Abstract: Information and communication infrastructures underwent a rapid and extreme decentralization process from a world of statically and partially connected central servers to web of millions of information sources loosely connecting one to another (peers). The need of collaboration among peers requires dynamic capabilities of negotiating agreements on common interpretations within the context of a given task. What is (relatively) new is the focus on semantics because now information has to be sharable in an open environment, where interacting agents do not necessarily share a common understanding of the world at hand, as used to be the case in traditional enterprise information systems. For this reason, many researches were done on the approaches and techniques of semantic routing. This paper represents an effort to state the different techniques used with the limits, advantages, disadvantages and future work of each.

Key words: semantic routing, peer to peer (p2p), ontology, query routing

INTRODUCTION

The research area of Peer-to-Peer (P2P) systems is fairly new in the field of distributed computing, the idea of Peer-to-Peer (P2P) computing offers new opportunities for building highly open distributed systems. The emergence of collaboration among peers requires dynamic capabilities of negotiating agreements on common interpretations within the context of a given task [ABE 04]. This is typical for instance of peer-based systems, characterized by a set of independent peer parties without prior reciprocal knowledge and no degree of relationship, that dynamically need to cooperate by sharing their resources (e.g., data, documents, services). These types of systems are decentralized since authorities are defined to manage a comprehensive view of the resources shared by all the nodes in the system, due to the high dynamism and variability of collaboration and sharing requirements (Figure 1). On the opposite, each peer is responsible of providing the knowledge description of the resources to be shared, usually through ontology. Furthermore, peers act as independent agents and interact by semantically matching their respective knowledge (i.e., peer ontology) with the aim to discover potential collaboration partners with similar contents. As a result, each peer becomes aware of its semantic neighbors (i.e., nodes storing similar contents) that “emerge” from the knowledge discovery process.

By definition, Semantic Routing is a method of routing which is more focused on the nature of the query to be routed than the network topology. Essentially semantic routing improves on traditional routing by prioritizing nodes which have been previously good at providing information about the types of content referred to by the query. In order to provide scalable infrastructures for peer communications, P2P semantic routing protocols are being proposed based on the idea to address query propagation by selecting as query recipients the peers that are most likely to provide relevant results according to the query content. In this respect, some P2P semantic routing protocols are being proposed in the literature with the aim to address query propagation on the basis of the local context of the requesting peer [HAA 04] [BOR 05a] [ZEI 05]. However, most of the existing approaches are based
on the simplifying assumption to adopt a global repository of knowledge where mappings among the distributed peer knowledge are maintained. Actually, a challenging issue regards the need of advancing the existing semantic routing protocols by combining ontology-based peer knowledge models and ontology matching techniques for providing query forwarding on a real semantic basis and in a completely decentralized way. Furthermore, additional issues are related to the scalability question.

The main limits that affect traditional P2P routing protocols derive from two opposite goals that need to be pursued. On one side, queries need to reach the largest number of peers in order to increase the chances to locate the target file. On the other side, the need of scalability requires that the generated number of messages should be as lower as possible.

When the P2P network is targeted to knowledge sharing, query routing become even more complex than traditional file-sharing systems due to the different levels of semantic complexity that need to be considered in query processing [EHR 04a]. In this respect, content-based and semantic routing techniques are being proposed to optimize the tradeoff between the need to maximize the effectiveness of the discovery phase and to minimize the network traffic in a P2P system.

1. P2P semantic routing techniques

Semantic query routing techniques are required to improve effectiveness and scalability of current discovery and search processes for resource sharing in P2P systems. In this direction, the notion of P2P Semantic Link Network is introduced in [ZHU 04a] to emphasize the need of typed semantic links specifying semantic relationships between peers in order to maintain information about nodes with similar contents. Each peer defines its own XML Schema (source schema) describing the contents to share and adopts SOAP-based messages to communicate with the other members of the network.

1.1. Query routing through good peers

To provide better service in P2P systems, several research projects investigated how the traffic generated due to the broadcast protocol in a Gnutella system could be minimized while still preserving the quality of search results. One such approach is based on the idea of learning who the “good peers” are in a P2P system [RAM 02]. Peers that have sent a “good” response to a peer’s request are entered in a special list by the peer, following the assumption that these peers may also have good resources for subsequent queries in this area. Peers are deemed as “good peers” of peer C when these peers have responded the most to queries of peer C, regardless of the semantics of the query. The peer sends subsequent requests to the peers in its list.

We note that the proposed approach is suited for file-sharing P2P networks and only supports string-based matching functionalities when comparing the relevance of a file with respect to a given target query. Moreover, no check is performed by the requesting node on the real relevance of the collected replies, and then importance measures can indicate misleading interest similarity values. The reputation evaluation model provides a subjective and poor grained interpretation of the file relevance that is not related to the semantic affinity existing between files and queries. However, experimental data shows that structuring the network according to the peer interest similarity can be promising. In particular, the approach can provide interested results in terms of traffic load reduction as a high number of files matching the query target can be discovered also with low TTL values.

