Service Matching and Discovery in P2P Semantic Community *

Devis Bianchini Valeria De Antonellis Michele Melchiori Denise Salvi

Università degli Studi di Brescia - Dip. di Elettronica per l'Automazione
Via Branze, 38 - 25123 Brescia - Italy
{bianchin|deantone|melchior|salvi}@ing.unibs.it

Abstract. Recently, more and more web services are being made available and their use can potentially highly increase cooperation in P2P systems, where different partners intend to share and exchange digital resources. A major concern, in this framework, is related to the development of semantic driven techniques for description and organization of available services in semantic communities for discovery purposes. We propose an ontology-based Semantic Driven Service Discovery approach in P2P systems, where peers are organized in a semantic community to support effective automation of service discovery. In the approach ontologies are exploited both to improve precision and recall of search results by means of an innovative service matchmaking strategy and to organize services in the community by means of intra-peer and inter-peer semantic links. Some experimental results are briefly discussed.

1 Introduction

Recently, an ever-growing number of web services are being made available to increase cooperation and in P2P systems, where different partners intend to share and exchange digital resources. Automation of service discovery is a crucial aspect in this context. P2P systems are featured by lack of a common understanding of the world and high heterogeneity in resource description. A major concern is related to the development of ontology-based techniques for description and organization of available services in semantic communities for discovery purposes. However, members of a P2P semantic community cannot be constrained to use the same reference ontology. In our application scenario, each peer of the community adopts its own ontology (called *peer ontology*) to express domain knowledge related to service descriptions.

During community constitution, we assume that a peer called promoter spreads out a manifesto containing a relevant portion of its peer ontology, expressing a core domain knowledge for the community. A peer that aims at joining

* This work has been partially supported by the ESTEEM PRIN Project (http://www.dis.uniroma1.it/∼esteem/index.html) and TEKNE FIRB Project (http://www.tekne-project.it/), funded by the Italian Ministry of Education, University and Research.

the community matches its own peer ontology against the manifesto and replies to the promoter. Once the community is established, services registered on the community members can be searched through ontology-based techniques. When the peer receives a service request expressed in terms of a different peer ontology, an innovative ontology-based service matchmaking procedure must be applied and strategies to forward the request over the P2P network must be exploited to improve discovery efficiency and avoid network overloading.

In particular, in this paper we propose the Semantic Driven Service Discovery approach for open P2P systems, that applies the service matchmaking strategies introduced in [5]: (i) to improve discovery results working with different peer ontologies in a distributed scenario; (ii) to improve performances through the identification of semantic connections between the peers of the community (called *inter-peer semantic links*) by comparing their service descriptions. This ontological framework extends functionalities of traditional UDDI Registry, improving keyword-based searching strategies.

The paper is organized as follows: Section 2 shows how different peer ontologies can be used to enable service discovery in P2P systems; Section 3 presents the adopted matchmaking models, while Section 4 explains how these models are used to define semantic links for optimization purposes; Section 5 gives some experimental results; Section 6 compares our approach with related proposals in literature; finally, in Section 7 final considerations and future work are discussed.

2 Exploitation of peer ontologies

We consider a P2P semantic community, where each peer can be a service provider or a broker where service descriptions are published in a UDDI Registry extended through its own peer ontology to provide semantic knowledge related to service descriptions. Brokers are in charge of receiving service requests, applying matchmaking models on services published locally and forwarding requests to the other brokers. Service descriptions represent functional aspects, based on the UDDI and WSDL standards, in terms of service category, service functionalities (operations) and input/output messages (parameters). Each peer ontology is constituted by: (i) a *Service Functionality Ontology (SFO)*, that provides knowledge on the concepts used to express service functionalities (operations); (ii) a *Service Message Ontology (SMO)*, that provides knowledge on the concepts used to express input and output messages (parameters) of services. In general, a service request description could contain terms not defined in the peer ontology of the broker to which the request is sent. To deal with this problem, the peer ontology is extended by a *thesaurus* that provides terms and terminological relationships (synonymy SYN, broader/narrower term BT/NT and related term RT) with reference to names of concepts in the ontology. In this way, the resulting extended registry (called *Semantic Peer Registry*) enables enlarged matchmaking capabilities when looking for correspondences between elements in service descriptions and concepts in the ontology. Furthermore, we assume that a broker can serve more semantic communities, that is, its registry can contain descriptions of services offered by different communities.

The joint use of the peer ontology and the thesaurus allows to define the two basic concepts presented in the following.

