

# Web Page Ranking using Domain based Knowledge

Sutirtha Kumar Guha  
Seacom Engineering College,  
Howrah, India  
sutirthaguha@gmail.com

Anirban Kundu  
Netaji Subhash Engineering  
College,  
Garia, India  
anik76in@gmail.com

Rana Dattagupta  
Jadavpur University,  
Kolkata, India  
ranadattagupta@yahoo.com

**Abstract**—In this paper, a new ranking technique is introduced to rank Web pages categorically. Web pages are categorized as ‘Primary’ and ‘Secondary’. It is considered that available Web page URLs are resided in a predefined database. ‘Primary’ Web pages are selected based on the keywords. These Web pages are ranked by a newly introduced equation. It is evident that all matched Web pages are not selected by this keyword selection procedure. Hence, unmatched Web pages are checked and named as ‘Secondary’ Web pages. These Web pages are ranked by another new equation. Ranking procedure of ‘Primary’ and ‘Secondary’ Web pages is different based on the different selection process. Final outcome is obtained by merging and ranking all matched Web pages. Wide number of Web pages having matched contents is obtained by implementing this proposed procedure.

**Keywords:** *Web page ranking; Primary Web pages; Secondary Web pages;*

## I. INTRODUCTION

Immense usage of Web sphere has led to a situation where role of Search Engine is invaluable in navigating through the vast expanse of the Web world to find the most relevant information. It is highly desirable to get most accurate and wide number of information for a particular query in Search Engine environment.

Primary objective of a search engine is to collect relevant and useful Web pages from the extensive database of the search engine and display them in a chronological order for the user. Desired result may not be obtained every time as the searching procedure is executed mechanically. It may happen that URL that may not tally fully with the user given query but match a few keywords may cater to users’ requirements.

In typical Search engine environment Web pages having matched keywords are fetched from the predefined database and ranked. Wide range of relevant Web pages having unmatched keywords are omitted in this procedure. Neglected Web pages having relevant information are covered by the proposed method. Relevant Web pages are fetched and displayed irrespective of keyword matching.

Primary and Secondary Quotient values are introduced to consider and rank all relevant Web pages irrespective of keywords matching with respect to the user query. Similarity index concept is proposed to measure the rank

of Web pages having partially matched keyword Web pages having partially matched keywords are collected by nearest keywords and similar Web pages concept as discussed.

Rest of the paper is organized as follows: Section II represents related works; proposed system framework has been described in Section III; Section IV contains analytical study; Experimental result is shown in Section V; Section VI concludes the paper.

## II. RELATED WORK

In real-world scenario different interrelated domains are ranked where available training data is not evenly distributed [1]. In [2] probability based ranking method is proposed. Automatic sentence extraction using graph-based ranking algorithms is introduced in [3]. Source of evidences are used to produce distinct concept based ranking that are finally combined into a final ranking. Potential of collaborative search method to produce better search performance for end user is described in [4]. Graph-based ranking method is proposed in [5] to rank email correspondents according to their degree of expertise on subjects of interest.

Probabilistic approach in ranking is introduced in [6] to fetch domain information retrieval. Probabilistic Gaussian Process classifiers are introduced to rank documents and models accurately [7]. The learning framework is applied to the settings of aggregating full rankings and aggregating top-k lists, demonstrating significant improvements over a domain-agnostic baseline in both cases [8]. Different features of a Web page is used to rank the matched pages in [9]. Session in Web pages is examined to avoid the unwanted noise in [10]. Web page ranking procedure using Semantic Similarity and HITS algorithm is introduced in [11]. A modified version of HIT algorithm is proposed in [12] to measure Web page rank. In [13] a new rank process is proposed to rank Web pages that are not connected with Search Engine. Web page rank is calculated based on data mining concept is introduced in [14].

