



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

OPTIMIZED SEARCH ENGINE TECHNIQUES WITH SERVER DATABASE SEARCH FACILITY

¹Sanjay Sharma, ²Shrinath Tailor, ³Amit Kumar Sharma, ⁴Toofan Mukherjee, ⁵Dinesh Jain

^{1, 2, 3, 4, 5}Assistant Professor,

¹Computer Science and Engineering,

¹Sri Balaji College of Engineering & Technology, Jaipur, India

Abstract: As the trend of Internet is increasing the education and the learning process is also moving towards the data that is available on wikis and on the internet. To search the data we have Search Engines but what Search Engines lack to render us is the required keyword specific data on one go. Along with that all the secret information are also published on the internet that can be searched by any Search Engine. This implies a risk on the openly available data to get subjected to misuse. In this paper we present an overview of the methods to prevent that misuse and give effective and efficient search results to the users

Keyword- Search engine, Metadata, vod assisted Peer-to peer.

I. INTRODUCTION

The Search Engines present in the World Wide Web are capable of rendering efficient search by matching the Characters given as an input query with the metadata available in the head part of every webpage. We provide here different Fields to enter a query on what our user wants to find. Such as in order to search a .pdf[1] file related to any website on Google.com[2] we give keywords such as Inurl:www.example.com File type:. pdf. But here the main problem is that it shows results along with the websites that have www.example.com also written in their contents. We should provide along with meta search, a method to search databases also in order to receive the queried text from the specific website we need to get that data. Providing filtering options helps us in searching only the specified website for the file we need to extract.

II. MOTIVATION

Consider you are searching for all the .pdf documents in the Microsoft.com website, for which you enter the query in Google.com In URL www.microsoft.com File type:.pdf. This gives you around billions of pages that are contained in other websites also with the URL name www.microsoft.com and having .pdf as their file extension. This is not what you expected or meant while entering your query. A new method would be like, entering the query and specifying the file extension in different field along with the URL that you wish to search with. This search will import all the links available in the particular website's server with the file extension of .pdf and will give only those results that are received from the particular website. This will filter the extra bundle of data received from the whole World Wide Web.

III. EXISTING TECHNIQUES

3.1. In Existing Search Engines

The currently existing search engines in the web market provide us with the facility to search for files with keywords like In URL: to link the URLs of different websites in the search results, File type: to search the type of extension.

3.2. Meta search engines

A meta search engine is a search tool that sends user requests to many other search engines and/or databases, combining the results into a single list or presenting them according to their source. Meta search engines allow users to enter search criteria once and to simultaneously access multiple search engines. Meta search engines work on the assumption that the web is too vast for any single search engine to index it all and that it is possible to achieve more accurate search results by integrating the results of many search engines. This will also save the user from having to separately use several search engines.

3.3 Operation on Engine of meta search:

Meta search engines construct what is referred to as a virtual database. A physical archive or online catalogue is not compiled by them. Instead, they take a user's request, pass it to several other heterogeneous databases and then compile the results in a homogeneous manner based on a specific algorithm.

There are no two identical meta search engines. Others are only searching for the most popular search engines, while others are also looking for lesser-known engines, newsgroups, and other databases. They also vary in how the findings and the quantity of engines used are described. Some will list results according to the database or search engine. Others return results based on importance, frequently hiding which search engine returned the results. By removing redundant hits and grouping the most important ones at the top of the list, this helps the user.

Search engines also have distinct methods of expecting requests to be sent. For instance, some search engines allow the word "AND" to be used, while others need "+" and others only need space to combine terms. When sending them, the stronger meta search engines attempt to synthesize requests appropriately.