Definition 1 (Name Affinity coefficient). *Given the thesaurus $\mathcal{TH}$, the Name Affinity coefficient between two terms $t, t' \in \mathcal{TH}$, denoted by $NA(t, t')$, is: (i) 1.0 if $t = t'$; (ii) $max_l(\tau(t \rightarrow^l t'))$ if $t \neq t' \wedge t \rightarrow^l t', l \geq 1$, where $t \rightarrow^l t'$ denotes a path of terminological relationships from t to t'; (iii) 0.0 otherwise. A weight $\sigma_{tr} \in [0, 1]$ is associated to each kind of terminological relationship tr, in order to evaluate its implication for name affinity; in our experimentation, $\sigma_{SYN} = 1$, $\sigma_{BT/NT} = 0.8$ and $\sigma_{RT} = 0.5$. The function $\tau(t \rightarrow^l t') = \prod_{k=1}^{l}(\sigma_{tr_k}) \in [0, 1]$ defines the strength of $t \rightarrow^l t'$ as the product of the weights of all terminological relationships in the path. Since between two terms in the thesaurus there can exist more than one path, the one with the highest strength is chosen. We say that t and t' have name affinity ($t \sim t'$) if and only if $NA(t, t') \geq \alpha$, where $\alpha > 0$ is a threshold given by experimental results to select only terms with high values of the name affinity coefficient.*

Definition 2 (Affinity-based subsumption test). *Given an atomic concept C in the peer ontology $\mathcal{PO}$, we define the set of terms in the thesaurus that have name affinity with a given concept as $C_{\mathcal{TH}} = \{T \in \mathcal{TH} \mid T \sim C\}$. Analogously, we define the set of concepts of $\mathcal{PO}$ that have name affinity with a term T in $\mathcal{TH}$ as $T_{\mathcal{PO}} = \{C \in \mathcal{PO} \mid T \in C_{\mathcal{TH}}\}$.*
Given the peer ontology $\mathcal{PO}$, the thesaurus $\mathcal{TH}$ and a pair of terms $T1$ and $T2$ used in service descriptions to denote service elements, $T1$ is subsumed by $T2$ with respect to $\mathcal{TH}$, denoted by $T1 \sqsubseteq_{\mathcal{TH}} T2$, if and only if there exists $C \in T1_{\mathcal{PO}}$ and $D \in T2_{\mathcal{PO}}$ such that $C \sqsubseteq D$ is satisfied in $\mathcal{PO}$. Note that we pose $T1 \equiv_{\mathcal{TH}} T2$ if both $T1 \sqsubseteq_{\mathcal{TH}} T2$ and $T2 \sqsubseteq_{\mathcal{TH}} T1$ hold.

3 Service matchmaking model

The Name Affinity coefficient and the affinity-based subsumption test are exploited to perform service comparison according to different matchmaking models [5]: (a) a *deductive model*, based on deduction algorithms for reasoning on service functional descriptions [4]; (b) a *similarity-based model*, where the degree of match between services is measured [6]. These matchmaking models are combined together to produce better searching results according to precision and recall values, as shown in Section 5.

In the deductive model, the affinity-based subsumption test is exploited to determine the match type between a service request $\mathcal{R}$ and a supplied service $\mathcal{S}$, considering separately service elements (categories, operations, parameters), as summarized in Figure 2. A formal description of these match types to infer the match between two service descriptions is given in [5]. During a pre-filtering phase, categories are used to select a set of services, called *candidate services*,

