scheme in which the service information is disseminated with the usage of a DHT-based semantic overlay networks. This scheme combines DHT and SON and can be used in distributed basis schemes such as Grids and Clouds. The scheme presented in this work [25] allows semantic-driven query answering in DHT-based network by constructing a SON over a DHT. For enabling semantic service matchmaking over the combined DHT/SON network, a semantic similarity equation for obtaining the similarities between concepts has also been defined. The authors claim that this scheme improves accuracy of search and network traffic.

Heine F. et al. [26] proposed an ontology-based search scheme using DHTs for resource discovery in grid. A P2P overlay provides for a resource catalogue with the aid of DHT algorithms. Nodes give resource descriptions in ontologies according to the description logic, and each node can query the network for resources. The node's ontology is maybe incomplete, but it can be completed by ontologies of other nodes. The authors claim that this work has a good scalability for large number of concepts and nodes.

Li J. [27] proposed a grid resource discovery approach based on the Semantic Communities. In this approach, a Semantic structure which is based on small world is used for grouping the similar nodes; hence, the request for the resource is only sent to on the related nodes. Furthermore, this work proposes a new algorithm for effective resource information integration and searching in the network with the aid of semantic small-world. The author claims that this work vastly improves the search expressiveness, scalability, and accuracy.

Di Modica G. et al. [28] proposed a service discovery for service oriented architecture which uses semantic P2P structure. It consists of two steps; semantic overlay management and service discovery. In clustering step, nodes are grouped together in the overlay based on their resources semantic similarities. In service discovery step, a Semantic Query is routed to the nodes which are similar to the query from the semantic point of view. This work uses a semantic similarity function which is explained in [29] for computing the semantic similarities between entities.

Su Zhenglian et al. [30] proposed a model for web service discovery which uses semantic technology, fuzzy theory and multi-phase matching. This paper seeks to discover web services which have some imprecise feature, and perform fuzzy based reasoning. There are two types of services in this plan; concrete service level which performs all the service details and abstract service level which explains the major typical function of a set of services on concrete service level.

## IV. THE PROPOSED APPROACH

As p2p is used in our architecture, JXTA [31] is chosen as an infrastructure of our model. One of the features of JXTA is to grouping nodes based on some criteria such as the semantic similarity between their services. So JXTA provides some good facilities for node grouping. Another feature of JXTA gives a good facility in discovery phase. As JXTA has a two-layer architecture, the major search process is just in super peers so the overhead of the network will be reduced.

The features of semantic Overlay Network are used in the proposed model architecture. In the proposed model by considering delay as a major parameter of fuzzy system, we can expect to obtain an overlay with enhanced measure in Average Path Length and Clustering Coefficient. The amount of semantic similarity between nodes services is the other parameters of the fuzzy system. In fact we use delay, bandwidth and semantic similarity as the input parameters of fuzzy system to create semantic Overlay Network (SON).

Our scheme consists of two phases. Nodes grouping and resource (service) discovery. We use fuzzy theory in both phases. Our scheme uses a hybrid P2P structure where the nodes are divided into groups; in each group there is a coordinator, which holds the semantic similarity of the whole group. In grouping phase, nodes are divided based on the bandwidth of each node, delay and the service semantic similarities between each node and the super node of the groups. These three criteria are used for the input of the fuzzy system; SON will be created according to the output of the fuzzy system. A request must be routed to groups which are near the semantic characterization of the request itself. In service discovery phase, after finding the adequate group fuzzy logic is used with three parameters. Nodes with the most adequate computational capabilities, the adequate free space and the closest semantic similarity with the request will be selected for the request.

In our proposed model, a set of nodes are organized in a grid environment which follow the modeling and construction of ontology strategies in Web Services and annotation is performed on the web services of each node based on ontology. In other words, Concepts of each node are identified and will be placed in the ontology.

Our proposed method can be used in a Geographic Information System, whose job is to resolve issues such as identifying geographical locations that are close to each other. Grid environment will be divided to a geographical map according to geographical areas; and then its usage can be improved by grouping the nodes with similar resources and discovering them from the semantic point of view. We have only defined a few geographic concepts in this work. These concepts are: Latitude, longitude, country, street address, city, state/province, postal code, geographic operators, geographic objects, climate conditions and soil information. We have not needed go any further than that, so far.

### A. *nodes grouping and creating Semantic Overlay Network*

First the first node joins the network and creates the first group. It will be identified as the super peer of the group.

