On-the-fly Collaboration in Distributed Systems through Service Semantic Overlay

Devis Bianchini, Valeria De Antonellis, Michele Melchiori, Denise Salvi
Università di Brescia
Dip. Elettronica per l’Automazione
Via Branze, 38
25123 Brescia - Italy
{bianchin|deantonel|melchior|salvi}@ing.unibs.it

ABSTRACT
In the recent years distributed architectures and P2P technology have been adopted to better support effective collaboration among networked organizations. According to the P2P paradigm, many autonomous peers need to cooperate under highly dynamic conditions by sharing their resources (such as data and services), without sharing a common semantic model or ontology and without having a-priori knowledge about each other. In such a scenario, for effective collaboration, new advanced tools are required to support semantic representation and discovery of distributed resources such as data and services. In this paper we focus on semantic collaboration and service discovery in P2P systems. We propose the construction of a service semantic overlay over the P2P network, that is, a dynamic conceptual map across peers that provide similar services and constitute synergic service centers.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

Keywords
Service Semantic Discovery, P2P, Ontology-based matchmaking

1. INTRODUCTION
In the recent years, service-oriented paradigm evolved towards P2P technologies to better support organization collaboration in dynamic and open environments. Many autonomous peers need to cooperate by sharing their resources (such as data and services), without sharing a common semantic model or ontology and without having a-priori knowledge about each other. Moreover, peers collaborate under highly dynamic conditions, joining and leaving the network at any moment and the computation to search for services on the network should be distributed to different peers to ensure scalability. According to this view, in peer-based collaborative systems, effective communication and exchange of data and services under highly dynamic conditions are challenging activities. From the service perspective, availability of semantic representation is a key issue to support effective discovery and integration/composition of services. Ontologies provide the benefits of formal specifications and inference capabilities apt to improve semantic matching. For service analysis and automated matchmaking, we have proposed a methodology and the Semantic Driven Service Discovery (SDSD) approach characterized by a novel hybrid matchmaking and discovery environment [4]. However, in a P2P scenario, new advanced tools are required to support semantic representation and discovery of distributed resources.

In literature, P2P semantic-driven resource discovery has attracted much attention from Web services and Semantic Web area and relies on several efforts in related research fields, such as data integration and emergent semantics in P2P environments [3] to go beyond limitations of centralized architectures. Recent works involved structured [6] and unstructured P2P architectures [1]. The absence of any global ontology to overcome semantic heterogeneity has been faced recurring to centralized data structures to classify peer registries [7], to mediation-based architectures, where mediator super-peers store available services and have their own reference ontologies, to the adoption of semantic communities, intended as sets of peers that share similar interests and reach a semantic agreement on shared resources [2].

Our approach focuses on unstructured architectures without constraining to global ontologies/registries or mediation-based architectures. In particular, we propose the construction of a service semantic overlay over the logical network, that is a conceptual dynamic model of peers that provide similar services and constitute synergic service centers.

In this paper we emphasize the role of the peer knowledge model in service discovery. The paper is organized as follows: Section 2 describes the service semantic overlay and how to apply semantic matchmaking techniques to establish it; Section 3 illustrates how to maintain the service semantic
overlay in a dynamic P2P-based environment; the service discovery process based on the semantic overlay is sketched in Section 4 with the help of some examples extracted from the application scenario, while Section 5 presents concluding remarks and future work.

1.1 Application Scenario

Let consider an application scenario, where healthcare organizations (laboratories, drug-stores, research centres, hospitals) and doctors collaborate to provide useful services (i.e., drug ordering, diagnosis services, on-line consultation) in emergency situations, in which patient care services must be quickly found (for example, for anaphylactic shock treatment), but peers keep joining and leaving the network (for example, a doctor connects and disconnects from the network with his/her laptop). Organizations and doctors constitute the nodes of a P2P network, where each peer can play two roles: (i) it makes available discovery functionalities for services provided on the network (broker); (ii) it supplies a service, publishing it on a broker (provider). Brokers and providers can also act as requesters, searching for services on the network. Brokers are in charge of establishing and maintaining semantic links towards other brokers and performing service search.

2. PEER KNOWLEDGE AND THE SERVICE SEMANTIC OVERLAY

In the P2P-SDSD framework independent peers cooperate in a common domain by sharing services without any a-priori reciprocal knowledge. In such a scenario, no centralized authorities are defined to provide a comprehensive view of the shared services in the collaborative network. Each peer has local knowledge about the services it handles and local P2P knowledge about peers in the network with similar services.

2.1 Peer knowledge conceptual model

Each collaborative peer in the network has a local knowledge infrastructure constituted by: (i) UDDI Registry, where services are published together with their WSDL document and service providers register themselves; (ii) Peer Ontological Knowledge, that provides a conceptualization of abstract service operations and I/O parameters through a given domain ontology; a conceptualization of service categories through a Service Category Taxonomy (SCT), extracted from available standard taxonomies (e.g., UNSPSC, NAICS) and a conceptualization of peer neighborliness through semantic overlay.