Entity-based similarity
$$ESim(\mathcal{R},\mathcal{S}) = \frac{2 \cdot A_{tot}(IN_{\mathcal{R}}, IN_{\mathcal{S}})}{\lvert IN_{\mathcal{R}}\rvert + \lvert IN_{\mathcal{S}}\rvert} + \frac{2 \cdot A_{tot}(OUT_{\mathcal{R}}, OUT_{\mathcal{S}})}{\lvert OUT_{\mathcal{R}}\rvert + \lvert OUT_{\mathcal{S}}\rvert} \in [0,2]$$
$IN_{\mathcal{R}}, IN_{\mathcal{S}}$ - input parameter names of $\mathcal{R}$ and $\mathcal{S}$ $OUT_{\mathcal{R}}, OUT_{\mathcal{S}}$ - output parameter names of $\mathcal{R}$ and $\mathcal{S}$ $A_{tot}(IN_{\mathcal{R}}, IN_{\mathcal{S}}) = \sum_{in^i_{\mathcal{R}} \in IN_{\mathcal{R}}, in^j_{\mathcal{S}} \in IN_{\mathcal{S}}} NA(in^i_{\mathcal{R}}, in^j_{\mathcal{S}})$ $A_{tot}(OUT_{\mathcal{R}}, OUT_{\mathcal{S}}) = \sum_{out^i_{\mathcal{R}} \in OUT_{\mathcal{R}}, out^j_{\mathcal{S}} \in OUT_{\mathcal{S}}} NA(out^i_{\mathcal{R}}, out^j_{\mathcal{S}})$
Operation similarity
$$OpSim(op^i_{\mathcal{R}}, op^j_{\mathcal{S}}) = NA(op^i_{\mathcal{R}}, op^j_{\mathcal{S}}) + \frac{2 \cdot A_{tot}(IN^i_{\mathcal{R}}, IN^j_{\mathcal{S}})}{\lvert IN^i_{\mathcal{R}}\rvert + \lvert IN^j_{\mathcal{S}}\rvert} + \frac{2 \cdot A_{tot}(OUT^i_{\mathcal{R}}, OUT^j_{\mathcal{S}})}{\lvert OUT^i_{\mathcal{R}}\rvert + \lvert OUT^j_{\mathcal{S}}\rvert} \in [0,3]$$
$IN^i_{\mathcal{R}}, IN^j_{\mathcal{S}}$ - input parameter names of the i-th operation of $\mathcal{R}$ and the j-th operation of $\mathcal{S}$ $OUT^i_{\mathcal{R}}, OUT^j_{\mathcal{S}}$ - output parameter names of the i-th operation of $\mathcal{R}$ and the j-th operation of $\mathcal{S}$
Functionality-based similarity
$$FSim(\mathcal{R},\mathcal{S}) = \frac{2 \cdot \sum_{i,j} OpSim(op^i_{\mathcal{R}}, op^j_{\mathcal{S}})}{\lvert OP(\mathcal{R})\rvert + \lvert OP(\mathcal{S})\rvert} \in [0,3]$$
$OP(\mathcal{R}), OP(\mathcal{S})$ - operations of $\mathcal{R}$ and $\mathcal{S}$
Global similarity
$$GSim(\mathcal{R},\mathcal{S}) = w_1 \cdot NormESim(\mathcal{R},\mathcal{S}) + (1-w_1) \cdot NormFSim(\mathcal{R},\mathcal{S}) \in [0,1]$$
$w_1 \in [0,1]$ - weight used to assess the relevance of each kind of similarity $NormESim(), NormFSim()$ - $ESim()$ and $FSim()$ normalized to the range $[0,1]$

Fig. 1. Similarity Coefficients between service descriptions $\mathcal{R}$ (request) and $\mathcal{S}$ (supply).

that have at least one associated category related in any generalization hierarchy with the category of $\mathcal{R}$ in peer ontology.

Similarity-based model applies the Name Affinity coefficient between service elements to quantify the service similarity. In particular, `exact` and `plug-in` matches denote that the supplied service completely fulfills the request and service similarity is set to 1.0 (full similarity); if `mismatch` occurs, the similarity value is set to 0.0; finally, `subsume` and `intersection` matches denote partial fulfillment of the request and similarity coefficients exposed in Figure 1 are computed. Matching results are only those services such that the match type is not `mismatch` and $GSim$ value is equal or greater than a global similarity threshold δ.

MatchType	Description
Exact	$\mathcal{S}$ and $\mathcal{R}$ have the same capabilities, that is, they have: (i) equivalent operations; (ii) equivalent output parameters; (iii) equivalent input parameters
Plug-in	$\mathcal{S}$ offers at least the same capabilities of $\mathcal{R}$, that is, names of the operations in $\mathcal{R}$ can be mapped into operations of $\mathcal{S}$ and, in particular, the names of corresponding operations, input parameters and output parameters are in any generalization hierarchy in the peer ontology
Subsume	$\mathcal{R}$ offers at least the same capabilities of $\mathcal{S}$, that is, names of the operations in $\mathcal{S}$ can be mapped into operations of $\mathcal{R}$; it is the inverse match type with respect to `plug-in`
Intersection	$\mathcal{S}$ and $\mathcal{R}$ have some common operations and some common I/O parameters, that is, some pairs of operations and some pairs of parameters, respectively, are related in any generalization hierarchy in the peer ontology
Mismatch	Otherwise