The authors found that this reduces traffic in the network as peers now send queries selectively to other peers. Also this selective dispatching of queries does not affect the search results at the end and good responses are obtained in a short interval. This works when the user is consistently searching in a single type of semantic query. However, users search often for more than one semantic category at a time [IAM 02]. So when the user changes its query semantics, then
these “good peers” will not be able to provide the responses. The peer will find new “good peers” for this type of query again, but the peers of the old query type will be dropped.

1.2. The REMINDIN’ query routing algorithm

REMINDIN’ (Routing Enabled by Memorizing INformation about Distributed INformation) algorithm is proposed to provide a query routing mechanism based on confidence ratings [STA 06] for the SWAP (Semantic Web and Peer-to-peer) platform [BRO 03a]. In SWAP, each peer provides a local node repository (i.e., an ontology) containing the RDF(S) statements that describe either data or conceptual information of the peer [BRI 03]. Furthermore, the SWAP storage model also capture meta-information about the RDF(S) statements in order to specify where the statement came from and how much resource-specific confidence and overall confidence is put into these statements and peers, respectively. The SWAP query model is based on the SeRQL language where a query consists of the RDF(S) triple (i.e., subject, object, predicate) [BRO 03b].

We note that relying on peer replies to train the routing protocol behavior enables a passive learning approach where peers become aware of relevant knowledge location without affecting the network with an additional traffic overhead. Moreover, the idea that confidence and expertise measures are directly associated to resources allows to specify information about peer relevance at different levels of granularity. Furthermore, additional metrics (e.g., trust and reputation of nodes) can be considered for flexibly providing a more accurate measure of peer relevance. On the other hand, we observe that semantic resource matching and efficient query processing are crucial issues for providing effective knowledge sharing.

1.3. Socialized.Net and the Seers search protocol

The Seers search protocol is defined to enforce P2P semantic query routing by creating an overlay network where nodes are organized according to their interests [BOR 04]. In the Seers search infrastructure, each shared resource is described through a meta-document that provides a set of corresponding metadata expressed in the XML syntax. When a query is submitted to the system, a target meta-document is provided by specifying the searched metadata and one or more matching meta-documents are returned as a result by the replying peers. In particular, a reply is represented by a meta-document describing an existing resource whose metadata match the target meta-document. In order to limit a sudden burst of replies to the requesting peer, all matching meta-documents are routed back, following the reverse path of the query. After the discovery of the matching meta-documents available in the network, the requesting peer can acquire the corresponding resources through a traditional HTTP connection. The Seers search protocol is based on three reference policies: the matching policy, the transmission policy, and the life cycle policy.

The Seers search protocol has been refined in [BOR 05b] where the Socialized. Net infrastructure is presented to support a mechanism for query forwarding based on preferences and reputation. A social metaphor based on human gossiping is considered as a reference model that is followed by each peer of the network to identify the reliable nodes that can be selected for sharing interaction and to improve semantic routing efficiency. In the Socialized.Net infrastructure, when peers are selected for query forwarding, neighboring nodes are ranked by combining the traditional Seers measures (i.e., state, interests) with preference and reputation information. In particular, preference is a local rating of neighbor nodes based on statistics and user input, while reputation is an asynchronous gossip between neighbor nodes regarding their behavior.

We note that the use of a social metaphor based on human gossiping for defining the Socialized. Net infrastructure enforces an adaptive query routing protocol. In particular, the proposed Seers search mechanism is based on a soft learning mode where logical connections are established according to the observed peer behavior in query replies and where active advertising methods are limited to reputation information exchange. In this respect, we observe that the use of soft or completely passive learning modes should be preferred for improving semantic routing efficiency as they contribute to reduce the traffic overhead introduced by peers for measuring remote peer expertise. However, the main concern with Socialized.Net is that this system only supports metadata-based peer content representation and thus traditional string-based matching functionalities that are suited to work in file-sharing P2P systems.

1.4. Directed Breadth First Search

Directed Breadth First Search (DBFS) selectively forwards queries to peers who are seen to have returned good results for previous queries [YAN 02]. For this kind of intelligent routing, the peer has to keep information on the neighboring peers about how many hops that peer had to forward the query to obtain results, the number of results returned by that peer, and the length of the message queue of that peer. All this information has to be stored and processed by a peer when it wants to request a file. Since the neighboring peers do not change, the query message travels the entire path until it finds the peer that has the document. This approach still generates a lot of messages. The authors do not explain how a peer will know what constitutes a good result. If the number of results is taken to be an indication of good results, then a peer that returns irrelevant but many results would be considered a good candidate.