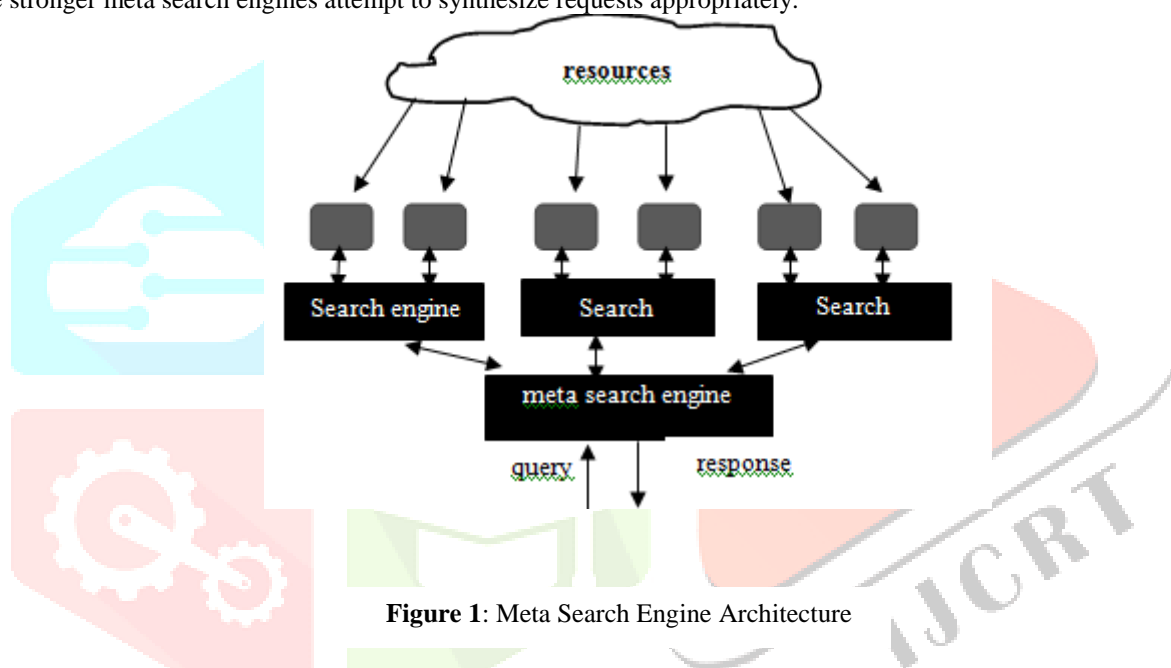


Figure 1: Meta Search Engine Architecture

IV. IMPLEMENTATION

Other than Federated search and other Meta-Search Engine, this method of searching will help us in gathering data on the basis of relevance to the exact material from the user defined source. Now the basic model of a search engine gives us results on the basis of number of hits the particular source website gets from the users, which diminishes the chances of getting exact data.

The architecture of this Server-Index Search Engine is shown in Figure 2. As shown the data is fetched from the specific server of the mentioned URL. For the sake of relevant indexing, we intend to use Meta-Search Engines too in order to provide facilities to the user which will allow browsing relevant links and then redefining the search on the basis of that link.

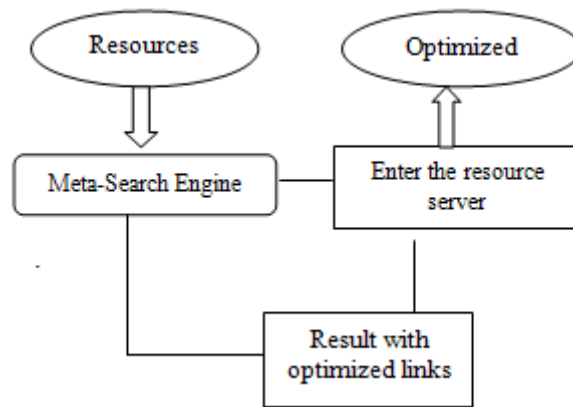


Figure2: Flow of new Server Oriented Search

The Meta-Search Engine will fetch relevant data from the resource, then the optimized results received due to the SEO (Search Engine Optimization) of websites, a particular/relevant link can be selected on the choice of the user or the resource can be accessed directly by entering the resource server.