Fig. 2. Classification of matches

4 Definition and deployment of semantic links

We propose optimization strategies based on the use of semantic links between services. These strategies aim at reducing the number of service comparisons according to the matchmaking models introduced above, finding all local services that fulfill the request and decreasing the network overload. Two kinds of semantic links can be established: (a) semantic links between services belonging to the same peer (*intra-peer semantic links*); (b) semantic links between services belonging to different peers (*inter-peer semantic links*). In the latter case, the related peers are referred as *semantic neighbors*. A peer p_i is a semantic neighbor of another peer p_j with respect to a service description $\mathcal{S}_i$ if there exists a service description $\mathcal{S}_j$ published on p_j such that S_i and S_j are related by an inter-peer semantic link. A *semantic link* is defined as a 4-uple:

$$sl = < source, destination, MatchType, GSim > \quad (1)$$

where *source* and *destination* identify the services related through the semantic link, `MatchType` the match type and `GSim` the global similarity value. Semantic links are established once the semantic community is established. To do this, it sends a probe service request for each service it wants to make sharable; this probe service request contains the description of the service functional interface (categories, operations, I/O parameters). The probe service request is sent to all the other peers of the semantic community. Each peer receiving the probe service request matches it against its own service descriptions by applying the matchmaking techniques explained in the previous section. Note that, in a P2P context based on UDDI technology, a service is identified by: the *businessKey* and *serviceKey* in the UDDI Registry and the IP address of the host on which the UDDI Registry is located.

Service functional descriptions related by means of semantic links are stored in a service ontology for each peer of the semantic community. In particular, the peer knows its services, intra-peer semantic links between them and inter-peer semantic links towards its semantic neighbors. Intra-peer semantic links are exploited during service discovery on the broker that receives the request, to prune the set of candidate services and reduce the number of service comparisons. Inter-peer semantic links are used to select a subset of semantic neighbors of the broker that receives the request, making request forwarding over the community more efficient.

Pruning of candidate services. In this phase, we exploit previously established matching results and intra-peer semantic links to infer a match or a mismatch between request and candidate services. Our purpose here is to avoid the application of the hybrid matchmaking model if this is not strictly necessary, but at the same time to identify each offered service matching the request. Therefore, the expected result of the pruning process is an improvement of performances in the proposed approach. From an operative point of view, we suppose that the match type between the request $\mathcal{R}$ and a candidate service $\mathcal{S}_1$ offered on a given peer has been already established. Now, intra-peer semantic links relating $\mathcal{S}_1$ with other offered services can be considered. For instance, if match($\mathcal{R}$,$\mathcal{S}_1$) evaluation results in a `plug-in` match and there exists a semantic link between $\mathcal{S}_1$ and $\mathcal{S}_2$ that states that match($\mathcal{S}_1$,$\mathcal{S}_2$) is a `plug-in`, we can infer that also the match($\mathcal{R}$,$\mathcal{S}_2$) is a `plug-in`. Generally speaking, given a match type between $\mathcal{R}$ and $\mathcal{S}_1$ and a semantic link between $\mathcal{S}_1$ and $\mathcal{S}_2$ it can occur that:
- exactly one match type between $\mathcal{R}$ and $\mathcal{S}_2$ can be inferred, therefore it is not necessary to apply the matchmaking model to evaluate the match between $\mathcal{R}$ and $\mathcal{S}_2$;
- a mismatch between $\mathcal{R}$ and $\mathcal{S}_2$ can be inferred; also in this case it is not necessary to apply the matchmaking model;
- more than one match type between $\mathcal{R}$ and $\mathcal{S}_2$ can be inferred, allowing to restrict the set of possible match types;
- nothing can be inferred from the available information.

Figure 3 provides the details on all the different cases that can occur and shows the matches or mismatches that can be inferred in each case. The column headers of the table describe match types associated to a semantic link between $\mathcal{S}_1$ and $\mathcal{S}_2$. The row headers describe possible match types between $\mathcal{R}$ and $\mathcal{S}_1$. An cell is empty if nothing can be inferred in association with a given column and row.

Selection of semantic neighbors. Given a service request $\mathcal{R}$ sent to a peer p, the peer searches for suitable services in its own Semantic Peer Registry and retrieves a list $CS = \{\langle \mathcal{S}_1, GSim_1, mt_1 \rangle, \ldots \langle \mathcal{S}_n, GSim_n, mt_n \rangle\}$ of services with corresponding similarity values $GSim_i \geq \delta$ and match type mt_i different from `mismatch`.