Different Web mining algorithms used in page ranking are discussed in [15]. Page rank algorithm to rank and track Web pages is proposed in [16.] Ranking procedure based on a variant of the h-index for directed

information networks is proposed in [17]. This variant is used to measure the weightage of a Web page [17]. Different semantic ranking approaches based on semantic are analyzed in [18] to identify some special features of the approaches. A domain specific ranking model for a broad-based search engine is proposed in [19]. Page rank procedure to rank Web pages based on user favorite domain is proposed in [20]. User specific domain based searching method is proposed in [21]. Domain and domain details are given by the user and user query is searched from the user given domain [21].

In [22] seed URLs are classified and different domains are formed. User query is analyzed and sent to the specific domain to fetch matched Web pages [22]. A relation based ranking algorithm is proposed in [23]. Reformulation of Linear rank, Total rank and Generalized Hyperbolic Rank is proposed in [24]. Different Web page ranking algorithms based on machine learning are discussed in [25]. Various ranking procedure implemented in digital libraries are analyzed in [26].

### III. PROPOSED APPROACH

Different domain specific searching and ranking techniques are proposed by the researchers. In this paper, domains are constructed based on direct matching of user query and Web page keywords. In existing approaches, domains are formed based on the contents of the Web pages. Domain selection in existing approaches is made in server side. In our approach, existing Web pages are categorized based on the user query at runtime. Wide range of matched Web pages is fetched for the user. Web pages having minimum possibility of matching with the user query are fetched and displayed by this approach.

In this paper, it is considered that Web page URLs are resided in a predefined database. User given query is matched with the Web page URL database and based on the similar Web pages are displayed rank wise. These Web pages are named as 'Primary' Web page. The matching procedure is explained in Fig. 1.

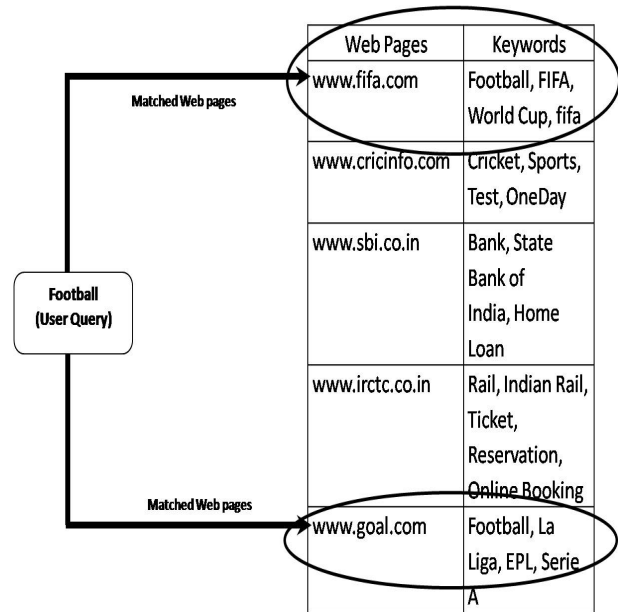


Fig. 1. 'Primary' Web page selection based on Keyword

'Primary' Web pages are ranked by Primary Quotient value based on Equation 1. Highest rank is indicated by maximum quotient value.

$$R_1(W) = \left( \frac{1}{\sum PR(W)} + \frac{\sum S_w}{N_w} \right) \quad \dots\dots\dots (1)$$

Where,

$R_1(W)$  = Primary Quotient Value of Web page 'W'

$PR(W_i)$  = Page Rank of Inbound Web page in 'W'

$S_w$  = Visitors' session in Web page 'W'

$N_w$  = Total number of visitors' in Web page 'W'

It is evident that a Web page containing relevant information with respect to the user query may not have similar keywords. Keywords are considered as an indicator not as decisive factor. Our paper is aiming this dilemma of Web page selection based on keyword matching criteria. This drawback is handled in this paper by introducing 'Secondary' Web page concept. 'Secondary' Web pages are considered as the Web pages whose keywords do not match with the user query but contents match with the user requirement. These Web pages are selected based SimilarityIndex value as shown in Equation 2.

Web page selection process is depicted in Algorithm 1.

Algorithm 1: matchKeyword (K,  $W_p$ )

Input: User given query as 'K' and Web page that is not selected as 'Primary' and named as ' $W_p$ '.