1.5. BuddyWeb

In the “Buddy Web” approach [WAN 04a] [WAN 04b], routing is done based on similarity of interest. The main reason for building a “BuddyWeb” was to
redistribute the amount of traffic in and out of the university network, which costs the university money. For a given query, there is always the possibility that someone else on campus has already done the same query and has already received results for it. By caching previous search results and searching first through them, the university can reduce cost and utilize the network better. The similarity between the user’s interests and the files’ semantics that is used for routing inside the network is calculated in several ways. One way is based on the meta-data of the file, e.g. the title tags downloaded by the user. Another way is by storing words that the user selected while browsing through its contents in a vector. By summing up points for the words, which have already been assigned some weight by some weighting scheme, a calculation expressing the value of interest is performed. A few servers maintain the interest value along with the peer’s address. A peer calculates similarity with other peers in the network by comparing its own interest value to those of the others. When a peer is querying, the query is first sent to the underlying network, called BestPeer. The system sends the query to those peers whose interest values are close to the querying peer. Once results are found, they are sent to the querying peer. If the system does not find any results then it sends the request outside the network.

An external search takes place in the traditional manner. A peer does not learn if the peers who have been found to be similar are useful or not. The authors do not mention how long a cache is stored. If a peer is searching vigorously, it may overwrite the cache with new documents and older documents will not be found.

1.6. P2P semantic link networks

A P2P Semantic Link Network (P2PSLN) is a direct network, where nodes are peers or P2PSLN, and edges are typed semantic links specifying semantic relationships between peers [ZHU 04]. Different types of semantic links can be established between two nodes (i.e., equal-to, similar-to, reference, implication, subtype, sequential, empty, and null) according to a set of pre-defined reasoning rules. Each peer defines its own XML Schema (source schema) describing the contents to share and adopts SOAP-based messages to communicate with the other members of the network. When a peer joins a P2PSLN, it will first identify the semantic relationship between itself and the nodes in the network. Moreover, the entering peer acquires the XML Schema of the identified nodes in the network (target schemas) and analyzes them recursively in depth-first order with the intention to identify the semantic mappings (i.e., semantic node mapping, semantic clique mapping, and semantic path mapping) between the elements in its source schema and the elements in the incoming target schemas. The similarity between the source and the target schemas can be measured by the methods of cycle analysis and functional dependency analysis as proposed in [ABE 03a].

Upon receiving a query, a peer will autonomously forward the requirement to relevant peers according to the types of the semantic links as well as the similarity between elements and structures of peer schemas. QoP (Quality of Peer) techniques, such as number of returned results, response time, traffic overhead, precision, and recall, are considered to manage inconsistent data in returned data flows. The P2PSLN approach is compared with the Breadth First Search (BFS) and the RandomWalk Search (RWS) by means of simulation techniques. The experiment illustrates that the BFS routing policy achieves the highest recall rate, but P2PSLN can ensure the best performance in terms of generated traffic.

We note that the adoption of reasoning-based semantic mappings among nodes constitute a first step towards an actual query semantic routing protocol. However, the main problem of P2PSLN is that the semantic overlay of semantic mappings is statically defined with the gradual entrance of peers in the system. In particular, an entering node defines semantic mappings only with the initially discovered peers. Such an approach has two main drawbacks. First, the network topology does not change according to the peer contents. This way, the number and the relevance of the semantic mappings can be poor in case those weakly similar peers are neighbors, thus affecting the query routing capabilities of the entire network. Furthermore, a P2PSLN lacks of flexibility as semantic mappings need to be recomputed when changes occur in peer source schemas. In this respect, moving P2PSLN towards an adaptive approach for peer organization could contribute to increase flexibility by allowing nodes to get closer according to their content similarity.

1.7. The intelligent search mechanism

The Intelligent search mechanism (ISM) is defined in [ZEI 05] as a novel mechanism for information retrieval in P2P networks with the aim to improve traditional techniques by efficiently finding the most relevant answers to a given query rather than the largest number of answers. The underlying idea of ISM is that a peer that has a document relevant to a given query is also likely to have other documents that are relevant to other similar queries. In this respect, for each query in an ISM-based system, a peer exploits the replies of past queries and estimates which of the known peers are more likely to return relevant results to the actual query, and propagates the request to those peers only. ISM techniques are entirely distributed and each node can make local and autonomous decisions without coordinating with any other peers and with no centralized authority. In an ISM-based system, each node maintains a collection of documents (text, audio, and video) that it is willing to share with the other members of the system. Moreover, for each document in a node, a set of metadata (e.g., author, title) is maintained along with a number of associated keywords. Searches occur when a node submits to the ISM-based system a query containing one or more target keywords of interest. Each receiving node
matches the target keywords in the query with the keywords associated to its own documents and replies by returning those documents that are associated to at least one matching keyword. In order to match queries and document keywords, string-based query similarity metrics can be used, such as the cosine similarity function [BAE 99].

We note that ISM has been specifically conceived for P2P information retrieval and thus it is suited for document- and file-sharing rather than for knowledge-sharing. In particular, a lack of expressive power characterizes the ISM approach due to the keyword-based mechanism adopted for metadata representation and to the string-based matching techniques. However, the idea to use past interactions to rank known peers and to train the behavior of the routing protocol can have an interesting impact also on more expressive knowledge-sharing infrastructures.