If a service $\mathcal{S}_i \in CS$ presents an `exact` or a `plug-in` match with the request, then $\mathcal{S}_i$ satisfies completely the required functionalities and it is not necessary to forward the service request to semantic neighbors with respect to $\mathcal{S}_i$. Otherwise,

match($\mathcal{S}_1$,$\mathcal{S}_2$) **match($\mathcal{R}$,$\mathcal{S}_1$)**	**exact**	**plug-in**	**subsume**	**intersection**	**mismatch**
exact	(exact) 1.0	(plug-in) 1.0	(subsume) $GSim(\mathcal{S}_1, \mathcal{S}_2)$	(intersection) $GSim(\mathcal{S}_1, \mathcal{S}_2)$	(mismatch) 0.0
plug-in	(plug-in) 1.0	(plug-in) 1.0	- -	(plug-in OR intersection OR mismatch) -	(mismatch) 0.0
subsume	(subsume) $GSim(\mathcal{R}, \mathcal{S}_1)$	(exact OR plug-in OR intersection OR subsume) -	(subsume) -	(subsume OR intersection) -	(subsume OR intersection OR mismatch) -
intersection	(intersection) $GSim(\mathcal{R}, \mathcal{S}_1)$	(plug-in OR intersection) -	(intersection OR subsume OR mismatch) -	- -	(intersection OR subsume OR mismatch) -
mismatch	(mismatch) 0.0	(plug-in OR intersection OR mismatch) -	(mismatch) 0.0	(plug-in OR intersection OR mismatch) -	- -

Fig. 3. Exploitation of intra-peer semantic links to prune the set of candidate services.

if $\mathcal{S}_i$ presents a `subsume` or an `intersection` match with the request, the peer p forwards the request to those peers that are semantic neighbors with respect to $\mathcal{S}_i$. Peer p does not consider semantic neighbors that present a `subsume` or an `exact` match with $\mathcal{S}_i$, because this means that they provide services with the same functionalities or a subset of $\mathcal{S}_i$ functionalities and they cannot add further capabilities to those already provided by $\mathcal{S}_i$ on the peer p. This phase is repeated for every $\mathcal{S}_i \in CS$. Semantic neighbors which present inter-peer links with any service $\mathcal{S}_j$ stored on p, but not included in CS, are discarded since they are not relevant with respect to $\mathcal{R}$. Each selected semantic neighbor sn presents a set of k inter-peer semantic links with some services on p that are suitable for $\mathcal{R}$. It is described as $\langle sn, \{\langle \mathcal{S}_1, GSim_1, mt_1\rangle, \ldots \langle \mathcal{S}_k, GSim_k, mt_k\rangle\}\rangle$, where $\mathcal{S}_1 \ldots \mathcal{S}_k \in CS$ and have a semantic link with some services stored on sn, featured by $GSim_1 \ldots GSim_k$ similarity degree and $mt_1 \ldots mt_k$ type of match, respectively. Note that this formulation holds even if sn has more than one service related to the same service $\mathcal{S}_i \in CS$.

Since the relevance of sn does not depend only on the similarity associated to the semantic links between p and sn, but also on the similarity degree between $\mathcal{S}_i \in CS$ and $\mathcal{R}$, the harmonic mean is used to combine these two aspects. Therefore, the relevance of a semantic neighbor sn is defined as:

$$r_{sn} = \frac{1}{k} \sum_{i=1}^{m_{sn}} \frac{2 * GSim_i * GSim(\mathcal{R}, \mathcal{S}_i)}{GSim_i + GSim(\mathcal{R}, \mathcal{S}_i)} \quad (2)$$

Relevance values are used to rank the set of semantic neighbors in order to filter out not relevant semantic neighbors (according to a threshold-based mechanism) and to further constrain the request forwarding (according to a token-based strategy). The harmonic mean is used here to combine the similarity values between the request and locally offered services with the similarity values of semantic links.

5 Experimental evaluation

Our approach is supported by the COMPAT MatchMaker Version 1.0, implemented in Java as a Web application. The MatchMaker exploits reasoning facilities of RACER OWL-DL reasoner Version 1.9. We applied the MatchMaker on a set of 480 web services specified in WSDL1.1 and covering touristic, *e*-banking, ERP and geographic application domains and a set of 10 queries formulated using both terms extracted from the peer ontology and terms that are not defined as atomic concepts of the ontology.