Output: Selection of 'Secondary' Web page

Step 1: Consider a variable 'A' as Digging Factor having initial numeric value '2'.

Step 2: Consider the parent Web page of  $W_p$  and named

that Web page as  $W_1$

Step 3: If  $W_P$  is the root Web page then go to Step 6.

Else, go to Step 4.

Step 4: Calculate SimilarityIndex based on Equation 2,

$$I_s = \frac{(K_{FM} + 0.5) \times K_{0.5}}{K_N(W)} \quad \dots\dots (2)$$

where,

$I_s$ = Similarity Index

$K_{FM}$ = Number of Fully Matched Keywords

$K_{0.5}$ = Number of 50% Matched Keywords

$K_N(W)$ = Total Number of Keywords in Web page  $W$

Step 5: If SimilarityIndex is more than 50% then, tag  $W_P$  as ‘Secondary’ Web page.

Go to Step 7.

Else, go to Step 6.

Step 6: Consider the parent Web page of  $W_1$  and named that Web page as  $W_1$ . Increment the value of  $A$  by 1. Go to Step 7.

Step 7: If, there is no Web page URL left in the database go to Step 7.

Else, Consider another Web page and go to Step 1.

Step 8: End.

‘Secondary’ Web page is selected from the Web pages that do not contain any matched keywords. Hence, the parents of those Web pages are considered. It is presumed here that all features of parent Web pages are inherited by the child Web pages. Hence, relevant contains are resided in the Web page whose parent Web page have matching keywords. Digging Factor  $A$  is used to track the level of digging the hierarchy for any particular Web page in bottom-up direction.

‘Secondary’ Web pages are ranked Secondary Quotient value as shown in Equation 3. Highest rank is indicated by maximum quotient value.

$$R_2(W) = \frac{\left( \frac{1}{\sum PR(W_1)} + \frac{\sum S_w}{N_w} \right)}{D_F} \quad \dots\dots (3)$$

Where,

$R_2(W)$  = Secondary Quotient value of ‘Secondary’ Web page ‘ $W$ ’

$PR(W_1)$  = Page Rank of Inbound Web page in ‘ $W$ ’

$S_w$  = Visitors’ session in Web page ‘ $W$ ’

$N_w$  = Total number of visitors’ in Web page ‘ $W$ ’

$D_F$  = Digging Factor

It is evident from Equation 3 that ‘Secondary’ Web page rank is constructed by dividing the ‘Primary’ Web page Rank process by Digging Factor. Hence, same features, conditions are considered for both ranking process. Importance of ‘Secondary’ Web page is relatively less compare to same ranked ‘Primary’ Web page due to lack of keyword matching. This is reflected in

‘Secondary’ Web page ranking. Hence, all selected Web pages are merged and displayed according to their rank calculated by Equation 1 and 3.

#### IV. ANALYTICAL STUDY

Let consider the following case study at any time instance as shown in Fig. 2.

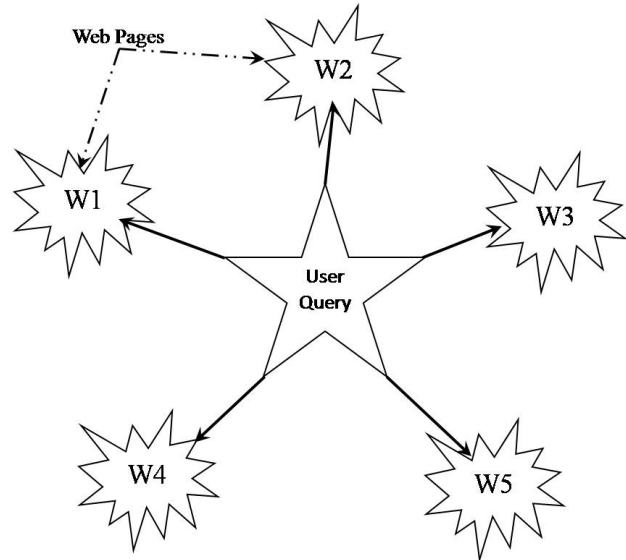


Fig. 2. User query and available Web pages

It is considered in our case study that user query is sent to five Web pages named as  $W_i$ . These five Web pages are categorized as ‘Primary’ and ‘Secondary’ based on the proposed method. Classification and other details of the Web pages are depicted in Table I.

