ATLAAS-P2P: A Two-Layers Architecture for Approximated Search in P2P

Ranieri Baraglia
ISTI - CNR
ranieri.baraglia@isti.cnr.it

Patrizio Dazzi
ISTI - CNR
patrizio.dazzi@isti.cnr.it

Matteo Mordacchini
ISTI - CNR
matteo.mordacchini@iit.cnr.it

Laura Ricci
University of Pisa
ricci@di.unipi.it

ABSTRACT
ATLAAS-P2P is a two-layered P2P architecture for developing systems providing resource aggregation and approximated discovery in P2P networks. Such systems allow users to search the desired resources by specifying their requirements in a flexible and easy way. From the point of view of resource providers, this system makes available an effective solution supporting providers in being reached by resource requests.

1. SYSTEM DESCRIPTION
The process of discovery useful resources in a P2P network is highly conditioned to the query formulation mechanism [3]. Users should be enabled to easily express their needs and an efficient query resolution mechanism should be able both to efficiently find significant resources and to limit the number of messages exchanged. Common techniques for searching resources in P2P systems are based on range queries over a set of different attributes [4]. However, the amount of resources in a P2P network could be very large and heterogeneous, and users knowledge about the available resources could be not enough accurate to allow them to properly formulate their queries. Actually, a user is more likely able to define an “ideal” resource that satisfies her needs and ask to the search system to find resources close to such entity. Indeed, she would avoid the need to specify precise ranges on all the attributes. She would simply provide an example of what better suits her requirements.

This mechanism would simplify the work for users and would lead to a more efficient exploitation of the search system. Moreover, from the point of view of resource providers, which aim is to be found by users, it would provide an effective infrastructure for resource advertising.

ATLAAS-P2P consists in a P2P system that provides to the user a flexible way to express her requirements and an effective solution for letting resource providers be reached by users requests. It is based on a two-layer architecture, where peers in the network represent the resources of providers. The lower layer is an unstructured, gossip-based, P2P network [1, 2] allowing peers to efficiently gather in logical groups of nodes representing similar resources. The role of this layer is to automatically capture the affinities existing between resources belonging to different providers and put them in common communities. Those communities distribu-

Figure 1: Overall Architecture

The structured network has been extended to support approximated searches over the community representatives [5]. Users can submit to this network the queries by providing sample resources consisting in prototype of resources they are searching for. Most similar representatives are found and are selected for forwarding the query inside the community they represent in a gossip fashion. This permits to give to users suitable alternative to resources they are searching for when none of the resources available in the system matches the user request.

The overall architecture of ATLAAS-P2P is sketched in Figure 1. Peers (circles) form distinct communities built on a similarity basis in the unstructured gossip-based layer. Each community elects a representative, denoted with an L in the figure. Each representative is in charge of register itself on

L
L
DHT with LSH support
PUT
PUT
PUT
GET
GET
GET
Client
Resource

L

L

DHT with LSH support
PUT
PUT

Client
Resource
the higher structured layer. Users of the systems (rhombus) can query the structured network searching for the resources they need. Results will consist of the most significative community profiles and their representatives. The representatives will act as entry points to further forward queries to the peers of the represented community.

This architecture simplifies the task of searching for peers which profile is similar to the one given in input by a user, to the task of searching for communities described by a profile similar to the input one. This reduces both the amount of comparisons to perform and the amount of peers to contact. As a consequence, also the amount of generated network traffic decreases.

ATLASS-P2P ability to return significant resources has been tested using a dataset of 200 word domain labels organised in a hierarchical structure built by exploiting the WordNet domain [6].

F-Measure Comparison with ERGOT

Figure 2: Experiments

The content of this dataset has been used for generating textual descriptions, such description have been assigned to 5000 peers according to a Zipf distribution for building the peer profiles.

ATLASS-P2P performances are presented by Figure 2 that compares them to the ones provided by ERGOT [7], a solution aiming at the same goal but based on semantic overlay networks.

2. REFERENCES