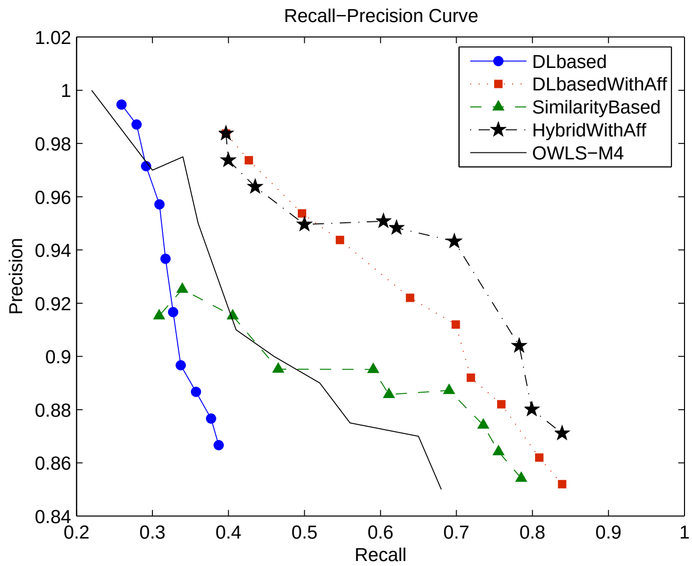

Fig. 4. Comparison of hybrid, deductive and similarity-based matching models on the basis of recall-precision performances.

Figure 4 shows the Precision-Recall curves by applying the deductive (with and without affinity-based subsumption test), the similarity-based and the hybrid matchmaking models, compared with another hybrid approach exposed in [7]. We evaluated the average values of precision and recall over the service requests by varying the global similarity threshold δ to filter out relevant results. Application of deductive matchmaking model presents high precision values, further improved by the joint use of peer ontology and thesaurus in the affinity-based subsumption test. Similarity-based matchmaking model is featured by lower precision values, although it is better than traditional IR metrics due to the use of similarity coefficients explicitly meant for service comparison. The best performances are obtained by applying the hybrid matchmaking model. Figure 5 depicts the response time for different matchmaking models, showing that the hybrid matchmaking performs quite linearly with the number of considered services. Therefore, improvement in precision and recall offered by matchmaking model is paid in term of a higher response time with respect to other approaches. However, further experimentations aim at demonstrating that the application of the proposed optimization rules based on semantic links will improve the overall performances of the discovery approach due to the lower number of comparisons among service requests and advertisements.

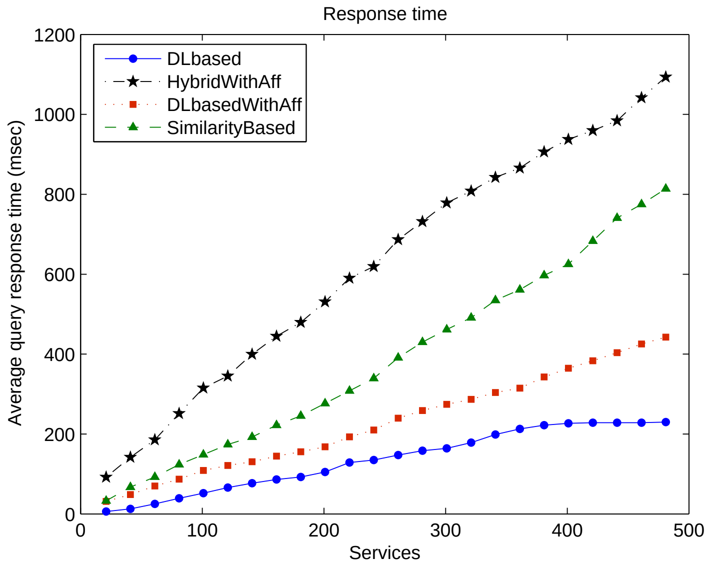

Fig. 5. Average response time of hybrid, deductive and similarity-based matching models.

6 Related work

Semantic description of services in a P2P context and semantic links between peers based on services they provide is a crucial aspect to improve effectiveness and performance of peer discovery. Several proposals have been made in literature to enable semantic-enhanced service discovery. Their main features are summarized in Table 1. Some of these approaches use centralized ontologies: GloServ [1] defines a predefined *skeleton ontology*, represented as a taxonomy of concepts each of them representing a category of services, that in turn is associated to an high level server. Each server represents a P2P network of nodes organized via a *Content Addressable Network (CAN)* [10]. These peers provide services belonging to the category represented by the ontological concept associated to their own server. The *skeleton ontology* is a centralized structure, replicated and cached in each of the high level servers.