When the second node joins the network the semantic similarity between its services and the super peer services is calculated with the aid of [29] and will be considered as the first input parameter of fuzzy system. The second parameter is delay. We assume nodes which are near to the super peer have the low delay. Nodes near to the super cluster of a group geographically are put in the group for reducing delay. And finally bandwidth is the third parameter of the fuzzy system. These three criteria are being placed in the low, medium and high intervals which overlap. The output number of fuzzy system determines if the node can be joined to the group or not. If the node does not join the group it will create a new group and will be the super peer of this new group. Then the third node joins the network. The semantic similarity between its services and the super peers of each groups are be calculated and uses as the first parameter of fuzzy system. Its distance to the super peers is the second parameter which determines delay. Bandwidth is again the third parameter. According to the output of the fuzzy system nodes will be joined to the groups. This scenario is repeated for the next nodes. Figure 1 shows the node grouping in our scheme. In each phase after joining the new node, super peer will be selected based on the bandwidth of the nodes belonging to the group. In another word, a node with the highest bandwidth will be chosen as the super peer. As all the major searching process is done between the super peers, this strategy will improve the efficiency of the network. Each super peer has the knowledge of all the nodes which are inside to it. The semantic similarity of each group is calculated dynamically after joining the new node with the aid of the formula mentioned in [29].

Our model is based on fuzzy logic which uses some parameters as the input of fuzzy system. These values are some numbers which represent semantic similarity between the services of nodes, delay and bandwidth. Teases non-fuzzy numbers are the inputs of fuzzy inference system which is used for fuzzy reasoning. The output value of fuzzy inference system is a non-fuzzy number which represents if the node is adequate for joining the specific group or not. The figure 2 shows the used fuzzy inference system.

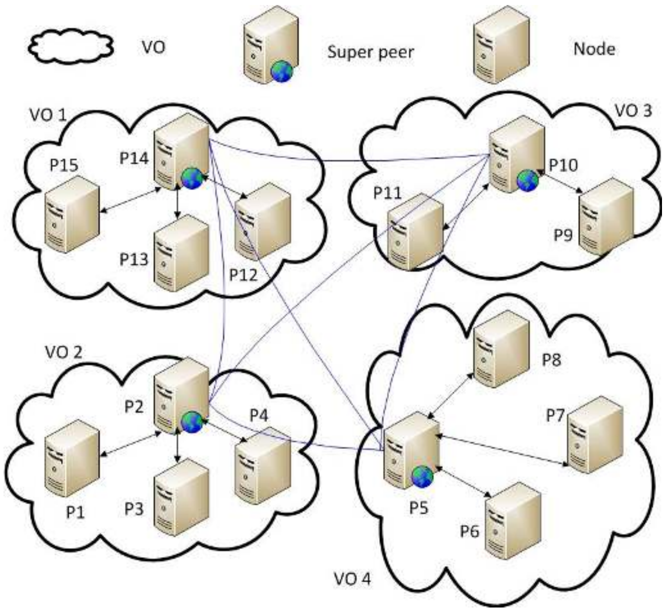


Fig. 1. the routine of node grouping

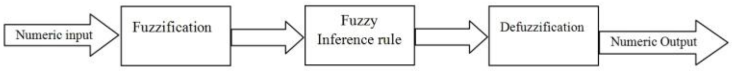


Fig.2. the structure of the fuzzy inference engine

There are two common types of fuzzy inference systems; Mamdani and Sugeno [32]. Mamdani inference system is used in our model due to its simplicity. Fuzzy inference system consists of five phases; fuzzification of the input values, implementation of the fuzzy operator in the antecedent, implication from the antecedent to the consequent, aggregation of the consequents across the rules, and defuzzification. The first step is receiving the inputs and determining the degree to which they belong to each of the adequate fuzzy sets through membership functions [33]. For doing this, three overlapping fuzzy sets will be created. Values between 0 and 0.5 are in low distance range, values between 0.2 and 0.8 are in the medium distance range and values between 0.5 and 1 are placed in high distance range. It is much better to select the intervals in a way that the end point of low be the starting point of high. A *membership function* is a curve that defines how each point in the input space is mapped to a membership degree between 0 and 1 [34]. $\mu$ represents the membership degree which is a number between 0 and 1. In general we have the following equation (1).

$$\mu_A(x) = \text{Degree}(x) \text{ in } A \quad (1)$$

$$\forall\, x \in X : \ \mu_A(x) : X \rightarrow [0,1]$$

Figures 3, 4 and 5 show the fuzzy sets for semantic, delay and bandwidth parameters which are created by using Matlab fuzzy logic toolbox [35].

