Querying Peer-to-Peer Networks Using P-Trees

Adina Costea    Prakash Linga    Johannes Gehrke    Jayavel Shanmugasundaram

Department of Computer Science
Cornell University
{adina,linga,johannes,jai}@cs.cornell.edu

Abstract

Peer-to-peer (P2P) systems provide a robust, scalable and decentralized way to share and publish data over the Internet. However, from a database point of view, P2P systems only provide very rudimentary query facility; they support just equality queries for files with a given name or keyword searches. We believe that future P2P systems will require complex querying functionality because users will publish data as semantically rich XML documents. In this paper, we focus on providing system support for a richer class of queries over P2P networks. We propose an index structure called the P-tree index, which is a distributed index over the data stored in a P2P network. The P-tree index provides support for efficient evaluation of range queries in addition to equality queries. We describe how to maintain a P-tree under insertions and deletions of data items, and we evaluate its performance in an extensive simulation study. Our results show the efficacy of our approach.