TABLE I. DETAILS OF CLASSIFIED WEB PAGES

Web Page( $W$ )	Classification	Inbound Web page( $W_i$ )	Page rank of Inbound Web page( $PR(W_i)$ )	Total Session in $W(\sum S_w)$	Total Number of Visitors in $W(N_w)$
$W_1$	UNMATC HED	$W_{111}$	6	16	3
		$W_{112}$	15		
		$W_{113}$	7		
$W_2$	Primary	$W_{121}$	5	26	8
		$W_{122}$	3		
		$W_{123}$	10		
		$W_{124}$	4		
$W_3$	UNMATC HED	$W_{131}$	11	20	6
		$W_{132}$	6		
		$W_{133}$	12		
$W_4$	Primary	$W_{141}$	3	8	3
$W_5$	UNMATC HED	$W_{151}$	9	13	3
		$W_{152}$	14		

Rank of ‘Primary’ Web pages are calculated based on

Equation 1 and shown in Table II.

TABLE II. RANK CALCULATION OF 'PRIMARY' WEB PAGES

Web page (W)	$\sum(PR(W_i))$	$\sum S_w$	$N_w$	$R_1(W)$ value
$W_2$	22	26	8	3.29
$W_4$	3	8	3	2.99

'Secondary' Web page selection is depicted in Fig. 3.

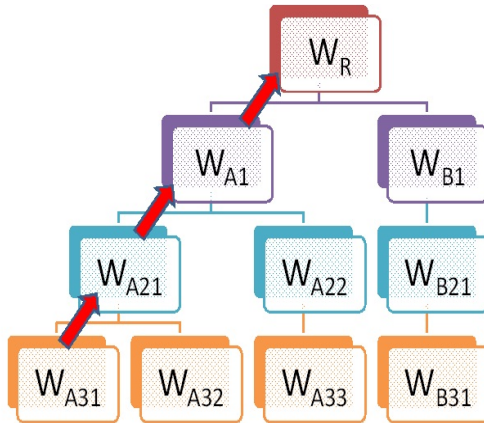


Fig. 3. 'Secondary' Web page selection

'Secondary' Web pages are selected based on our proposed method. Selection method for Web page  $W_{A31}$  is shown in Fig. 3.  $W_{A21}$  parent Web page of  $W_{A31}$  is inspected first to find matched keywords with respect to user query. If no such matching keywords found then parent  $W_1$  Web page of  $W_{A21}$  is inspected. The process

is going on until matched keywords found. If matched keywords are not found till the root Web page, then next Web page will be checked. Similar process will be implemented for  $W_{A32}$ ,  $W_{A33}$  and  $W_{B31}$ . Selected 'Secondary' Web pages and their details are shown in Table III.

TABLE III. RANK CALCULATION OF 'SECONDARY' WEB PAGES

Web page	$\sum(PR(W_i))$	$\sum S_w$	$N_w$	Digging Factor ( $D_F$ )	$R_2(W)$ value
$W_1$	20	16	3	3	1.793
$W_5$	23	13	6	2	1.101

Four matched Web pages are found according to our proposed approach. Rank value of these Web pages and their rank is shown in Table IV.

TABLE IV. RANK OF MATCHED WEB PAGES

Web pages	Rank value	Rank
$W_1$	1.793	03
$W_2$	3.29	01
$W_4$	2.99	02
$W_5$	1.101	04

## V. EXPERIMENTAL RESULT

'Primary' and 'Secondary' Web pages rank calculation method is demonstrated by considering the URL 'http://www.capexindia.com/IRL'. Web pages of the considered URL are oriented as shown in Fig. 4.