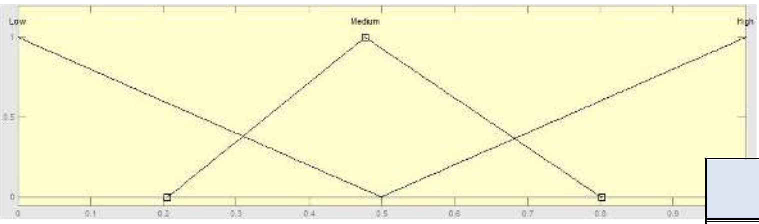


Fig.3. Fuzzy sets for semantic parameter

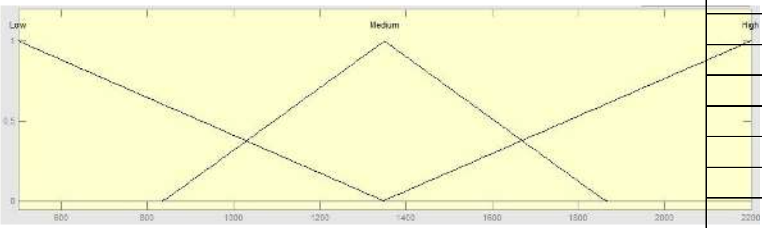


Fig. 4. Fuzzy sets for bandwidth parameter

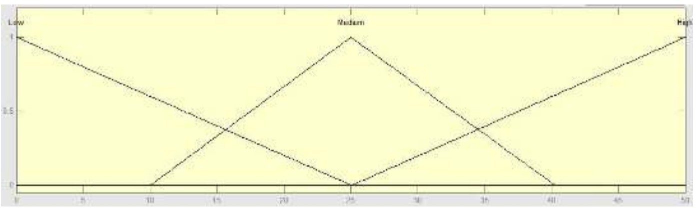


Fig. 5. Fuzzy sets for delay parameter

Table 1 some examples of fuzzy rules

| Semantic | Delay | Bandwidth | Result |
|---|---|---|---|
| Low | Medium | Low | Medium |
| Low | Medium | Medium | Low |
| Low | Medium | High | Low |
| Low | High | Low | Low |
| Low | High | Medium | Low |
| Low | High | High | Medium |
| Medium | Low | Low | Medium |
| Medium | Low | Medium | Medium |
| Medium | Medium | Low | Medium |
| Medium | Medium | Medium | Medium |
| Medium | High | Low | High |
| Medium | High | Medium | Medium |
| High | Low | Low | Medium |
| High | Low | Medium | High |
| High | Low | High | Medium |
| High | Medium | Low | High |
| High | Medium | Medium | High |
| High | Medium | High | High |

For example in figure 3, with the semantic similarity 0.4, the membership degree for low interval is 0.25, for medium interval is 0.75 and for high interval is 0. These values are used for fuzzy rules in fuzzy reasoning phase.
The fuzzy rules in Mamdani inference system can be created based on past experiences. In our model we assume that we have the following rules as shown in table 1.

According to the input parameters some rules will be fired. The fired rules should be integrated in a way that a decision can be made based on the aggregation of the fired rules. In Aggregation of fired rules phase the fuzzy sets that show the outputs of each fired rule are integrated into a single fuzzy set. This single fuzzy set is the input for the defuzzification phase. The output of defuzzification phase is a non-fuzzy number. There are five common defuzzification methods; centroid, bisector, middle of maximum, largest of maximum, and smallest of maximum [36]. Centroid method [37] is the most common used method. Centroid method is something like calculating the average of courses which is used for the fired rules. $\mu$ is similar to the number of units and represents the membership degree of fired rule. Equation 2 represents the centroid method.

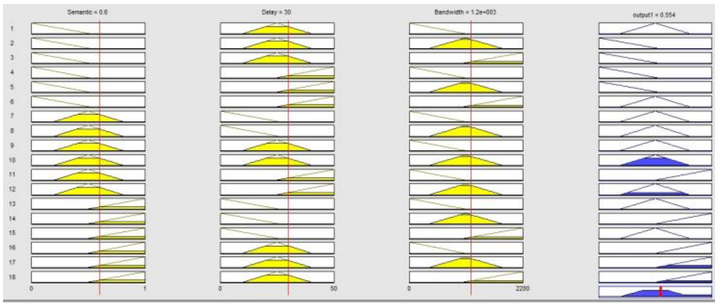

Fig.6. The process of aggragition of fired rules and deffuzzification of them in node grouping phase.

$$\alpha = \frac{\int_Z \mu_A(x) z dz}{\int_Z \mu_A(x) dz} \tag{2}$$

This non-fuzzy number is used for determining if the node is adequate for joining the group or not. If it is upper than a predefined threshold it will be joined the group. The value of threshold is based on each Virtual organization policies and the past experiences. Figure 6 shows the aggregation of fired rules and the defuzzification phase which is performed by Matlab fuzzy logic toolbox.

*B. service discovery*

When a request for service is issued, it should be directed to the group(s) that is (are) most relevant for the request. A request is directed to a group only if their degree of similarity is higher than a threshold. In most of the semantic similarity functions, this threshold is 0.69; so this value is used for the semantic similarity threshold of our plan. As we mentioned earlier the semantic similarity of the entire group is calculated and maintained in the super cluster of the group based on [29]. The request will be send to the super peer of the groups. The super peer calculates semantic similarity. If this value is higher than the threshold the request will be propagated to the peers belonging to the group itself or directed to other super clusters.

