

# Information Retrieval in Peer-to-Peer Systems

**Wolf-Tilo Balke**

L3S Research Center and University of Hannover,  
Hannover, Germany  
*balke@l3s.de*

## **Abstract**

Applying information retrieval techniques in peer-to-peer infrastructures is a demanding challenge. Peer-to-peer systems are already being used for a vast number of applications in content exchange, but mostly searching is done by simple keyword lookups. In contrast, information Retrieval techniques vastly differ from retrieval in pure data management systems. Information retrieval means that not only some more or less matching objects have to be retrieved, but a list of the best matching objects over the entire network given a user's information needs. Since the 1960ies the information retrieval community considers ways to efficiently and effectively query document collections and designs dedicated retrieval systems like e.g. SMART. Since the content offered in P2P networks currently evolves from simple files that can be adequately described by some meta-data tags, to collections of complex documents that need somewhat more detailed descriptions, the development of refined retrieval techniques has to be fostered.

Usually a query is seen as a (possibly weighted) set of keywords a user specifies to express his/her information need. Documents that contain those (or sufficiently similar) keywords are considered to be relevant to the user's information need as expressed by the query. Thus, for information retrieval in peer-to-peer infrastructures the challenge is not only to retrieve documents efficiently, but also to effectively find a set of best matching objects. Generally speaking in this scenario retrieval effectiveness can only be traded for improved efficiency to a very limited degree. Moreover, in IR scenarios collection-wide statistical information is heavily used to improve retrieval effectiveness clashing with the very distributed nature of peer-to-peer scenarios. During the talk we will discuss the main issues and point to possible solutions and some ongoing research.

(Dagstuhl Seminar Series Nr 06131. 26.03.-29.03.06  
*A. Joseph, R. Steinmetz, I. Stoica, K. Wehrle:  
Peer-to-Peer-Systems and -Applications*)