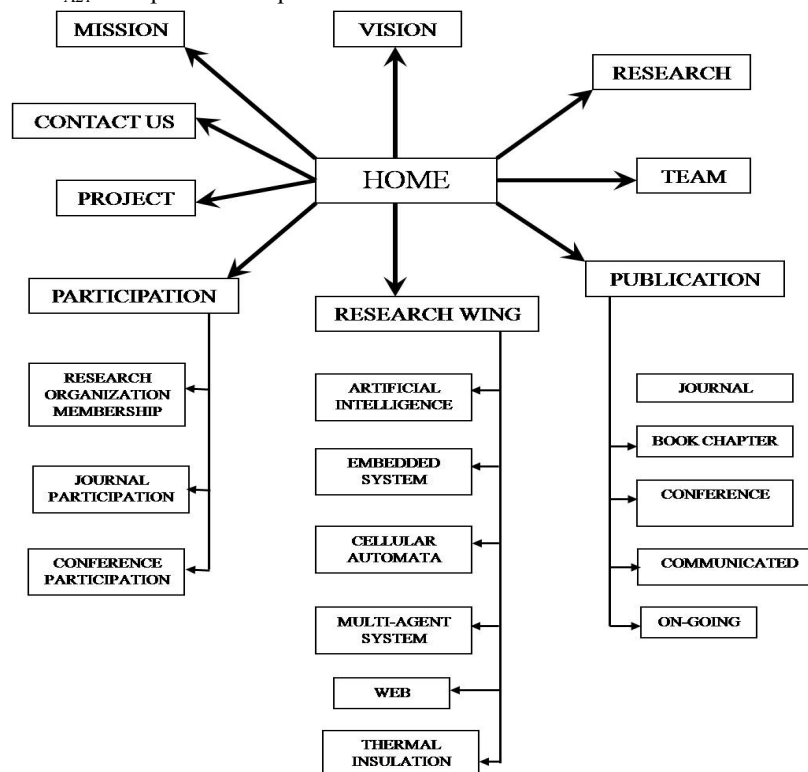


Fig. 4. Web page orientation of URL 'http://www.capexindia.com/IRL'

Let user query at time instance  $t_i$  is 'Multi-Agent'.  
Details of classified Web pages are shown in Table V.

TABLE V. DETAILS OF CLASSIFIED WEB PAGES FOR USER QUERY 'MULTI-AGENT'

Web Page	Web Page(W)	Classification
WI	<a href="http://www.capexindia.com/IRL/index.htm">http://www.capexindia.com/IRL/index.htm</a>	UNMATCHED
WM	<a href="http://www.capexindia.com/IRL/mission.htm">http://www.capexindia.com/IRL/mission.htm</a>	UNMATCHED
WV	<a href="http://www.capexindia.com/IRL/vision.htm">http://www.capexindia.com/IRL/vision.htm</a>	UNMATCHED
WR	<a href="http://www.capexindia.com/IRL/research.htm">http://www.capexindia.com/IRL/research.htm</a>	MATCHED
WAI	<a href="http://www.capexindia.com/IRL/centre_for_AI.htm">http://www.capexindia.com/IRL/centre_for_AI.htm</a>	UNMATCHED
WES	<a href="http://www.capexindia.com/IRL/centre_for_ES.htm">http://www.capexindia.com/IRL/centre_for_ES.htm</a>	UNMATCHED
WCA	<a href="http://www.capexindia.com/IRL/centre_for_CA.htm">http://www.capexindia.com/IRL/centre_for_CA.htm</a>	UNMATCHED
WMAS	<a href="http://www.capexindia.com/IRL/centre_for_MAS.htm">http://www.capexindia.com/IRL/centre_for_MAS.htm</a>	MATCHED
WWEB	<a href="http://www.capexindia.com/IRL/centre_for_web.htm">http://www.capexindia.com/IRL/centre_for_web.htm</a>	UNMATCHED
WTI	<a href="http://www.capexindia.com/IRL/centre_for_TI.htm">http://www.capexindia.com/IRL/centre_for_TI.htm</a>	UNMATCHED
WT	<a href="http://www.capexindia.com/IRL/team.htm">http://www.capexindia.com/IRL/team.htm</a>	UNMATCHED
WPJ	<a href="http://www.capexindia.com/IRL/published_journal.htm">http://www.capexindia.com/IRL/published_journal.htm</a>	MATCHED
WPBC	<a href="http://www.capexindia.com/IRL/book_chapter.htm">http://www.capexindia.com/IRL/book_chapter.htm</a>	MATCHED
WPC	<a href="http://www.capexindia.com/IRL/published_conference.htm">http://www.capexindia.com/IRL/published_conference.htm</a>	MATCHED
WCP	<a href="http://www.capexindia.com/IRL/communicated_papers.htm">http://www.capexindia.com/IRL/communicated_papers.htm</a>	UNMATCHED
WOGP	<a href="http://www.capexindia.com/IRL/on_going_papers.htm">http://www.capexindia.com/IRL/on_going_papers.htm</a>	MATCHED
WMEM	<a href="http://www.capexindia.com/IRL/member.htm">http://www.capexindia.com/IRL/member.htm</a>	UNMATCHED
WJP	<a href="http://www.capexindia.com/IRL/journal_participation.htm">http://www.capexindia.com/IRL/journal_participation.htm</a>	UNMATCHED
WCPART	<a href="http://www.capexindia.com/IRL/conference_participation.htm">http://www.capexindia.com/IRL/conference_participation.htm</a>	UNMATCHED
WPROJ	<a href="http://www.capexindia.com/IRL/projects.htm">http://www.capexindia.com/IRL/projects.htm</a>	MATCHED
WCON	<a href="http://www.capexindia.com/IRL/contact.htm">http://www.capexindia.com/IRL/contact.htm</a>	UNMATCHED