After determining the group(s), some peers which have the following two characteristics are chosen; first, their semantic similarity should be more than the predefined threshold. Second, they should have adequate computational capabilities. These two parameters are the input parameters for fuzzy system. In fact we use fuzzy theory with two parameters to discover the adequate resources. Computational capacity of the nodes is a parameter which should be considered in P2P based models because normally the jobs launched in P2P overlays have low communications between nodes and high execution of time. Figure 7 and 8 show respectively the fuzzy sets for computational capacity parameter and fuzzy sets for output variable which are created by using Matlab fuzzy logic toolbox. For semantic parameter we use the same fuzzy sets which are used in node grouping phase (figure 3).

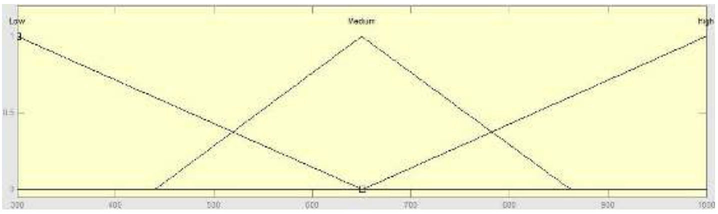


Fig. 7. Fuzzy sets for computational capacities

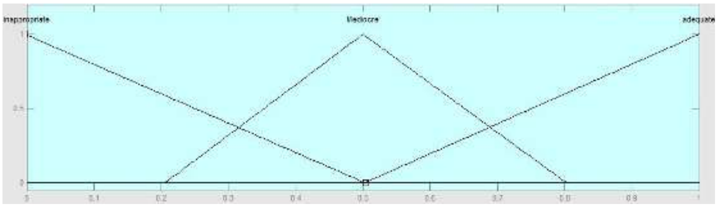


Fig. 8. Fuzzy sets for output variable

By utilizing these fuzzy sets the membership degree of each interval which is used in fuzzy reasoning is obtained. For fuzzy reasoning the rules in table 2 are used.

According to the input parameters some rules will be fired which will be integrated in aggregation of fired rules phase. The output of aggregation of fired rules phase is a single fuzzy set which is the input for the defuzzification phase. The output of defuzzification phase is a non-fuzzy number which is used for determining if the node has adequate services or not. If the output value is higher than a predefined threshold it shows that the node has adequate services. The value of threshold is again based on each Virtual organization policies and the past experiences. Figure 9 shows the aggregation of fired rules and the defuzzification phase which is performed by Matlab fuzzy logic toolbox.

Table 2- Fuzzy rules for service discovery phase

| Semantic | computational capacities | Results |
|---|---|---|
| Low | Low | Low |
| Low | Low | Medium |
| Low | Low | High |
| Low | Medium | Low |
| Low | Medium | Medium |
| Low | Medium | High |
| Low | High | Low |

## V. PERFORMANCE EVALUATION

In this section we express the results of the performed simulations. In these experiments the proposed method is compared with the similar approaches in several evaluation criteria with Gridsim simulator [38] and Matlab [39]. Grid resource ontology concepts and their semantic similarity threshold values have been outlined in GridSim. We have used a protégé ontology editor and framework [42] to build these Ontologies.

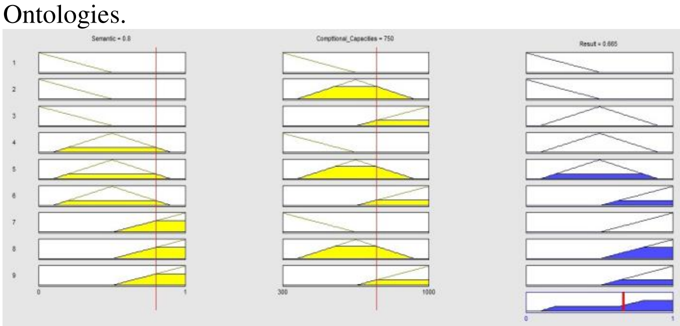

Fig.9. The process of aggragition of fired rules and deffuzzification of them in service discovery phase.

Grid resource ontology is used with the help of protégé OWL APIs. Subsequently, the semantic distance between these ontology concepts is evaluated with the help of the formula mentioned in [29]. To define these values in our Gridsim model, .csv file has been combined in our matching class and then the semantic threshold value under the geographic information specification such as soil information has been picked. We use the following parameters for the next experiments:

Total Concepts = 10

Semantic Threshold= 0.7

For simulation, the protocols of JXTA are not implemented in Gridsim. However, to keep some similarities with JXTA protocols, the message size of the Gridsim is considered as an approximate length in bytes of the JXTA headers and message length to take into account these protocols. In another word, there is no need to implement the protocols such as JXTA in

Gridsim. It should be taken into account that when a big Grid infrastructure is implemented in GridSim, the low level protocols, such as JXTA protocols are not necessary to implement. Only the organizations, the users and the management of resources (high level) are the relevant to the simulation.