1.8. Hierarchical semantic routing for Grid resource discovery

A hierarchical semantic routing algorithm for grid networks is proposed with the aim at exploiting query content and peer knowledge to drive routing decisions [JLI 05]. In particular, the nodes of the network are grouped into independent clusters that are organized in tree-based structures. In such an approach, each node encodes its available resources and provides the corresponding RDF metadata representation. The adoption of RDF as resource representation language implies that the node metadata repositories can be queried by relying on existing RDF query languages (e.g., SPARQL [PRU 06], RDQL [SEA 04]). For semantic routing purposes, nodes periodically exchange their respective RDF metadata and, thus, a concise representation is required. To this end, each node summarizes the verbose RDF syntax used for resource representation in a space-efficient bitmap according to the hashing-based Bloom Filter technique [BLO 70]. It is important to stress that the Bloom filter technique introduces a certain degree of approximation in the node resource representation.

The Bloom Filter summarizations are exploited to aggregate nodes in tree-base clusters by relying on the Prinkey method [PRI 01]. In such a method, a node joining the system is inserted in a tree structure and provides to its parent a summarization (i.e., bitmap) of its RDF metadata. Each non-leaf node n maintains a routing table containing a number of Bloom filters: one bitmap for its local resources and one bitmap for each direct child. The node n merges through logical OR its own bitmap with the child bitmaps and produces a new comprehensive bitmap as a result. The merged bitmap is sent to the parent of the node n. This way, each non-leaf node n in the tree maintains a comprehensive bitmap describing the RDF metadata of all the nodes in the branch rooted by n. Subsequently, the cluster root provides a summarized view of all the knowledge stored in the cluster.

1.9. Query routing through semantic mappings

The routing by mapping mechanism is proposed in Mandreoli [MAN 06] to address semantic query propagation in a Peer Data Management System (PDMS) [HAL 04]. The approach is based on the idea to enhance the sharing functionalities of a peer-based system by combining the semantic richness of PDMSs with the routing capabilities of pure P2P architectures. To this end, each peer is defined as an autonomous and independent agent and provides an OWL ontology describing the informative contents of its sources. Two adjacent peers (i.e., neighbors) are logically connected through a set of semantic mappings that keeps track of the semantic affinity among the ontology concepts of the two peers. As a result, the system is organized as an overlay network where two nodes are connected through a set of semantic correspondences between concepts. Furthermore, semantically rich query models are adopted as in traditional PDMS. In particular, a query is specified by a requesting peer p in terms of its relation terminology and it is submitted to the other peers that can evaluate whether they can reply with relevant knowledge. A receiving peer q evaluates the query by rewriting it according to the semantic mappings between p and q ontologies and by executing the query on its ontology. When a peer joins the network, it has to discover the semantic mappings between its ontology and the respective ontologies of the neighbor peers by performing appropriate schema matching operations. In particular, the matching operations are based on terminological and structural comparisons on the elements of the two ontologies and produce a set of semantic mappings between concepts as a result. Thus, each concept in the ontology of the joining peer is associated to the most similar concepts in the neighbor schemas. Furthermore, each mapping is characterized by a numerical score in the range [0, 1] that quantifies the level of semantic similarity of the two concepts connected by the mapping. The mapping score is computed according to a fuzzy-based similarity function and represents the level of affinity of the two concepts.

1.10. Semantic Query Routing in SenPeer

SenPeer is an unstructured P2P system based on an organization of peers around super-peers according to their semantic domains. In SenPeer, the knowledge of a peer is represented through a semantic network (called sGraph) which can represent data conforming to various data models. To interact, super-peers maintain expertise tables describing data at the semantically linked (super-)peers and define semantic mappings between their content descriptions. These mappings are the basis of semantic overlay network where peers having similar schemas form a semantic neighborhood. This semantic overlay is exploited further to address query propagation. The routing technique is applied in the presence of several data models provided that there are wrappers which undertake the transformation of the query to a suitable query language for each data source.
The peers advertise the content they want to share by their expertise and discover acquainted peers by schema matching techniques. The peers’ expertises and the semantic mappings are the basis of a semantic overlay organizing the network into semantic domains. The introduction of superpeers in combination with expertise tables and domain-aware query routing reduces the workload of peers by distributing queries to the appropriate subset of peers.

This semantic routing mechanism provides better results in terms of recall and precision, reduces query distribution effort significantly compared to the baseline.

1.11. H-Link semantic routing

H-Link semantic routing approach is designed to exploit the results of an ontology matchmaking process for providing a semantic overlay network where peers having similar contexts are recognized and interlinked as semantic neighbors [MON 08]. In particular, H-LINK aims at advancing the existing semantic routing protocols by combining ontology-based peer context descriptions and ontology matching techniques for providing query forwarding on a real semantic basis, in a completely decentralized way. Furthermore, H-LINK aims at enforcing scalability in query forwarding by introducing a credit-based mechanism where the approximate number of desired replies is specified rather than the non-scalable number of hops to cross. H-LINK relies on the results of the knowledge discovery interactions among peers to train the query routing behavior and to identify the peers with similar knowledge that are called semantic neighbors. Query recipients in H-LINK are then selected according to confidence and semantic neighborliness values associated with the discovered semantic neighbors. This way, H-LINK aims at supporting query propagation by exploiting semantics-based criteria, like semantic affinity, rather than topology-based parameters, like peer distance and bandwidth. The results obtained in the experiments show that the H-LINK mechanism succeeds in improving the effectiveness of traditional P2P query protocols by providing interesting results in terms of scalability at the same time.