Table 1: Comparison between approaches for P2P service semantic discovery

System	Network architecture	Semantic infrastructure	Service publishing	Service discovery
METEOR-S [12] focus on: - semantic overlay	hybrid P2P; special peers are introduced to handle a global ontology;	- a centralized *registries ontology* to classify peer registries; - local *domain specific ontologies* to provide service semantics; - no semantic links between peers; - semantically enriched WSDL to describe services	- service semantic descriptions are kept in semantically enhanced UDDI registries	two phases: (a) browsing of the *registries ontology* to find the proper registry; (b) deductive-based service profile matching based on I/O comparison
DAML-S for P2P [9] focus on: - service matchmaking	pure P2P	- local DAML-S ontologies to describe services; - no semantic links between peers	- service semantic descriptions are kept in local UDDI registries extended with DAML-S ontologies	deductive-based service profile matching based on I/O comparison
WSPDS [3] focus on: - semantic overlay	pure P2P	- local ontologies to provide service semantics; - semantic links between peers based on similarity of services they provide; - service semantic descriptions are made according to semantically enriched WSDL;	- service semantic descriptions are kept in semantically enhanced local UDDI registries	deductive-based service profile matching based on I/O comparison
GloServ [1] focus on: - semantic overlay - network architecture	hybrid P2P network, with high level servers and peers in a Computer Addressable Network (CAN) organized according to a concept taxonomy	- centralized ontologies (*skeleton ontology* and *thesaurus*) replicated in each high level service; - CAN lookup table to store information about peer contents inside the CAN networks; - service semantic descriptions are made according to semantically enriched WSDL	- services are provided by peers organized in the CAN networks	two phases: (a) browsing of the taxonomy of concepts to find the proper high level server; (b) keyword-based search through the CAN lookup table
WSMO-based P2P [11] focus on: - network architecture	pure P2P; network is organized as n-dimensional hypercube (according to the HyperCup model) with the purpose of improving routing efficiency	- hypercube infrastructure built according to an ontology; - peers are clustered according to an ontology concept and form a hypercube dimension; - service semantic descriptions are made according to the WSMO model	- service descriptions are kept in the WSMX registry; - service organization is flat	keyword-based matching on the non-functional service description
Artemis [2] focus on: - semantic overlay - medical domain specific services	hybrid P2P, with mediator super-peers: each peer registers to a mediator and sends it both advertisements and requests	- each peer has a coarse-grained *Service Functionality Ontology (SFO)*, to classify services, and a fine-grained *Service Message Ontology (SMO)*, to annotate services with medical concepts, based on medical information standards; each mediator super-peer keeps reference SFO and SMO and stores mappings among these and corresponding ontologies of the peers	- service descriptions are stored in UDDI registries on the mediator and annotated with ontology concepts saved in an ebXML registry on the mediator	(a) a peer sends a request to its reference mediator expressed in terms of its own ontologies; (b) the mediator uses ontology mappings to find matching services in its local registries; (c) the mediator also forwards the request to the other mediators

Service discovery is organized in two phases: (a) browsing of the taxonomy of concepts to find the proper high level server; (b) keyword-based search through the CAN lookup table.

WSPDS [3] describes a P2P network where peers have local DAML-S ontologies to provide service semantics and semantic links with other peers based on similarity of services they provide. When a request is submitted to a peer, it searches for local matching results and forwards the request to all the semantic neighbors, independently of the current request or the local results of the query. Also in WSMO-based P2P [11] the centralized ontology is used to organize peers of the P2P network as a n-dimensional hypercube (according to the HyperCup model); service semantic descriptions are made according to the WSMO model [8]. Service descriptions are stored in a registry and there is no use of semantic links between peers. METEOR-S [12] uses a centralized organization of peer registries, where service descriptions are kept in UDDI Registries semantically enhanced with local *domain specific ontologies*, while a centralized *registries ontology* is used to classify peer registries. During the discovery process, the *registries ontology* is browsed to find the proper registry to which submit the request. Sometimes, to overcome the heterogeneities between ontologies, it is required a manual-defined mediator-based architecture. For example, ARTEMIS [2] defines a network of peers, each of them has a coarse-grained functionality ontology to classify services and a fine-grained message ontology to annotate services with medical concepts, based on medical information standards. Each peer registers in a mediator super-peer the services it provides. A peer sends a request to its reference mediator expressed in terms of its own ontologies; mediator uses ontology mappings to find matching services in its local registries and also forwards the request to other mediators. Rarely semantic links between peers are considered and so the request is broadcasted on the network increasing its overload. In DAML-S for P2P [9] it is considered a network of peers, where each member stores a local DAML-S ontology to describe services; service semantic descriptions are kept in local UDDI registries extended with DAML-S ontologies. No semantic links are maintained between peers that provide similar services. Moreover, if a peer does not satisfy a request a flooding mechanism is used to find candidate services on the other peers of the network.