User hit data for the query 'Multi-Agent' for the given URLs are shown in Fig. 5.

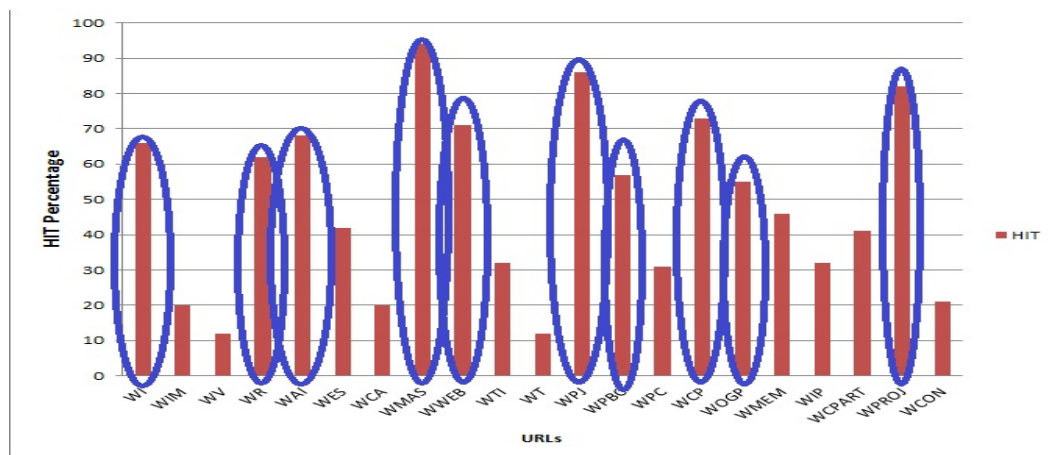


Fig. 5. User HIT data

It is evident from Figure 5 that relevant URLs for query 'Multi-Agent' are as follows:

<http://www.capexindia.com/IRL/index.htm>  
<http://www.capexindia.com/IRL/research.htm>  
[http://www.capexindia.com/IRL/centre\\_for\\_AI.htm](http://www.capexindia.com/IRL/centre_for_AI.htm)  
[http://www.capexindia.com/IRL/centre\\_for\\_MAS.htm](http://www.capexindia.com/IRL/centre_for_MAS.htm)  
[http://www.capexindia.com/IRL/centre\\_for\\_web.htm](http://www.capexindia.com/IRL/centre_for_web.htm)  
[http://www.capexindia.com/IRL/published\\_journal.htm](http://www.capexindia.com/IRL/published_journal.htm)  
[http://www.capexindia.com/IRL/published\\_journal.htm](http://www.capexindia.com/IRL/published_journal.htm)  
[http://www.capexindia.com/IRL/book\\_chapter.htm](http://www.capexindia.com/IRL/book_chapter.htm)

