



## Research Paper

# UPDATION OF URLS IN META SEARCH ENGINE DATABASE

Biraj Patel and Dr. Dipti Shah

### Address for Correspondence

G H Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University,  
Vallabh Vidyanagar, Gujrat, India

#### ABSTRACT

Existing meta search engines display search result on screen with no role of indexing. They basically send request for links on individual search engine and retrieve results and display aggregate result on screen based on score. In existing meta search engines there is no availability of index database for search process. In new proposed model the concept index database for link management is introduced. This paper discusses the need of updation of index database for link management in a meta search engine and discusses a model, which allows user to get search results from index database for input search text. A new model of the meta search engine comparatively retrieves speedily aggregate search results according to assigned rank in database.

**KEYWORDS** - search engine; meta search engine; index database; updation; search results.

#### I. INTRODUCTION

A meta search engine is a tool that retrieves the search results of multiple individual search engines. Each search engine on the Web has different strategies in terms of usage of different ranking methods to display links of web pages, using different resources. Users might have seen how these differences cause different search engines to return hugely different search results for the same input search text. To perform specific search for a query, users might need to use various individual search engines. Using a meta search engine users can search different engines at the same time, so user does not need to conduct the same search several times on different search engine. Meta search engines do not have their own databases for indexing of the Web information; instead, a meta search engine sends user input search text to several other search engines, search engines run the search text against their databases of the Web information and return results to the meta search engine. The meta search engine then returns consolidated results from all the search engines.

The problem with existing scenario is, meta search engines do not have their own databases for indexing purpose like general search engines and takes time for result retrieval. And because of that it causes problem of time out, which is there with mother of meta search engines “mamma” also. The results are combined from various individual search engines by meta search engine. They are displayed on screen based on ranking formula.

A new model of the meta search engine includes concept of databases for managing URLs and database will be updated periodically for effective search results like database updation with common search engines for indexed search results.

In database oriented meta search engine, to search text based contents first it will look in the database whether search text containing keywords are available in it or not. If user search keywords are already in database, then will retrieve web information from the database by sending simple query.[13] Otherwise, search query will be send to multiple individual search engines and retrieved result will be stored in the database and then through procedure web information from database will be displayed on user screen.[13]

Goal of using meta search engine is to get consolidated web results by sending queries to

multiple individual search engine. It is used to have limited number of hits on screen. Dealing with multiple individual search engines for this may retrieve large number of search results with duplicates. Instead let have database containing unique web results for specific search text will help user to get web information in an efficient way. [13] Communication between engine and database will be faster than communication between engine and the web is true but indexed updation is necessary component for this. So, in addition to this database introduction there is a requirement of periodical updating algorithms which will run queries and update indexed results periodically in the database like updating database of common search engines.

#### II. INDEXING

The purpose of storing an index is to optimize speed and performance in finding relevant documents for a search query. Without an index, the search engine would scan every document, which would require considerable time and computing power. [6] For example, while an index of 10,000 documents can be queried within milliseconds, a sequential scan of every word in 10,000 large documents could take hours. [6]

#### III. INDEXING IN SEARCH ENGINES

Search engine indexing collects, parses, and stores data to facilitate fast and accurate information retrieval. An alternate name for the process in the context of search engines designed to find web pages on the Internet is Web indexing. [6] The search engine index provide the primary structure for storing information about documents, including the information required to retrieve them: file system location, Web address (URL: Uniform Resource Locator), etc. [5]

Search engines employ varying strategies according to available resources and editorial or business objectives. Some simply crawl the entire Web on a fixed schedule (biweekly, monthly, etc.), re-indexing content as they go, others employ strategies based on factors like the expected frequency of changes (based on prior visits or site type, such as news media) and site quality heuristics. [5]

Whatever strategy is used to achieve optimal freshness, search engines are designed for incremental, differential data collection and index updating, and there is no technical barrier to high performance search engines performing updates for billions of pages. [5]

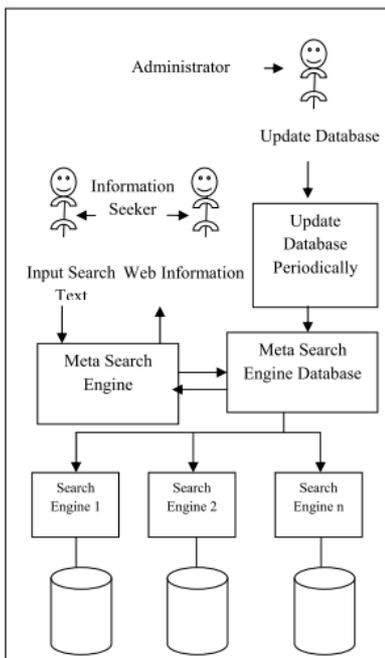
In existing scenario, meta search engines reuse the indices of other services and do not store local index, whereas cache-based search engines permanently store the index along with the associated details. Unlike full-text indices, partial-text services restrict the depth indexed to reduce index size. Larger services typically perform indexing at a predetermined time interval due to the required time and processing cost, while agent-based search engines index in real time. [6]

