# Ontology-Based Search and Broadcast in HyperCuP

Mario Schlosser Michael Sintek Stefan Decker Wolfgang Nejdl[1]
Computer Science Department, Stanford University
{schloss, sintek, stefan, nejdl}@db.stanford.edu

## 1 Introduction

Peer-to-peer networks are envisioned to be deployed in a wide range of applications, including the Semantic Web [2]. However, unorganized P2P networks exhibit characteristics that severely hamper their scalability to a large number of nodes, among them long search times, network traffic overload and traffic hotspots. In [1], we present HyperCuP, a topology for P2P networks which allows for very efficient search and broadcast, as well as a distributed construction and maintenance algorithm which organizes peers in a P2P network automatically into the desired topology. On this poster, we present a concept extending this scheme with ontology-based routing which can be exploited for directed search and broadcast on HyperCuP-based P2P networks.

## 2 HyperCuP: Shaping Up P2P Networks

A P2P network based on HyperCuP organizes its peers into a deterministic graph topology which features symmetry (i.e. every node can equally originate searches and broadcasts), low network diameter, a limited node degree (of logarithmic order to the amount of peers in the network) and load balancing of traffic in the network, avoiding traffic hotspots. We arrive at this by constructing a hypercube with all available peers, as depicted in Figure 1.

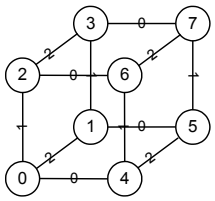


**Figure 1. A Hypercube topology with dimension 3**

Basically, to construct a hypercube of base 2 and dimension d, for each dimension the set of nodes is split into two partitions 0 and 1, which are defined by the d-th digit (0 or 1) in the binary encoding of the node number (from 0 to $2^d$-1), giving d dimensions for the hypercube and thus d neighbors for each node. Searching and broadcasting on this topology guarantee that each node receives the message exactly once, plus it exhibits a network diameter logarithmic to the amount of nodes. For this scheme, we provide a distributed construction and maintenance algorithm which permits nodes to join and leave the network arbitrarily (also for numbers of nodes not equal to $2^d$), with a low message overhead (logarithmic to the amount of nodes in the network), and not requiring any central servers or super nodes.

## 3 Ontology-Based Search and Broadcast

Usually, each peer in a P2P network can provide content that is associated with particular topics. These topics can be arranged as concepts in a global ontology, hierarchically classified by is-a links. It would be desirable to restrict the broadcast of a query message to peers that can potentially provide information related to the concepts asked for in the query. We address this problem by constructing more than one hypercube in the network: One hypercube is created for each ontology concept, creating a cluster of peers which carry information related to the concept. The wiring among the hypercubes mainly follows the is-a links in the ontology. Moreover, to answer queries consisting of logical combinations of concepts (conjunctions, disjunctions and negations), hypercubes are additionally organized internally: Inside the hypercubes, subclusters naturally exist – hypercubes consist of cubes, planes (e.g. peers 0, 1, 2, 3 in Figure 1) etc. Peers that are member of the hypercube associated with, say, concept A are now assigned to a subcluster that is associated with a logical combination of concepts they supports too. To equally fill all subclusters of a hypercube, the association of subclusters with logical concept combinations is chosen carefully, using the ontology much like a decision tree. If any peer now issues a query, it is routed to a cluster representing one of the non-negated concepts in the query, and then broadcasted only in the subcluster that is associated with the same logical combinations of concepts the query is aiming at. We contrast this approach to solutions such as [3].

## 4 References

[1] M. Schlosser, M. Sintek, S. Decker, W. Nejdl. HyperCuP – Shaping Up Peer-To-Peer Networks. *Submitted to DISC 2002*, April 2002.

[2] Nejdl, W. et al. Edutella: A P2P Networking Infrastructure based on RDF. In *Proceedings of the 11th International WWW Conference*, May 2002, Hawaii.

[3] S. Ratnasamy et al. Routing Algorithms for DHTs: Some Open Questions. In *Proceedings. of 1st International Workshop on Peer-to-Peer Systems*, March 2002.

[1] On leave from University of Hannover, Germany.