In the UDDI Registry, a Provider peer is registered as businessEntity, that provides businessServices, that are associated to corresponding WSDL document through the tModel mechanism. WSDL standard distinguishes among the abstract service interface (operations and I/O parameters) and concrete bindings. From the abstract service interface a semantic description (Semantic service interface class) is generated, where the operations and I/O parameters are concepts of the domain ontology. Mappings between the operations and I/O parameters in the WSDL interface and the corresponding concepts in the semantic service interface are maintained. Finally, each semantic service interface is classified into Service Categories defined in the SCT.

Semantic service interfaces are related to each other according to semantic links between similar services. In particular, an Intra-peer semantic link relates a pair of similar services published on the same peer, while an Inter-peer semantic link relates a locally published service with a similar service published on a Semantic neighbor of the peer. In both cases, semantic links are established by applying semantic matchmaking techniques [4] and are labeled with matching information, as explained in the next sections. The peer ontological knowledge is represented using OWL-DL language. The conceptual model of the peer knowledge infrastructure is shown in Figure 1.

![Figure 1: Peer knowledge infrastructure conceptual model.](image)

2.2 Matching techniques for semantic link definition

Each broker is endowed with an ontology-based matchmaker used to support service discovery and semantic link definition. In [4] we defined a hybrid matchmaking strategy. In the hybrid model, a deductive matchmaking model is combined with a similarity-based model to compare services on the basis of their semantic-enriched functional interface.

The SCT and the domain ontology are exploited by the matchmaker to identify matching services. Moreover, the domain ontology is augmented by a thesaurus containing terms that are related by terminological relationships (as synonymy or hyponymy) to the names of concepts and categories. By means of the thesaurus, matching capabilities based on the domain ontology and categories are extended.

The deductive matchmaking is used to qualify the kind of match MatchType($S_1$, $S_2$) between two semantic service interfaces $S_1$ and $S_2$. According to this matchmaking model, it is possible to state if $S_1$ and $S_2$ provide the same functionalities ($S_1$ Exact $S_2$), if $S_1$ provides additional functionalities with respect to $S_2$ ($S_1$ Extends $S_2$) or vice versa, if there is a non empty intersection between functionalities provided by $S_1$ and $S_2$ ($S_1$ Intersects $S_2$) or if $S_1$ and $S_2$ have nothing in common ($S_1$ Mismatch $S_2$).

In case of partial overlapping among service functionalities (Extends or Intersects) the similarity-based matchmaking model is used to quantify service similarity $Sim(S_1, S_2) \in [0, 1]$ through coefficients properly defined to compare service interfaces. Otherwise, if $S_1$ Exact $S_2$ or $S_1$ Mismatch $S_2$, $Sim(S_1, S_2) = 1.0$ or $Sim(S_1, S_2) = 0.0$, respectively. Two semantic service interfaces are similar (denoted with $S_1$ $\approx$ $S_2$) when MatchType($S_1$, $S_2$) is not Mismatch and $Sim(S_1, S_2)$ $\geq$ $\delta$, where $\delta$ is a similarity threshold experimen-


2.3 On-the-fly semantic collaboration and service semantic overlay definition

The semantic service interfaces and the set of semantic links between them constitute the service semantic overlay built on the logical network overlay, that is, the P2P network. Each broker has its own local view (the links to semantic neighbors) on the semantic overlay. The service semantic overlay can be seen as a global collaboration view that dynamically emerges from local on-the-fly interactions between peers. In fact, we define a probe collaboration model in which a probe service query is submitted by a peer with the intention to find peers with semantically related services.

Example 2.1. Figure 2 shows an example of service semantic overlay as viewed by a peer $P_X$. Peer $P_X$ recognizes $P_Y$ as its semantic neighbor through a sequence of local interactions between $P_X$ and $P_Y$. In fact, to find semantic neighbors, the peer $P_X$ sends a probe service request for each service $S_{iX}$ published on it to other peers connected at the logical network overlay (in the example, $P_Y$ and $P_Z$). The probe service request contains the interface of $S_{iX}$ and the IP address of $P_X$. The peer $P_Y$ and $P_Z$ match the probe service request against their own semantic service interfaces by applying the matching techniques and obtain for each comparison the MatchType and the similarity value. If the MatchType is not MISMATCH and the similarity value exceeds the threshold $\delta$, an inter-peer semantic link can be established. In this example, only $P_Y$ finds a semantic service interface matching the request (in this example, $S_{iY}$) and replies to $P_X$ sending only $S_{iY}$ together with the MatchType and the similarity value. $P_X$ establishes an inter-peer semantic link, denoted with isl$_{P_X \rightarrow P_Y}$ ($S_{iX}, S_{iY}$), and $P_Y$ is recognized as a semantic neighbor of $P_X$.

![Figure 2: A portion of the service semantic overlay.](image)

3. NETWORK EVOLUTION AND SERVICE SEMANTIC OVERLAY ALIGNMENT

The P2P network evolves through time with peers joining or leaving. Evolution occurs both at logical and semantic overlay.