[http://www.capexindia.com/IRL/communicated\\_papers.htm](http://www.capexindia.com/IRL/communicated_papers.htm)  
[http://www.capexindia.com/IRL/on\\_going\\_papers.htm](http://www.capexindia.com/IRL/on_going_papers.htm)  
<http://www.capexindia.com/IRL/projects.htm>

In this work Web pages are classified based on their fitness with the keywords. Keywords of the Web pages are not shown in table due to inadequate space. Hence, rank of 'Primary' Web pages is calculated based on Equation 1 and shown in Table VI.

TABLE VI. RANK CALCULATION OF 'PRIMARY' WEB PAGES FOR USER QUERY 'MULTI-AGENT'

Web page (W)	$\Sigma(PR(W_i))$	$\Sigma Sw$	$Nw$	$R_i(W)$ value
<a href="http://www.capexindia.com/IRL/research.htm">http://www.capexindia.com/IRL/research.htm</a>	04	14	04	3.5
<a href="http://www.capexindia.com/IRL/centre_for_MAS.htm">http://www.capexindia.com/IRL/centre_for_MAS.htm</a>	04	32	03	10.91



http://www.capexindia.com/IRL/published_journal.htm	04	26	05	5.45
http://www.capexindia.com/IRL/book_chapter.htm	04	07	03	2.58
http://www.capexindia.com/IRL/on_going_papers.htm	04	07	03	2.58
http://www.capexindia.com/IRL/projects.htm	04	09	02	4.75

Secondary Web pages are selected by our proposed approach as discussed in Section 3. Selected

‘Secondary’ Web pages are ranked. Rank calculation for ‘Secondary’ Web page is shown in Table VII.

TABLE VII. RANK CALCULATION OF ‘SECONDARY’ WEB PAGES

Web page	$\Sigma$ (PR(W <sub>i</sub> ))	$\Sigma$ S <sub>w</sub>	N <sub>w</sub>	Digging Factor (D <sub>f</sub> )	R <sub>2</sub> (W) value
http://www.capexindia.com/IRL/index.htm	04	24	10	01	2.65
http://www.capexindia.com/IRL/centre_for_AI.htm	04	21	02	03	3.58
http://www.capexindia.com/IRL/centre_for_web.htm	04	26	02	03	4.41
http://www.capexindia.com/IRL/communicated_papers.htm	04	32	03	03	3.63

Hence, finally matched URLs are ranked as shown in Table VIII.

TABLE VIII. RANK OF THE WEB PAGES

Web pages	Web pages (URL)	Quotient Value	Rank
WR	http://www.capexindia.com/IRL/research.htm	3.5	07
WMAS	http://www.capexindia.com/IRL/centre_for_MAS.htm	10.91	01
WPJ	http://www.capexindia.com/IRL/published_journal.htm	5.45	02
WPBC	http://www.capexindia.com/IRL/book_chapter.htm	2.58	09
WOGP	http://www.capexindia.com/IRL/on_going_papers.htm	2.58	10
WPROJ	http://www.capexindia.com/IRL/projects.htm	4.75	03
WI	http://www.capexindia.com/IRL/index.htm	2.65	08
WAI	http://www.capexindia.com/IRL/centre_for_AI.htm	3.58	06
WWEB	http://www.capexindia.com/IRL/centre_for_web.htm	4.41	04
WCP	http://www.capexindia.com/IRL/communicated_papers.htm	3.63	05

In real time scenario user interest towards a specific Web page is reflected by user hit data of that Web page. Hence a comparison between users hit data and

calculated rank value would be considered as a comparative study of proposed approach and real time scenario. The comparative study of all Web pages is shown in Fig. 6.