1.12. Other Techniques

The idea to combine the self-organization of unstructured P2P systems with the efficiency of DHT like infrastructures represents the basic motivation of P-Grid [ABE 03b], a peer-to-peer lookup system based on a virtual distributed search tree. As in traditional DHT systems, the virtual tree covers the search space and each peer holds part of the overall space (i.e., a tree branch) through a hash-based index. A bitstring mechanism is provided to summarize the overall information for the tree branch the peer is responsible for. The distinguishing feature of P-Grid is that the peer position within the tree is dynamically determined through negotiation with other peers. This way, the network topology autonomously reacts to peer changes for providing query adaptation, thus improving search efficiency. As for many other structured approaches, the main concern with P-Grid is related to the traffic overhead that is required for maintaining the tree structure and for dynamically determining the peer position.

In Sidourgos [SID 05], an original DHT-based framework is presented for providing efficient routing of expressive RDF/S queries. In particular, the approach is organized in a classical schema-based infrastructure where both queries and peer knowledge are described through RDF/S statements. As a novel contribution, an original encoding of RDF/S schema fragments is introduced for checking whether a peer schema is subsumed by a query. This way, the system succeeds in recognizing RDF/S schemas with similar semantics, and thus the associated peers are placed in the DHT topology, accordingly. Experimental results show that the main benefits of the approach are obtained in terms of scalability. On the opposite, flexibility represents the main drawback, since only RDF/S resource descriptions are supported.

Some research explores the interest-based behavior of peers. [SRI 03] Proposed a solution where peers loosely organize themselves into an interest based structure. Initially, the algorithm for querying is flooding or broadcast as in Gnutella. When replies are returned, the peer chooses a peer randomly from all the peers who replied and adds it to the “shortcut list”. Subsequent queries are sent to the peers in the “shortcut list”. If there are no responses obtained by using this technique then the peer uses the flooding algorithm again. Each peer assigns space for storage of this list. The results show that the “shortcut list” is effective in finding both popular and unpopular content (files), and that 45-90% of files is found quickly. But a drawback to this approach is that it will create a lot of traffic and time delay as the search may keep reverting back to the flooding algorithm when the document is not found through that single peer chosen from the subset of peers responding to a similar query. The shortcut list contains peers who have returned response regardless of the semantic category of queries. A modification to the model that the authors suggest is to add to the list all the peers who have replied to a query. But the benefit of using a short-cut list may be lost as peers just keep getting added and deleted when the peer changes its interest for searching since there is only a limited space in the list.

Another approach [KAL 02] proposes that the peer maintains a profile of all other peers who answered its requests in a local repository by keeping a list consisting of (query, peer) pairs. A query is defined by the set of keywords used for the search. All peers that respond to a query are added, together with the query, to the pairlist. When the peer initiates a new query q, it matches it with the previous ones stored in its pair-list and picks out the peers that have answered similar queries. This is achieved in the following way: In this model the neighboring peers connected to the peer remain the same from the start and only the profile is
updated. Each query in the list is assigned a similarity value relative to the present query \( q \). It then adds up all the similarity values for a peer and forwards the requests to only the subset of peers from its neighborhood, that are found to have the greatest value for similarity with respect to the present query. A random peer, a peer not found to have answered similar queries, is also chosen for forwarding to explore the network. The authors conclude that the traffic is reduced. The amount of storage is limited and a node uses a Least Recently Used (LRU) policy to decide which (query, peer) will remain in the local repository. Thus if a peer is searching consistently in a query type then, older pairs of a different query type will be deleted from the repository. If now the peer wants to search in a different query type, the peer has lost the best peers for that query.

Some authors [CRE 02] suggest creating groupings of peers, which advertise together the resources available to each member of the group. The idea is similar to creating super-peers like in KaZaA, but the index of files is not maintained by a superpeer but by each member of the group. Peers use indices to keep track of the number of documents that can be accessed through each of their connected peers in a semantic category and communicate to neighbors the total number of documents that they can access in a category. Thus each peer has information about which neighboring peer to choose when a query arrives, based on the total number of resources accessible through the peer.

The approaches described above explore the semantics of queries and implicitly model the interests of peers in different semantic areas. These approaches all selectively forward queries to peers based on the knowledge obtained and stored. Some research studies have used semantic routing of queries by classifying queries based on keywords.

NeuroGrid [JOS 02] uses keywords, derived from the file meta-data, for searching and then matching a query. Peers are associated with keywords based on the contents they store. When a peer issues a query, it looks for remote peers characterized by keywords similar to the query and forwards the request to those peers. The peer also changes its neighborhood by adding remote peers that have answered its queries. Thus peers learn about other peers, along with all the queries they have answered, and route queries according to this knowledge. However, storing all this information is expensive since the storage is bounded and there is a limit to how much information each node can maintain.