Logical network overlay evolution. The logical network overlay collects all the peer participating to the collaborative network: each node represents a peer and each link is a logical connection between two peers. In order to guarantee connection between peers an Overlay Management Protocol (OMP) is used, which defines specific procedures to join, leave and modify the logical network overlay. In our approach a shuffling-based OMP is chosen in order to allow more effective information diffusion among peers. As shown in the evolution statechart in Figure 3, when peer $P_X$ joins the collaborative network at the logical level for the first time (State 1-2), it receives a list of logical neighbors. Through the probe mechanism, $P_X$ sends a probe service request to each logical neighbor $P_Y$. If $P_X$ does not receive any reply, no inter-peer semantic links are established and $P_X$ remains connected only at the logical overlay (State 2). Otherwise, $P_X$ establishes inter-peer semantic links with its semantic neighbors, to which it is connected at the service semantic overlay (state 3). Probe service requests are sent according to a Time-To-Live (TTL) mechanism with a low TTL value to avoid network overload.

Service semantic overlay evolution. The service semantic overlay has to be managed with respect to events that occur to the collaborative peers: acquisition of a new peer in the logical network overlay, disconnection of a semantic neighbor from the network, publication/cancellation of a service in a registry.

The shuffling mechanism acts at the logical network overlay, but its effects are propagated at the service semantic overlay. If a peer $P_X$, connected at the service semantic overlay (state 3), acquires new logical neighbors that are not yet among its semantic neighbors, it activates the probe mechanism to check if the new logical neighbors are also semantic neighbors. Depending on the replies to probe requests, new inter-peer semantic links are established for alignment.

A peer $P_X$ recognizes that a semantic neighbor $P_Y$ becomes unavailable if a given number of messages sent to $P_Y$ are not answered. In this case the inter-peer semantic links toward semantic service interfaces published on $P_Y$ are removed from $P_X$. Publication/cancellation of a service are managed in a similarly way.

4. SERVICE DISCOVERY AND SERVICE SEMANTIC OVERLAY EXPLOITATION

A service request $S_R$ is formulated by specifying one or more service categories, the desired outputs and operations.
and the inputs that the requester is able to provide for service execution. It is sent to a peer \( P_X \) in the collaborative network and it is matched against semantic service interfaces published on \( P_X \) (local search).

After performing local search, \( P_X \) starts the distributed search forwarding the request \( S_R \) to its semantic neighbors with respect to \( S_{3_X} \) and \( S_{1_X} \), that is, semantic service interfaces that match locally. If no matching service has been found locally, \( P_X \) selects randomly a set of peers connected at the logical network overlay and forwards \( S_R \) to them. Distributed search can be performed by applying different forwarding policies based on inter-peer semantic links; the request is propagated following inter-peer semantic links until the search stops according to a Time To Live mechanism. In Figure 4 a minimal policy is depicted. In this case, a request \( S_R \) sent to peer \( P_X \) is partially satisfied by \( S_{3_X} \) and \( P_X \) forwards \( S_R \) to semantic neighbors (\( P_Y \) and \( P_Z \)) with respect to \( S_{3_X} \). If no semantic neighbors exist for any matching service \( S_{3_X} \) on peer \( P_X \), the request is forwarded to one of the peers that are connected to \( P_X \) in the logical network overlay (randomly chosen). In this example, \( S_R \) is sent to \( P_Y \) and \( P_Z \). On peer \( P_Z \) the service \( S_{1_Z} \) completely satisfies the request (by applying the matching model) and there is no need to further propagate \( S_R \). On peer \( P_Y \) the request is only partially satisfied by \( S_{2_Y} \), but semantic neighbors of \( P_Y \) (in this case, \( P_H \)) do not add further capabilities with respect to \( S_{2_Y} \) and also in this case the request is not further propagated.

![Figure 4: Distributed search.](image)

According to the minimal policy, search over the semantic overlay stops when matching services which fully satisfy the request have been found. Exhaustive policies can be applied following the same rules, but the search does not stop when matching services that fully satisfy the request are found: the request \( S_R \) is forwarded to semantic neighbors to find other services that could present, for example, better non functional features. In [5] a detailed presentation of different forwarding rules based on inter-peer semantic links is provided. Note that without the semantic overlay, the discovery process would rely on conventional P2P infrastructures and associated routing protocols for query propagation in the network (e.g. flooding). Exploiting the semantic overlay, it is possible to enforce query forwarding according to content similarities rather than to the mere network topology.

5. CONCLUSION

In this paper, we proposed the P2P-based Semantic Driven Service Discovery (P2P-SDSD) framework to enable cooperation and communication based on a semantic overlay that organizes semantically the P2P-integrated knowledge space and emerges from local interactions between peers. The semantic overlay can be seen as a continuously evolving conceptual map across collaborative peers that provide similar services and constitute synergic service centres in a given domain. The semantic overlay enables effective similarity-based service search and optimization strategies are defined for request propagation over the P2P network keeping low the generated network overload. For a preliminary evaluation of our approach, experiments are being performed [5] to demonstrate the better recall results of our distributed search with respect to other approaches, both using flooding-based strategies and applying a service semantic overlay, and to confirm that the use of the P2P-SDSD request forwarding policy results in an improved scalability.

6. REFERENCES