Recently, the need of information retrieval is especially increasing in corporations, universities and so on. In such organizations, it is very important to use fresh information. However, conventional search engines, which users usually use, spend very long term, e.g. one month, to update indexes of all documents because search engines are centralized systems. In centralized search engines, a www crawler collects documents from worldwide servers, and an indexer generates their indexes. Because of this, it is difficult for centralized search engine to satisfy the requirement. [7]

Organic listings of search engine are updated after long periods of time, the paid listings in it are updated much more regularly. Here, one can decide to perform a ranking check once per month, due to the fact that search engines normally update their ranking around the period of time. This check leads to new decisions; the result shows whether the measures of optimization have been successful, if not, they should be improved. It also shows if the results remain the same over a longer period of time or if an update is needed. [8]

**IV. THE METHOD**

Fig. 1 shows the new model of meta search engine with its own database and update database procedure as below.



**Fig.1 Meta search engine model with update database control**

Fig.2 shows PHP code for MySQL query, which updates indexed search results in meta search engine database.

```
select keyword from ".$sennms." where
DATEDIFF(NOW(),updated_at)>5
```

**Fig. 2 Update query for searched results**

Designed and implemented administrator side administrator panel manages various databases of

meta search engine. Developed module to update searched and indexed search results periodically. By clicking on single menu button administrator can update links in database effectively. Fig. 2 shows update query to update searched links, if difference of last updated date and current date is greater than 5; then updates records.

**V. CONCLUSION**

In existing meta search engines there is no concept of database for link management purpose. So, there is no requirement of any kind of operation also. In a new model of meta search engine concept of database for storage of search results is introduced, which enables user to get search results from the database. Also it allows administrator to update links of searched keywords in a database and eliminates problem related to working with old retrieved search results. Hence, this new model of meta search engine enables administrator to update search results available in database periodically and gives resultant links in a efficient time to user.

**REFERENCES**

1. Fossein Bidgoll (Editor-in-Chief), The Internet Encyclopedia, Volume 3, Copyright: 2004 by John Wiley & Sons, Inc., Page Title: Web Search Technology, PP 750
2. Jose Luis Vicedo, Patricio Martinez-Barco, Rafael Munoz, Maximilianosaiz Noeda (Eds.), Advances in Natural Language Processing, 4th International Conference, EsTAL 2004, Alicante, Spain October 2004, Proceedings, Copyright: Springer-Verlag Berlin Heidelberg 2004. (e-book)
3. Davidd Groth, Dorothy McGee, i-NET+TM, Study Guide, Second Edition, Copyright: 2002 SYBEX Inc. (e-book)
4. Foreword by Danny Sullivam, SEARCH ENGINE VISIBILITY, Shari Thurow, Copyright: 2003 by New Riders Publishing. (e-book)
5. Gregory Grefenstette, Laura Wilber, Search-Based Applications (At the Confluence of Search and Database Technologies), Copyright: 2011 by Morgan & Claypool
6. Digital Marketing Handbook (http://books.google.co.in/books?id=L4wr9mFq7nkC&pg=PA93&dq=updating+search+engines+index&hl=en&sa=X&ei=je4RUsfrHYHQrQfe6ICgBQ&ved=0CFgQ6AEwCA#v=onepage&q=updating%20search%20engines%20index&f=false)
7. G. Goos, J. Hartmanis and J. Van Leeuwen (Series Editors), Lecture Notes in Computer Science, Vol. 1966, Copyright: Springer-Verlag Berlin Heidelberg 2000
8. Dominik Mulhaupt, Internet marketing via search engines with an emphasis on platforms and pay-per-view portals (Diploma Thesis), Copyright: 2005 GRIN Verlag
9. Mr. Biraj Patel, Dr. Dipti Shah, Framework of Meta-Search Engine, Proceedings of the Multi-Conference 2011 (ICISD 2011) Universal-Publishers
10. Mr. Biraj Patel, Dr. Dipti Shah, Adaptive Systems and Search Engine Optimization, International Journal of Information and Computing Technology (RESEARCH@ICT) ISSN: 0976 – 5999 Volume-1 Issue – 2 May 2011 PP 14-15
11. Mr. Biraj Patel, Dr. Dipti Shah, Meta Search Ranking Strategies, International Journal of Information and Computing Technology (RESEARCH@ICT) ISSN: 0976 – 5999 Volume-2 Issue – 1 November 2011 PP 24-25
12. Mr. Biraj Patel, Dr. Dipti Shah, Ranking Algorithm For Meta Search Engine, International Journal of Advanced Engineering Research and Studies E-ISSN2249-8974, Vol. II/ Issue I/Oct.-Dec.,2012/39-40
13. Mr. Biraj Patel, Dr. Dipti Shah, Incorporation of Databases for Faster Meta Search Engine, International Journal of Advanced Research in Computer Science ISSN No. 0976-5697 Volume 3, No. 7, Nov-Dec 2012, 209-210