In another approach [TRI 02], documents are accompanied by keywords, which semantically represent the contents of the documents, and these keywords are classified according to semantic categories. The system is assumed to be logically organized so that clusters of peers are formed based on semantic categories. However, the researchers do not specify how this is achieved. Each category could have one or more clusters associated with it. Peers have knowledge about all the categories and the entire associated set or subset of peers in the system. When a search request comes in, a peer forwards queries to the cluster associated with the semantic category of the search keyword. If the document is not found, then the peer in the cluster sends the queries to other peers in the same cluster or another cluster associated with the same category. In this model each peer in the system has to know all the categories existing in the system and clusters associated with them, which is hard to achieve without a central server maintaining a list of categories and clusters and peers.

Semantic routing has also been used by [NGC 02] in a P2P system for sharing images. The authors propose a search strategy based on clusters of peers with similar properties. Each peer according to its interests calculates its signature value using a certain function. When a peer joins a network it learns about a few peers, called “random links”, through the Gnutella ping-pong messages and adds them to its list representing a peer-cluster. Queries by the peer are forwarded through these “random links”. A peer stores another peer in its peer-cluster list if the data they share are similar; such peers are called “attractive links”. Peers thus get clustered according to similar signature value. Thus, when a query is sent by a peer it gets checked against its signature value and if no close match is found, the query is forwarded randomly to peers. Once the query reaches the appropriate peer in a cluster, the query is broadcast using the peer’s “attractive link” lists inside the cluster.

2. P2P semantic routing Categories

Content-based and semantic routing techniques are being proposed to improve the tradeoff between the need to maximize the effectiveness of the discovery phase while minimizing the network traffic of a P2P system. In the following, we analyze and we compare the main existing approaches for P2P semantic routing. In particular, we have considered the approaches where the idea of peer context is supported in some way to describe the knowledge provided by a single and independent node. In this respect, the considered approaches are classified in two main categories according to the adopted peer-knowledge representation approach, namely metadata-based and ontology-based semantic routing techniques.

2.1. Metadata-based semantic routing techniques

Metadata-based semantic routing approaches support an implicit definition of peer context since each node participates to the system by providing a set of files to share together with a set of corresponding metadata (e.g. title, author) and keywords. In this respect, the context of a peer \( p \) is defined by the set of keywords associated to the shared files provided by \( p \). In such kind of P2P systems, syntactic matching techniques (e.g. string-based and keyword-based) are defined to evaluate the similarity of a peer file with a given target request. Existing metadata-based semantic routing approaches can be distinguished
according to the level of adaptivity characterizing the query forwarding mechanism. As an interesting example of metadata-based semantic routing approach, the Intelligent Search Mechanism (ISM) is presented in [ZEI 05] as a novel mechanism for information retrieval in P2P networks. For each query, a peer exploits the replies of past queries and estimates which of the known peers are more likely to return relevant results for the current query, and propagates the request to those peers only. Searches occur when a peer submits to the system a query containing one or more target keywords of interest. Each receiving node matches the target keywords in the query with the keywords associated to its own documents and replies by returning those documents that are associated to at least one matching keyword. In order to match queries and document keywords, traditional similarity metrics can be used, such as the cosine similarity function [BAE 99] that is proposed by the authors. Each peer maintains a profiling structure where it stores the answer observations related to the queries that were routed through it.

Profiling structures represent an element of locality to be exploited for selection of query recipients. A forwarding mechanism based on TTL is defined to enlarge the query scope when required. Similarly to ISM, the concept of ‘good’ peer is defined in [RAM 04] as the key element of an automatic interest-based mechanism where peers with similar interests are close and the similarity between peer interests decreases as the distance among the peers increases. To measure the interest similarity, each peer calculates the ‘percQueryHits’ value which measures the ratio of the hits provided by each of its immediate neighbors (i.e. directly connected peers) to the total number of collected query replies. The percQueryHits values are then exploited by a requesting peer to select the most responsive neighbors as query recipients. Syntactic matching techniques, such as string-based techniques, are then used to determine the query reply by comparing the target keywords of a query with the document metadata in the repository of each receiving peer. As another example, the Seers search protocol is defined in [BOR 04]. In the Seers infrastructure, a XML meta-document is used to describe each shared resource and a matching policy is provided to define how tags in different meta-documents should be matched against others. Furthermore, a transmission policy is used to describe how a query with a target meta-document should be propagated in the network.

In particular, each peer associates a ranking value to its neighbors. For each neighbor, the ranking value is computed on the basis of (1) the neighbor state (i.e. active, inactive, stale), and (2) the neighbor interests that are discovered by observing the meta-documents routed by the neighbor.

The Seers search protocol has been refined in [BOR 05a] where the Socialized.Net infrastructure is presented to support a mechanism for query forwarding based on preferences and reputation for peer ranking. Finally, the Neurogrid system is proposed in [JOS 02] as an adaptive and decentralized search system. In such an approach, semantic routing is intended as content-based query forwarding, and a learning mechanism is defined to dynamically adjust the relevance of known peers for each query. In particular, each peer maintains (1) a list of keywords associated to its shared files and (2) a frequency indicator associating a keyword with the number of hits-per-remote-peer collected in previous interactions. The frequency indicator is used to rank the remote peers for query recipient selection.