Original contribution of our approach relies on the application, in a flexible way, of similarity-based and deductive matchmaking models to increase precision and recall. Deductive service matching is performed with a joint use of information coming from peer ontologies and terminological aspects to deal with different reference ontologies in a P2P semantic community. Moreover, semantic links between services are exploited to speed up the discovery process and to make more efficient service request propagation on the network.

7 Concluding remarks

In this paper, we proposed the application of ontology-based strategies: (i) in an innovative service matchmaking technique, that combines different matchmaking models to improve discovery results; (ii) in the organization of services by means of semantic links, in order to improve performances during service discovery in a semantic community built on the top of a P2P network. Heterogeneity in P2P systems is considered, without constraining peers to use the same reference ontology. Further experimentation will evaluate the impact of the proposed approach on service discovery in P2P networks according to well-known parameters (such as network overloading) and concrete applications (e.g., scientific collaboration in medicine between healthcare organizations) will be studied.

References

1. K. Arabshian and H. Schulzrinne. An Ontology-based Hierarchical Peer-to-Peer Global Service Discovery System. *Journal of Ubiquitous Computing and Intelligence (JUCI)*, 2006.
2. The ARTEMIS Project: A Semantic Web Service-based P2P Infrastructure for the Interoperability of Medical Information Systems. `http://www.srdc.metu.edu.tr/webpage/projects/artemis/`.
3. F. Banaei-Kashani, C. Chen, and C. Shahabi. WSPDS: Web Services Peer-to-Peer Discovery Service. In *Proc. of the Int. Conference on Internet Computing (IC'04)*, pages 733–743, Las Vegas, Nevada, USA, 2004.
4. D. Bianchini, V. De Antonellis, and M. Melchiori. Capability Matching and Similarity Reasoning in Service Discovery. In *CAiSE Int. Workshop on Enterprise Modeling and Ontologies for Interoperability, EMOI 2005*, Porto, Portugal, June 2005.
5. D. Bianchini, V. De Antonellis, M. Melchiori, and D. Salvi. Semantic-enriched Service Discovery. In *IEEE ICDE Int. Workshop on Challenges in Web Information Retrieval and Integration, WIRI 2006*, Atlanta, Georgia, USA, 2006.
6. D. Bianchini, V. De Antonellis, B. Pernici, and P. Plebani. Ontology-based Methodology for *e*-Service discovery. *Journal of Information Systems, Special Issue on Semantic Web and Web Services*, 31(4-5):361–380, June-July 2006.
7. B. Fries, M. Khalid, M. Klusch, and K. Sycara. OWLS-MX: Hybrid OWL-S Service Matchmaking. In *Proc. of First International AAAI Symposium on Agents and Semantic Web*, Arlington, VA, USA, November 2005.
8. U. Keller, H. Lausen, and D. Roman. *Web Service Modeling Ontology (WSMO).* WSMO Working Draft, March 2004. `http://www.wsmo.org/2004/d2/v02/20040306/`.
9. M. Paolucci, K.P. Sycara, T. Nishimura, and N. Srinivasan. Using DAML-S for P2P Discovery. In *Proceedings of the Int. Conference on Web Services (ICWS2003)*, 2003.
10. S. Ratnasamy, P. Francis, M. Handley, R.M. Karp, and S. Shenker. A Scalable Content-Addressable Network. In *Proc. of the ACM SIGCOMM'01 Conference on Applications, Technologies, Architectures and Protocols for Computer Communication*, pages 161–172, San Diego, CA, USA, August 2001.
11. I. Toma, B. Sapkota, J. Scicluna, J.M. Gomez, D. Roman, and D. Fensel. A P2P Discovery mechanism for Web Service Execution Environment. In *Proc of the 2nd WSMO Implementation Workshop (WIW'05)*, Innsbruck, Austria, June 6th-7th 2005.
12. K. Verma, K. Sivashanmugam, A. Sheth, A. Patil, S. Oundhakar, and J. Miller. METEOR-S WSDI: A Scalable Infrastructure of Registries for Semantic Publication and Discovery of Web Services. *Journal of Information Technology and Management, Special Issue on Universal Global Integration*, 6(1):17–39, 2005.