In the first experiment, our model has been evaluated in term of the number of discovered resources. In this experiment the proposed method is compared with the semantic threshold based methods [28, 22]. In the semantic threshold base models, a predefined threshold is used for resource discovery. In another word, if the semantic similarity between the resource nodes and the request is higher than a threshold, the nodes will be selected as an adequate node for discovery. Consider a situation in which the semantic similarity between the request and a service node is a little lower than the threshold. In this case the node will not be discovered although it has good computational capacities. As in our approach we have used fuzzy theory with two parameters (semantic similarity and computational capacities) this problem is overcome and we can have more discovered resources. The result illustrates that although the precision of resource discovery is decreased a little, the number of adequate discovered resources is increased. As it is illustrated in figures 10 and 11, the semantic threshold based approaches discover no resources whose similarity is less than the threshold; however as our approach uses fuzzy rules which overlap some nodes which have lower similarity than the threshold but have good computational capacities will be discovered. In this experiment we choose 0.7 as the semantic similarity value. This value can be changed based on the user requirements and the management policies of each VO.

In the second experiment, our model has been evaluated in terms of response time. Response time is the time that a node takes to react to a request. Response time is reduced as we have considered delay in node grouping phase and computational capacities in discovery phase. In this experiment we calculate response time as a function of the number of discovered nodes. We compare this criterion of our approach with two semantic p2p based distributed discovery plan [25, 28] in which delay and computational capacities have no effect on discovery results. In [25] JXTA is used as a P2P protocol as well as Chord is used as a routing protocol in [28]. In our simulation we have passed up the routing protocols and JXTA and Chord is not implemented since the behavior of JXTA and Chord do not have any effect in the result of the simulation. In this way the response time for [25, 28] are similar to each other. We compared the response time of our approach with [25, 28] under the same number of discovered nodes. In this experiment the computational capacities criterion has more preference than the delay criterion. In [25, 28] nodes with any computational capacities are discovered but in our approach nodes with adequate computational capacities are discovered. Figure 12 shows this point. This figure shows the response time regarding to the number of discovered nodes. Clustering coefficient (CC) and Average Path Length (APL) are two criteria which are used for the third experiment. They are two major factors which determine the search expressiveness.

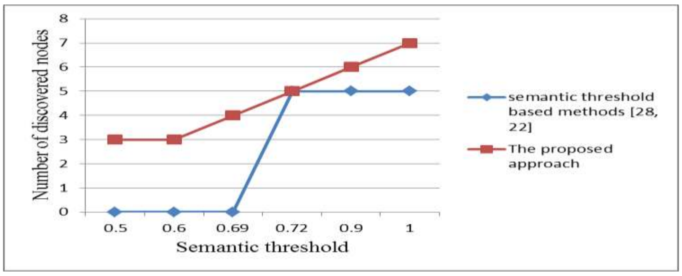


Fig. 10. The number of discovered resources regarding to different semantic threshold value.

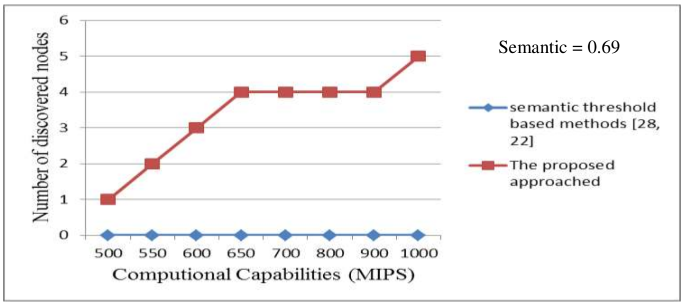


Fig. 11. The number of discovered resources regarding Computational complexity

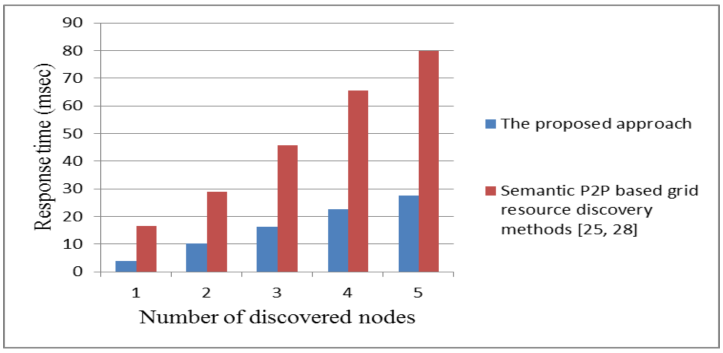


Fig. 12. Response time of our plan vs. other plans.

The clustering coefficient of a node is the ratio of the links between nodes within its neighborhood divided by the maximum number of possible links exists between them. Average path length is the average number of steps along the shortest paths for all possible nodes. We compare the proposed method with OntoSum [27] which is a semantic small world base grid resource discovery scheme and a random-walk based Gnutella scheme [41]. Since OntoSum is a semantic small world base scheme, it has great values in terms of APL and CC. The topology of network is a main factor for obtaining APL and CC; so we can claim that all schemes which use JXTA as an infrastructure have a similar value in terms of CC and APL with our plan. Figures 13 and 14 show the clustering coefficient and the average path length regarding to the number of nodes. As it is illustrated, the APL of the proposed approach is better than OntoSum, but it is a little higher than the random walk based scheme.