2.2. Ontology-based semantic routing techniques

Ontology-based semantic routing approaches support an explicit definition of peer context since each node participates to the system by providing an ontology describing its shared resources. In this respect, the ontology of a peer p defines the context of p and semantic-based matching techniques are defined to evaluate the similarity of a peer context with respect to a given target request. Existing ontology-based semantic routing approaches can be distinguished according to the supported ontology specification language and to the level of semantic complexity exploited by the adopted matching techniques. As an interesting example of ontology-based semantic routing approach, the REMINDIN’ (Routing Enabled by Memorizing INformation about Distributed INformation) algorithm is proposed in [EHR 04b]. Such an approach is defined in the framework of the SWAP platform where (1) each peer provides a local node repository (i.e. an ontology) containing the RDF(S) statements that describe either data or conceptual information, and (2) the SeRQL language is adopted for query formulation [11]. Furthermore, the SWAP storage model allows a peer to measure the confidence value of another remote peer p in terms of the number of relevant answers provided by p for each argument r. Such a confidence value is updated when new interactions are issued with the peer p on topics related to r. Receiving a query, a peer replies to a requesting node with the RDF(S) statements that satisfy the SeRQL query and that are contained in its local node repository. After the reply, a peer can forward the query to other peers in the network according to a TTL-based mechanism. In this respect, REMINDIN’ is invoked to select the most promising query recipients by ranking remote peers according to their stored confidence values. The REMINDIN’ approach has been refined where advanced techniques for peer confidence computation are illustrated. As another example, P2P Semantic Link Networks (P2PSLN) is presented in [ZHU 04b]. In a P2PSLN, different types of semantic links can be established between two peers (i.e. equal to, similar to, reference, implication, subtype, sequential, empty, and null) according to a set of predefined reasoning rules. Each peer defines its own XML Schema (source schema) describing the contents to share and adopts SOAP-based messages to communicate with the other members of the network. When a peer joins a P2PSLN, it acquires the XML Schema of a subset of network nodes (target schema) and analyzes them.
recursively through cycle analysis and functional dependency analysis with the intention to identify the semantic mappings between the source and the target elements. Upon receiving a query, a peer will autonomously forward the request to relevant peers according to its semantic links as well as the similarity between elements and structures of peer schemas. Similar to P2PSLN, the ‘routing by mapping’ mechanism is proposed in [MAN 06] to address semantic query propagation in a Peer Data Management System (PDMS) [HAL 04].

To this end, each peer provides an OWL ontology describing the informative contents of its sources and two adjacent peers (i.e. neighbors) are logically connected through a set of semantic mappings that keeps track of the semantic affinity among the ontology concepts of the two peers. In order to discover the semantic mappings between two peers, schema matching operations based on terminological and structural comparisons are performed on the respective peer ontologies.

Each resulting mapping is finally characterized by a numerical score in the range [0, 1] computed through a fuzzy-based similarity function. To support query submission/forwarding, a Semantic Routing Index (SRI) is maintained by each peer as a matrix where columns are the concepts of the local ontology and rows contain the neighbor peers as well as a summarized indication of their capability to provide relevant results for each local concept. In HAA 04], the authors define a peer selection approach based on a shared ontology and an expertise advertisement mechanism. In such an approach, each peer picks out the recipients of a given query by selecting the known peers whose expertise is similar to the query subject. The system is characterized by a common ontology that provides a shared conceptualization for the domain of interest of all the system nodes. Each peer defines and subsequently advertises a semantic description of its expertise that represents the knowledge base of the peer according to the common ontology. As a result of advertisement, each peer acquires a list of nodes with similar expertise that is exploited to rank query recipients according to the semantic similarity between peer expertise and the query subject.

3. Summary of Advantages and Disadvantages

I will list in this part the advantages and the disadvantages of the most common Semantic Routing Techniques.

