ABSTRACT

SANTTHOSH BABU SELVADURAI

Implementing a Metasearch Framework with Content-Directed Result Merging
(Under the guidance of Dr. Gregory T. Byrd)

A metasearch engine is a system that provides integrated access to multiple existing search engine mechanisms. Once a query is executed on a metasearch engine, the system passes the query to its participating component search engines, collects the individual results and merges them into a single ranked list. Metasearch engines increase the search coverage of the Web, help solve the scalability issues in searching the internet, and improve the retrieval effectiveness, and consequently the relevance, of results.

Result merging is a key constituent of metasearch engines. When results from several search engines are collected, the metasearch system has to merge them into a unified list. The effectiveness of the metasearch mechanism and the relevance of the result set are closely related to the result merging algorithm used.

The purpose of this research is to build a flexible, general purpose metasearch framework and explore a content-directed result merging approach to rank results. Here the content-direction is provided to the framework by the user in the form of documents or text artifacts.

A modular metasearch application programming interface (API) based on java has been implemented. The API framework provides interfaces and utilities to develop components of a metasearch system like segregators, scheduler, aggregators, and search service providers. A prototype metasearch engine has been built based on this framework to study the content-directed result merging algorithm.

Implementing a Metasearch Framework with Content-directed Result Merging

by

Santthosh Babu Selvadurai

A thesis submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the degree of
Master of Science

Computer Engineering

Raleigh, NC

2007

Approved by:

_____

Dr. Gregory T. Byrd, Chair of Advisory Committee

_____   _____

Dr. Munindar P. Singh   Dr. Yan Solihin