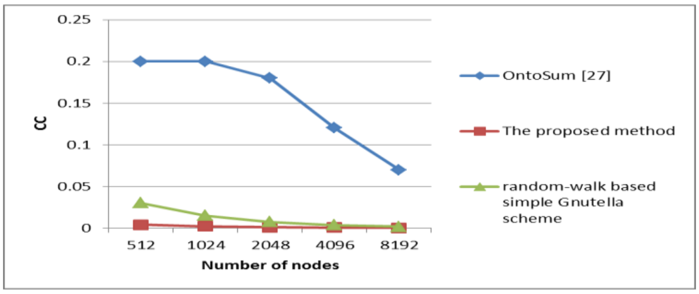


Fig. 13. Compression of clustering coefficient

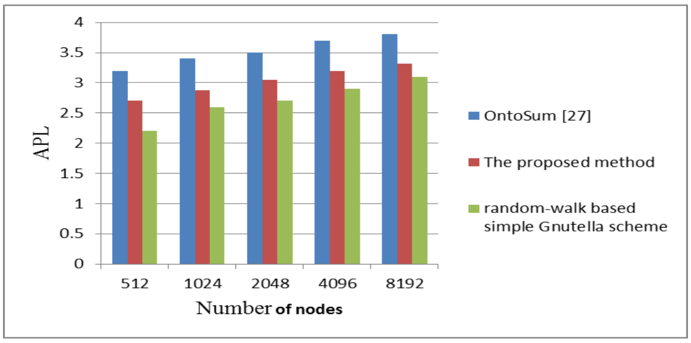


Fig. 14. Compression of Average Path Length

In the last experiment, the numbers of related and not related contacted nodes have been measured. We say a node is related if it provides resource that matches the request, and it is not related if it has been contacted but does not have any resource that matches the request. In Fig. 15 and 16, the numbers of relevant and not relevant nodes out of all visited nodes of our plan and a JXTA based plan [28] are shown for each query. In this experiment there are 100 nodes. As it is illustrated, the number of relevant nodes contacted in our plan is more than the similar plan [28] which uses semantic features for node grouping and service discovery. In both plans after determining the adequate group(s), the request will be propagated to the group's node. Since our plan discovered more adequate nodes, the number of related nodes contacted is more than [28]. In this experiment after node grouping phase, in our plan there will be 4 groups which consist of 18, 27, 32 and 23 nodes while there will be 4 groups which consist of 14, 17, 38 and 31 nodes in [28].

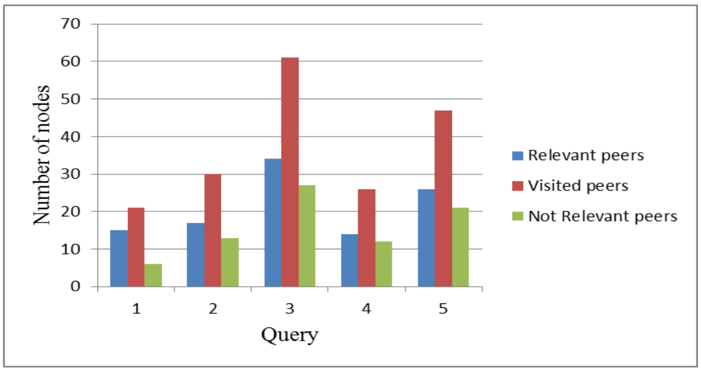


Fig. 15. Relevant vs. not relevant peers in our plan

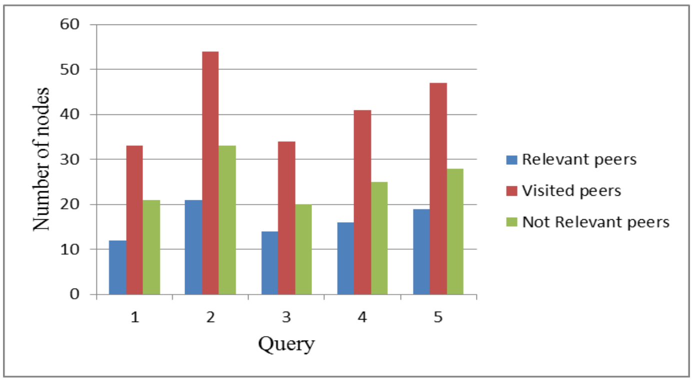


Fig. 16. Relevant vs. not relevant peers in [28]