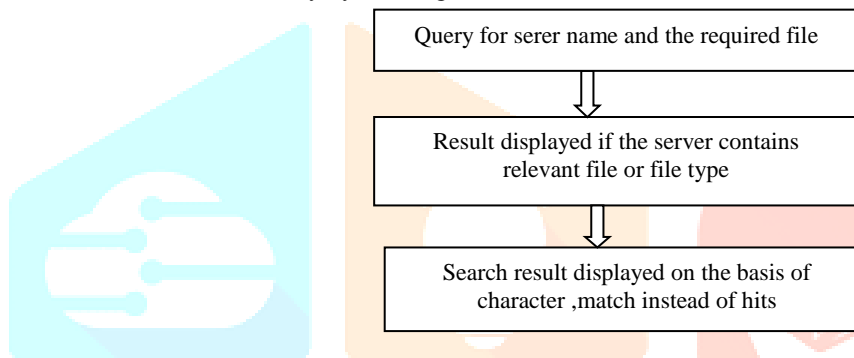


Figure3: Stepwise flow of filtered results

This method will match the queried character on with the files existing in the URL's Server, if not then results on the basis of Meta-Search will be displayed.

In total the results from the desired source only will be shown. In case the server is empty the results received from Meta-Search Engines will be further matched and shown to the user.

This process will allow more qualitative search results rather than quantitative search results. Instead of demanding for a needle in a haystack, demand for only the needle.

Server files are stored in the database of the server, which can be demanded for downloading through File Transfer Protocol (FTP). This will act as a dual purpose, server side search engine and meta-search engine. This reduces the possibilities of getting a dead link or bulky results which are very tedious and time consuming to filter manually. In Figure 4, the GUI has been implemented in order to give a wider view of the end product. In the page the user is required to enter the file type, which can be a Document, Media, Music, Compressed etc. The option in the extension type will behave according to the input of the File-Type Box. The extensions can be .pdf / .html / .chm / .epub / .doc / .xml etc. The extension will vary according to the chosen file type.

The URL Box allows you to enter your URL in which you want to specifically search your file, this can be left empty in order to perform meta-search of the input query. Entering a valid URL will result into the performing of Server side search, which will be bounded within the limits of the specified server only. If the specified file is not found in the server, the results will be evaluated on the basis of Meta-Search Results. The backend of this facility is a new alternative to the 'In URL:' keyword, used to search Meta links in the World Wide Web.

Base-Index Search

Search

Employee List

Select Extension Type (mpg, mov, pdf etc.) Select FileType (Video, Document, etc.) http://Enter URL to search in

Figure 4: Implemented GUI

REFERENCES

- [1] A. Newman, "Search engines used to attack databases", Technical report, Application Security, Inc. Available: http://www.appsecinc.com/presentations/Search_Engine_Attack_Database.pdf
- [2] Eric J. Glover, Steva Lawrance, William P. Birmingham and C. Lee Giles "Architecture of the Metasearch Engine that Supports User Information Needs", ACM.
- [3] K.Srinivas, P.V.S.Srinivas and A.Govardhan. "Web Service Architecture for a Meta Search Engine" International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011
- [4] K.Srinivas, P.V.S.Srinivas and A.Govardhan. "A Survey on the "Performance Evaluation of Various Meta Search Engines" IJCSI, Volume 8, Issue 3, No. 2, May 2011 Pages 359-364
- [5] Lee Underwood, "A Brief History of Search Engines"; www.webreference.com/authoring/search_history/.
- [6] DogPile metasearch engine. [tp://www.dogpile.com](http://www.dogpile.com).
- [7] Ritu Khare, Yuan An and Il-Yeol Song on Understanding Deep Web Search Interfaces: A Survey. SIGMOD Record, March 2010 (Vol. 39, No 1).