<table>
<thead>
<tr>
<th>Semantic Routing Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Query routing through good peers</strong></td>
<td>This approach can provide interested results in terms of traffic load reduction as a high number of files matching the query target can be discovered also with low TTL values. The selective dispatching of queries does not affect the search results at the end and good responses are obtained in a short interval.</td>
<td>This works when the user is consistently searching in a single type of semantic query. No check is performed by the requesting node on the real relevance of the collected replies. The reputation evaluation model provides a subjective and poor grained interpretation of the file relevance that is not related to the semantic affinity existing between files and queries.</td>
</tr>
<tr>
<td><strong>The REMINDIN’ query routing algorithm</strong></td>
<td>We note that relying on peer replies to train the routing protocol behavior enforces a passive learning approach where peers become aware of relevant knowledge location without affecting the network with an additional traffic overhead. The idea that confidence and expertise measures are directly associated to resources allows to specify information about peer relevance at different levels of granularity. Additional metrics (e.g., trust and reputation of nodes) can be considered for flexibly providing a more accurate measure of peer relevance.</td>
<td>Semantic resource matching and efficient query processing are crucial issues for providing effective knowledge sharing. The use of a traditional TTL-based mechanism for query forwarding.</td>
</tr>
<tr>
<td><strong>Socialized.Net and the Seers search protocol</strong></td>
<td>The use of soft or completely passive learning modes should be preferred for improving semantic routing efficiency as they</td>
<td>This system only supports metadata-based peer content representation and thus traditional stringbased matching</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Advantages</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Directed Breadth First Search</td>
<td>Many results would be considered a good candidate.</td>
<td>This approach still generates a lot of messages.</td>
</tr>
<tr>
<td>BuddyWeb</td>
<td>Reduces the amount of traffic in and out and reduces cost</td>
<td>A peer does not learn if the peers who have been found to be similar are useful or not.</td>
</tr>
<tr>
<td>P2P semantic link networks</td>
<td>P2PSLN can ensure the best performance in terms of generated traffic.</td>
<td>The network topology does not change according to the peer contents.</td>
</tr>
<tr>
<td>The intelligent Search mechanism</td>
<td>ISM has been specifically conceived for P2P information retrieval and thus it is suited for document and file-sharing rather than for knowledge-sharing.</td>
<td>The lack of expressive power characterizes the ISM approach due to the keyword-based mechanism adopted for metadata representation and to the string-based matching techniques.</td>
</tr>
<tr>
<td>Hierarchical semantic routing for Grid resource discovery</td>
<td>This approach can contribute to reduce both the network and the peers load, thus providing a greater scalability.</td>
<td>Sparse overlapping that typically emerges in real systems.</td>
</tr>
<tr>
<td>Query routing through semantic mappings</td>
<td>This approach actually defines a semantic-based infrastructure for enforcing query routing through the P2P network. However, mapping definition is an expensive operation in terms of computation time especially when the peers ontologies are large.</td>
<td>A joining peer may be interested in tolerating such a computation effort only in case of an expected stable presence in the network. This condition can rule out very dynamic peers.</td>
</tr>
<tr>
<td>H-Link semantic routing</td>
<td>H-LINK mechanism succeeds in improving the effectiveness of traditional P2P query protocols by providing interesting results in terms of scalability at the same time.</td>
<td>Unexpected H-Link behaviors were observed during experiments.</td>
</tr>
<tr>
<td>Semantic Query Routing in SenPeer</td>
<td>This approach provides better results in term of recall and precision, reduces query distribution effort significantly compared to the baseline. The introduction of superpeers in combination with expertise tables and domainaware query routing reduces the workload of peers by distributing queries to the appropriate subset of peers.</td>
<td>In some cases, flooding is unavoidable.</td>
</tr>
</tbody>
</table>

4. Conclusion and Future Work

Much future work is provided to enhance the semantic routing. As for protocols, a challenging issue is by combining ontology-based peer knowledge models and ontology matching techniques for providing query forwarding on a real semantic basis and in a completely decentralized way. Furthermore, additional issues are related to the scalability question. As for techniques many recommendations was listed for approximately each technique trying to solve the disadvantages.
A possible REMINDIN’ improvement in terms of flexibility can be introduced by combining SeRQL and associated syntactic and structural matching techniques with other existing ontology matching approaches, such as those based on linguistic and contextual comparisons or based on reasoning. Also, related to the use of a traditional TTL-based mechanism for query forwarding in REMINDIN’, given a target SeRQL query, the rankPeers algorithm returns the top-k relevant peers. Based on the assumption that more relevant peers are mutually connected, different values of TTL should be assigned to query recipients according to their ranking. This way, high TTL values can be selectively assigned to most relevant peers, thus contributing to reduce the overall network traffic.

As for Seers search, and in order to address knowledge-sharing P2P interactions, the Socialized.Net framework should be improved to provide semantic-based matching techniques capable of dealing with more expressive knowledge representation models (e.g., ontologies).

Moreover, the element of locality introduced with the profiling structure in the Intelligent Search Mechanism (ISM) should be taken into account to improve the scalability of the routing mechanism.

In Hierarchical semantic routing for Grid resource discovery, details regarding the RDF peer repositories adopted in the experiments are not specified and a more extensive experimentation should be provided to assess whether the approach can effectively work in different scenarios, especially with sparse overlapping in RDF statement distribution.

As for the semantic overlay network, it should be adaptive and should be devised to “near” in terms of hops the peers with similar contents.

More work must be done on evaluating the H-Link behavior when a different model for peer knowledge distribution is exploited. Also, interesting experiments should be intended to assess the variation of H-Link performance in terms of scalability when the ZIPF distribution is adopted. This choice is due to the fact that the ZIPF distribution provides a non-uniform distribution and a more realistic model of peer-knowledge assignment where some concepts are more queried than others and certain concepts have a higher likelihood to be assigned to peer ontology during the distribution phase.

Finally, in Semantic Query Routing in SenPeer, future work includes query planning, query optimization and answer fusion.

REFERENCES