## VI. Conclusion

In recent years a lot of attention has been paid to the utilization of semantic technology in grid resource discovery. Semantic Overlay Network is a novel overlay networks in which semantic technology is used for grouping nodes based on the semantic similarity between their resources. In this paper we propose a novel approach for grid resource discovery which is using fuzzy theory for creating semantic Overlay Network. Fuzzy theory is belongs to the intelligent approach which shows uncertainty in Phenomena. We also use fuzzy theory in resource discovery phase to discover the most adequate resources. In the proposed approach, Peers are grouped together in the network space based on fuzzy theory with three parameters; Delay, Bandwidth and characterization in the semantic space. Then the resource discovery is done by using some fuzzy input parameters like the computational capabilities of nodes and semantic similarity. We evaluated the performance of our approach with some distributed semantic based grid resource discovery models; the results of the experiments approved the efficiency of the proposed approach in term of scalability, precision, search expressiveness and response time.

ACKNOWLEDGMENT

The authors of this paper would like to thank Miroslaw Korzeniowski from Wroclaw University of Technology and Mr Damià Castellà (A researcher & research fellow at the University of Lleida) and Yunjia Li (PhD Student & Research Fellow at School of Electronics and Computer Science the University of Southampton) for their kindly comments and advice.

REFERENCES

[1]- B. Allcock, J. Bester, J. Bresnahan, A. L. Chervenak, I. Foster, C. Kesselman, S. Meder, V. Nefedova, D. Quesnal, S. Tuecke. “Data Management and Transfer in High Performance Computational Grid Environments,” Parallel Computing Journal, Vol. 28 (5), May 2002, pp. 749-771.

[2]- Foster, Ian, and Carl Kesselman, eds. “The grid 2: Blueprint for a new computing infrastructure,” Morgan Kaufmann, 2003.

[3]- Butt, Fouad, et al. “Scalable Grid Resource Discovery through Distributed Search,” arXiv preprint arXiv:1110.1685 , 2011.

[4]- Noghabi, Hossein Boroumand, et al. "An Optimized Search Algorithm for Resource Discovery in Peer to Peer Grid." Informatics and Computational Intelligence (ICI), 2011 First International Conference on. IEEE, 2011.

[5]- Xiong, Zenggang, et al. “Integrated agent and semantic p2p grid resource discovery model,” Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, in Proceeding of Eighth ACIS International Conference, Vol. 2. IEEE, 2007.

[6]- Rahman M, Ranjan R, Buyya R, Benatallah B. “A taxonomy and survey on autonomic management of applications in grid computing environments. Concurrency and Computation,” Practice and Experience 2011; 23(16), pp.1990–2019.

[7]- Foster, Ian, and Carl Kesselman. “Globus: A metacomputing infrastructure toolkit,” International Journal of High Performance Computing Applications 11.2,1997, pp. 115-128.

[8]- Amarnath, Balachandar R., et al. “Ontology-based Grid resource management,” Software: Practice and Experience 39(17) ,2009, pp. 1419-1438.

[9]- Rostam, Habib Esmaeelzadeh, Amir Masoud Rahmani, and Kamran Zamanifar.“Resource Management in Semantic Grid System Based on QoS,” in Proceeding of Second International Computer and Electrical Engineering Conference, Vol. 2. IEEE, 2009.

[10]- Weissman, J., and B-D. Lee. “The service grid: supporting scalable heterogeneous services in wide-area networks,” in Proceeding of Applications and the Internet Conference, IEEE, 2001.

[11]- Armbrust, Michael, et al.“A view of cloud computing,” Communications of the ACM 53(4), 2010, pp. 50-58.

[12]- Juric, Matjaz B., et al. “WSDL and UDDI extensions for version support in web services,” Journal of Systems and Software 82(8), 2009, pp.1326-1343.

[13]- Patil, Abhijit A., et al. “Meteor-s web service annotation framework,” in Proceedings of the 13th international conference on World Wide Web, ACM, 2004.

[14]- Martin, David, et al. “Bringing semantics to web services: The OWL-S approach,” Semantic Web Services and Web Process Composition, 2005, pp.26-42.

[15]- Akkiraju, Rama, et al. “Web service semantics-wsdl-s,” W3C member submission 7, 2005.

[16]- Li, Mei, Wang-Chien Lee, and Anand Sivasubramaniam. “Semantic small world: An overlay network for peer-to-peer search,” in Proceedings of the 12th IEEE International Conference, IEEE, 2004.

[17]- Jin, Hai, Xiaomin Ning, and Hanhua Chen. “Efficient search for peer-to-peer information retrieval using semantic small world,” in Proceedings of the 15th international conference on World Wide Web, ACM, 2006.

[18]- Hui, Ken YK, John CS Lui, and David KY Yau. “Small world overlay P2P networks,” 12fth IEEE International Workshop on Wuality of Service, IEEE, 2004.

[19]- De Roure, David, Nicholas R. Jennings, and Nigel R. Shadbolt. “The semantic grid: past, present, and future,” in Proceedings of the IEEE 93(3), 2005, pp. 669-681.

