

RANKING ALGORITHM FOR META SEARCH ENGINE

Mr. Biraj Patel and Dr. Dipti Shah

Address for CorrespondenceG H Patel Post Graduate Department of Computer Science and Technology,
Sardar Patel University, Vallabh Vidyanagar**ABSTRACT**

Various known meta search engines are built for access of multiple search engines. According to particular search keyword meta search engine returns a list of search results by avoiding duplicate entries. This article discusses a ranking algorithm for retrieved URLs for meta search engine.

1. INTRODUCTION

A meta search engine can be considered as a system that supports unified access to multiple existing search engines. [1] The meta search engine receives search keywords through user interface from user. Meta search engine queries multiple search engines then merges all search results returned from different search engines by eliminating duplicates.

For the retrieved results returned by different search engines to give rank to individual search result there is a requirement of standard communication between meta search engine and search engines.

2. ABOUT META SEARCH ENGINE SCOPE

A meta search engine has some intuitive appeals. First of all, it increases the search coverage of the Web. The Web is a huge information source, and each individual search engine may only cover a small portion of it. If user use only one search engine, user will never see those relevant pages that are not covered by that search engine. [2]

Meta search may also improve the search effectiveness. Each component search engine has its ranking algorithm to rank relevant pages, which is often biased. [2]

By combining the results from multiple search engines, their biases can be reduced and thus the search precision can be improved. [2]

The key operation in meta search engine is to combine the ranked results from the component search engines to produce a single ranking. [2]

The first task is to identify whether two pages from different search engines are the same. [2] If yes then easy way for duplicate removal and no need to think about ranking method.

The second task is to combine the ranked results from individual search to produce a single ranking, i.e., to fuse individual rankings. [2]

3. SEARCH RESULT POSITION ON A PAGE

There are two different meta search combination algorithms ones that use similarity of ranks returned by each individual search engine and ones that do not. [2]

4. PROBLEM DISCUSSION

Problem here in meta search engine is which search results should appear on first page with some rank, which should appear on second page with some rank, etc. If user uses individual search engine for single search query then each component of search engine has their own rank score. But here in case of meta search engine combining all search results is one job and another one is giving some rank to each search result. But problem here is how to produce a better combined ranking for effective results.

It is likely that search engines already use some techniques within their ranking mechanisms because a ranking algorithm need to consider multiple factors. So, there is a need of such kind of technique for meta search engine and new ranking method is discussed here.

5. ALGORITHM FOR RANKING

In meta search engine ranking can be done by focusing on rank positions on individual search results position on a particular page of search engine. For example, for a meta search engine if there is a use of three individual search engines then let have assumption that first page of each component has 5 search results (A, B, C, D and E) in which some of them may be common. For individual page results assign some integer value to the search result on a particular position.

If search result A is on first position of a page then assign rank to it 10, if it is on second position then assign rank to it 9 and if it is on last position then assign rank to it 1. There may be a chance that search result A on different search engines on different position of a page.

As per this example rank of A on meta search engine will be sum of result position on all three individual search engine pages. i.e., $(10+9+1=20)$. Similar way if search result B on one search engine on first position, B on second search engine second position and on third one it is on first position then its rank will be $(10+9+10=29)$.

Another alternate method wants to suggest using count function.

For example, if search result A is common of all three individual search engines part of meta search engine then the count will be 3. If it is common in any two of them then count will be 2. If it is in one of them then count will be 1. There is no focus on position on a page of individual component search engine. Then render results in descending order of count by avoiding duplicate entries.

There can be role of position of search result on a page of individual component search engine with this count function related method.

For example, there are five search results (A, B, C, D and E) on a page of 3 individual component search engines.

	Rank as per position				
	A	B	C	D	E
Search Engine 1	10	9	8	7	6
Search Engine 2	9	-	8	-	6
Search Engine 3	10	7	9	8	-

Search Result	Count
A	3
B	2
C	3
D	2
E	2

Formula suggested here:

$$\text{Rank}(x) = \sum_{i=1}^n P_i(x) / \text{count}(x) \text{ for all } i$$

Where, value of n = total number of search engines.

Search Result	Rank
A	9.66
B	8.00
C	8.33
D	7.50
E	6.00

As per above table meta search engine will render result in descending order of rank and will get their position as below.

Position #	Search Result
1	A
2	C
3	B
4	D
5	E

6. CONCLUSION

In this paper, ranking of retrieved URLs for meta search engine is discussed. The work involves, designing method for ranking retrieved results from different search engines for effective search results to increase reliability of meta search results.

REFERENCES

- [1] Advanced Metasearch Engine Technology by Weiyi Meng and Clement T. Yu, Copyright: 2011 by Morgan & Claypool. (e-book)
- [2] Web Data Mining (Exploring Hyperlinks, Contents and Usage Data) Second Edition by Bing Liu, Copyright: Springer-Verlag Berlin Heidelberg 2007, 2011. (e-book)
- [3] Advances in Data and Web Management, Joint International Conferences, APweb/WAIM 2009, Suzhou, China, April 2009 Proceedings; Qing Li, Ling Feng, Jian Pei, Sean X. Wang, Xiaofang Zhou, Qiao-Ming Zhu (Eds.) Copyright: Springer-Verlag Berlin Heidelberg 2009. (e-book)
- [4] Meta Search Ranking Strategies, published in the International Journal of Information and Computing Technology (RESEARCH@ICT) by Mr. Biraj Patel, Dr. Dipti Shah; ISSN: 0976-5999, Volume-2 Issue-1 November 2011.