



Xml Integration with Meta-Search Engine

¹Dr. Biraj Patel, ²Dr. Dipti Shah, ³Mr. Jay Patel

¹G H Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujrat, India

²G H Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujrat, India

³Anand Commerce College, BCA Department, Nr. Town Hall, Anand, Gujarat, India

Abstract - Existing meta-search engines display results on screen without indexing. They send request for links on individual search engine and retrieve search results and display aggregate result on screen based on some rank. In existing meta-search engines there is no concept of index database for search process using XML. In proposed model the concept of index database for link management is introduced using XML. This research paper discusses the need of effective index database for link management in a meta-search engine and discusses a new model for the same, which allows user to get search results from XML index database for user input search text. The proposed model of the meta-search engine comparatively retrieves aggregate search results according to assigned rank in XML database speedily.

Keywords – Meta-search engine; Search engine; Indexing; XML; Rank.

I. INTRODUCTION

A meta-search engine is a kind of search tool which sends user requests to several other search engines and/or databases and aggregates the results into a single list and displays them according to rank. Meta-search engines enable users to enter search text once and access several search engines simultaneously. Meta-search engines operate on the premise that the Web is too large for any one search engine to index it all and that more comprehensive search results can be obtained by combining the results from several search engines. This also may save the user from having to use multiple search engines separately. [1]

Meta-search engines create what is known as a virtual database. They do not compile a physical database of the web. Instead, they take user input search text, pass it to several other heterogeneous databases and then compile the results in a homogeneous manner using some algorithm. MetaCrawler and Mamma are some of the earliest meta-search engines. [1] No two meta-search engines are alike. Some search only the most popular search engines while others also search lesser-known engines and other databases. They also differ in how the results are presented and the quantity of engines that are used. Some of them list results according to search engine or database. And some of them benefit the user by eliminating duplicate hits and grouping the most relevant ones at the top of the list. [1]

The problem with existing scenario is, meta-search engines do not have XML databases for indexing purpose and takes time for result retrieval. 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 rank assigned to them. [4]

Extensible Markup Language (XML) is a kind of markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable format. [2] The design goals of XML emphasize simplicity, generality, and usability over the Internet. It is a textual data format with strong support via Unicode for different human languages. Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services. [2]

A new model of the meta-search engine includes concept of XML 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. [4] 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. [5] 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. [5]

II. XML INTEGRATION WITH MODEL

Unlike relational database management system (RDBMS), XML documents do not require a pre-defined schema / pattern. Thus an XML repository management system, could not receive new documents containing new elements without breaking the system. With a relational database management system, one could simply add a new table, but that would be a manual task of changing the schema / pattern. XML gives documents consisting data in a common format.

Example of basic XML code consisting some data (test.xml):

```
<?xml version="1.0" encoding="UTF-8"?>
< note>
< to>Radha</to>
< from>Krishna</from>
< heading>Reminder</heading>
< body>Do not forget me this weekend!</body>
< /note>
```

PHP code given as below to access and load data of test.xml file into an array form:

```
<?php
```

```
$xml=simplexml_load_file("test.xml");
print_r($xml);
?>
```

The output of the code above will be:

```
SimpleXMLElement Object ( [to] => Tove [from] => Jani [heading] => Reminder [body] => Don't forget me this weekend! )
```

Alternate way is to output the data from each element in the XML file:

```
<?php
$xml=simplexml_load_file("test.xml");
echo $xml->to . "<br>";
echo $xml->from . "<br>";
echo $xml->heading . "<br>";
echo $xml->body;
?>
```

The output of the code above will be:

```
Radha
Krishna
Reminder
Do not forget me this weekend!
```

In proposed model of meta-search engine XML document integration is suggested. XML document is used for indexing of search results retrieved by meta-search engine. XML document structure will look like structure given as below:

File name: mse_search_result1.xml

```
<?xml version="1.0" encoding="UTF-8"?>
< results>
< urltitle>Definition: System</urltitle>
< url>http://www.tempscontents.com</url>
< /results>
```

Now, one could store XML document on some physical storage location or in some database for future access.

III. CONCLUSION

In existing meta-search engines there is no concept of database for link management purpose. In the proposed model of meta-search engine concept of XML database (It is always in document form) for storage of search results is introduced, which enables user to get search results from the same in an effective manner.

REFERENCES

- [1] http://en.wikipedia.org/wiki/Metasearch_engine [Accessed on 16/09/2014]
- [2] <http://en.wikipedia.org/wiki/XML> [Accessed on 16/09/2014]
- [3] http://www.w3schools.com/php/php_xml_simplexml.asp [Accessed on 16/09/2014]
- [4] Mr. Biraj Patel, Dr. Dipti Shah, Updation of URLs in Meta Search Engine Database, International Journal of Advanced Engineering Research and Studies, E-ISSN2249-8974
- [5] 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

- [6] Fossein Bidgoll (Editor-in-Chief), The Internet Encyclopedia, Volume 3, Copyright: 2004 by John Wiley & Sons, Inc., Page Title: Web Search Technology, PP 750
- [7] 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)
- [8] Gregory Grefenstette, Laura Wilber, Search-Based Applications (At the Confluence of Search and Database Technologies), Copyright: 2011 by Morgan & Claypool