[20]- Flahive, Andrew, et al. “Ontology tailoring in the Semantic Grid,” Computer Standards & Interfaces, 31(5), 2009, pp. 870-885.

[21]- A.A. Allahverdiyev, “Application of Fuzzy-Genetic Algorithm for Solving an Open Transportation,” International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 7, Vol. 3, No. 2, pp. 119-123, June 2011.

[22]- A.A. Allahverdiyev, “Cargo Transportation Routing Under Fuzzy Conditions,” International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 6, Vol. 3, No. 1, pp. 45-48, March 2011.

[23]- Shaikh, Abdul Khalique, Saadat M. Alhashmi, and Rajendran Parthiban. “A semantic decentralized Chord-based resource discovery model for Grid computing,” in proceeding of 17th International Conference on Parallel and Distributed Systems (ICPADS), IEEE, 2011.

[24]- T. Andreasen, H. Bulskov and R. Knappe, “From ontology over similarity to query evaluation,” Proc. 2nd CologNET-ElsNET Symposium-Questions and Answer:Theoretical and Applied Perspective, 2003, pp. 39-50, Amsterdam, Holland.

[25]- Pirrò, Giuseppe, Domenico Talia, and Paolo Trunfio. “A DHT-based semantic overlay network for service discovery,” Future Generation Computer Systems, 2011.

[26]- Heine, Felix, Matthias Hovestadt, and Odej Kao. “Towards ontology-driven P2P grid resource discovery,” In Proceedings of 5th IEEE/ACM International Workshop on Grid Computing, IEEE, 2004.

[27]- Li, Juan. “Grid resource discovery based on semantically linked virtual organizations Grid Computing, Future Generation Computer Systems 26(3), 2010, pp. 361-373.

[28]- Di Modica, Giuseppe, Orazio Tomarchio, and Lorenzo Vita. “Resource and service discovery in SOAs: A P2P oriented semantic approach,” International Journal of Applied Mathematics and Computer Science, 21(2), 2011, pp. 285-294.

[29]- Bisignano, Mario, Giuseppe Di Modica, and Orazio Tomarchio. “Jaxson: A semantic P2P overlay network for web service discovery,” in Proceeding of World Conference on Services, IEEE, 2009.

[30]- Su, Zhenglian, Haisong Chen, Liang Zhu, and Yonghua Zeng. “Framework of Semantic Web Service Discovery Based on Fuzzy Logic and Multi-phase Matching,”,2012.

[31]- Gong, Li. “JXTA: A network programming environment,” Internet Computing, IEEE, vol. 5,No. 3, 2001, pp. 88-95.

[32]- Aguiar H, Junior O, Ingber L, Petraglia A, Rembold Petraglia M, "Fuzzy Modeling with Fuzzy Adaptive Simulated Annealing," Vol.35, Issue 3, pp. 973-980, springer, 2012.

[33]- Cheheltani, Seyyed Hadi, and Mohammad Mehdi Ebadzadeh. "Immune based fuzzy agent plays checkers game," Applied Soft Computing, 2012.

[34]- Sakthivel, N. R., V. Sugumaran, and Binoy B. Nair. "Automatic rule learning using roughset for fuzzy classifier in fault categorization of mono-block centrifugal pump," Applied Soft Computing, vol. 12, No. 1, 2012, pp.196-203.

[35]- Johanyák, Zs Cs, Domonkos Tikk, Szilveszter Kovács, and Kok Wai Wong. "Fuzzy rule interpolation Matlab toolbox-FRI toolbox," in proceeding of International Conference on fuzzy systems, IEEE, 2006, pp. 351-357.

[36]- Deng, Xingsheng, and Xinzhou Wang. "Incremental learning of dynamic fuzzy neural networks for accurate system modeling," Fuzzy Sets and Systems IEEE), vol.160, no. 7, 2009, pp.972-987.

[37]- Mendel, Jerry M. "On centroid calculations for type-2 fuzzy sets ," Appl. Comput. Math, vol. 10, no. 1, 2011, pp. 88-96.

[38]- Sulistio, Anthony, Uros Cibej, Srikumar Venugopal, Borut Robic, and Rajkumar Buyya. "A toolkit for modelling and simulating data Grids: an extension to GridSim," Concurrency and Computation: Practice and Experience vol.20, no. 13, 2008, pp. 1591-1609.

[39]- Etter, Delores M., and David C. Kuncicky. "Introduction to MATLAB," Prentice Hall, 2011

[40]- T. Hong, Chapter Fourteen: Performance, "Peer-to-Peer: Harnessing the Power of Disruptive Technologies," O'Reilly, 2001, pp. 203–241.

[41]- Gnutella website. http://gnutella.wego.com

[42]- Lasbleiz J., Brillet E., Decaux O., Duvauferrier R., "Staging disease with Protégé 4: Example of multiple myeloma," vol.32, Issue 6, Elsevier, 2011, pp. 329–331