Abstract

Because of the massive size and the speedy expansion of the Web, each search engine can only cover a small fraction of the Web. It is usually observed that many search engines cover different and overlapping portions of the Web. One most valuable way to increase the search coverage of the Web is to merge the coverage of multiple search engines. Model that performs such merging is called meta-search engine. A meta-search engine can be measured as a model that supports integrated access to multiple existing search engines. Using a meta-search engine, the user submits a search query to the meta-search engine, which passes the query to multiple existing search engines; when meta-search engine receives the search results returned from search engines, it merges these results into a ranked list and displays them on screen.

Research work was initiated for Designing new model for meta-search engine with new dimensions of resulting aspects. In this thesis a new model of the meta-search engine is discussed. The objective of research is to address the issues and problems exist with existing meta-search engines. A new model of meta-search engine has been developed for information retrieval in an efficient manner and introduced various search options with it, which are working dynamically.

This thesis is divided into eight chapters. The first chapter gives introduction of the web, search engine, search engine optimization, meta-search engine and meta-search engine optimization. It also discusses working of search engine and meta-search engine. Moreover, it discusses challenges in developing a new meta-search engine. In the last objective of the proposed research work has been discussed. It also furnishes layout of thesis.

Second chapter presents history of web surfing. It also presents initiative of search engine development, working of search engine, optimization technique used by search engine, limitations of search engine and
difference between search engine and meta-search engine. Moreover, it provides list of existing meta-search engines with their functionalities, differences and limitations. It also gives overview of different ranking methods. It also presents scope of research in this area.

Third chapter gives brief overview about each of the technology used for developing this new model of meta-search engine. Following technologies have been used for new meta-search engine development:

i. Operating system: ubuntu 09.04, ubuntu 12.04
ii. Open source server-side scripting language: PHP
iii. Scripting language: HTML
iv. Open source back-end: MySQL
v. Client-side scripting language: JavaScript
vi. Other technologies: CSS (Cascading Style Sheet) and Ajax

Fourth chapter describes the model of designed meta-search engine and presents work flow of a new model. It also discusses different search facilities and features of the new model in brief like, Default aggregate search, Selection search, LIKE search, and Direct search. In addition to these, it also introduces briefly about Administrator control. This meta-search engine also allows the user to search results related to images and video content. Also it presents need of stop word elimination process by sending input query to existing meta-search engine with or without stop words. The result shows that there is significant difference (more than 90%) in resultant links. The same feature is implemented in new model of meta-search engine and result shows that it retrieves resultant links almost same (> 90 % accuracy) with and without stop words. It also discusses about a new ranking formula for listing search results on screen by the proposed model.

Fifth chapter elaborates different search facility available to user. It gives brief about how does each of this facility work and what will be the output of each of it. It also measures response time of existing meta-search engines and new meta-search engine using webwait and pingdom
tools. Also, it shows PHP code to measure response time for search query by the new model. Moreover, it explains the use of firebug tool to measure load time performance of web page of existing meta-search engine and new meta-search engine. This analysis shows the performance of new model of meta-search engine is better than existing meta-search engines. In a new model of meta-search engine two new ranking formulas are introduced. One of them is static and other one is dynamic. Dynamic formula works on change in like/dislike values. The concept of database is used in a new model of meta-search engine, which improves efficiency of search result.

Sixth chapter discusses in detail various search facilities provided in a new model of meta-search engine like, aggregate search, selection search, LIKE search and direct search. It presents working of all search facilities with their sample screens and code.

Seventh chapter elaborates different administrator control available to administrator. It gives brief about how does each of this control work and what will be the outcome of each of it. It discusses concept of dynamic management of search engine’s URLs and its components in database to search the Web information from various search engines. This enables user to get search text from search engines listed in database and eliminates problem related to use of fixed number of search engines. It discusses database concept for management of stop word database, which allows administrator to manage stop words dynamically. It also discusses concept of updating links of searched keywords, which eliminates problem related to working with old retrieved search results. This new model of meta-search engine enables administrator to update search results available in database periodically. It also discusses creating merge table explicitly for merged results of search engine wise databases.

The eighth chapter discusses about impact of newly developed meta-search engine and future scope.
Research work in this era can be continued with different optimization techniques related to ranking formulas.

i. Organic search in meta-search engine through link analysis.
ii. New ranking formula for selection search.
iii. Normalization of indexed database based on count of keywords to improve efficiency and effectiveness of search results.
iv. Search results based on user profile or last access.
v. Dynamic management of keywords in meta-search engine.