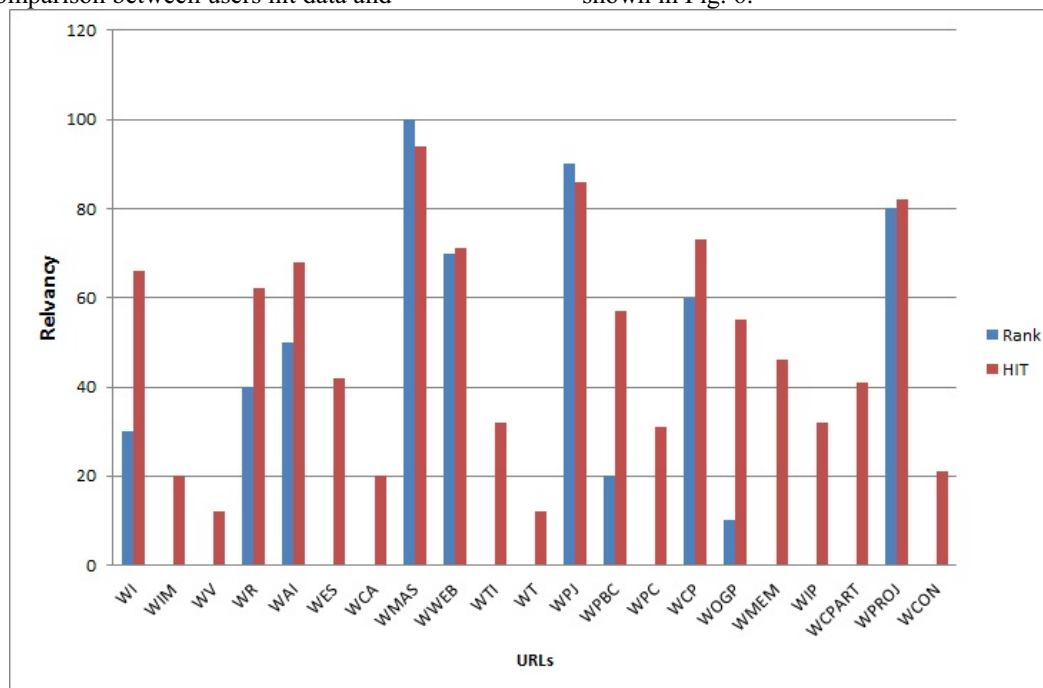


Fig. 6. Comparative study of proposed approach and ideal situation

It is evident from Figure 6 that more appropriate outcome is achieved by our proposed approach. It is observed that real time user favorite URLs and our ranked URLs are matched. Hence, most feasible result is achieved by our proposed approach.

## VI. CONCLUSION

In this paper Web pages are selected from predefined Web page URL database based on the user query. It is evident that a wide area of matched Web pages are displayed and ranked by our proposed method. More vast and satisfying searching result is obtained without affecting accuracy and efficiency from users' point of view by implementing this approach. 'Primary' Web pages are selected based on keyword matching. 'Secondary' Web pages are selected based on similarity between user query and keywords of parent Web pages of unmatched Webpages. Proper importance and distinction is maintained between 'Primary' and 'Secondary' Web pages by introducing Digging Factor. Large number of Web pages is selected and produced based on user query by our proposed method.

## REFERENCES

- [1] B. Wangy, J. Tang, W. Fan, Songcan Chen, Zi Yang, Yanzhu Liu, "Heterogeneous Cross Domain Ranking in Latent Space", 18th ACM Conference on Information and Knowledge Management (CIKM'09), Hong Kong, China, November, 2009, pp. 987-996.
- [2] F. Provost, P. Domingos, "Tree Induction for Probability-Based Ranking", Machine Learning, Volume: 52, Issue: 3, 2003, pp. 199-215.
- [3] R. Mihalcea, "Graph-based ranking algorithms for sentence extraction, applied for text summarization", the ACL 2004 on Interactive poster and demonstration sessions (ACLDemo'04), July, 2004
- [4] M. L. Silveira, B. Ribeiro-Neto, "Concept-based ranking: a case study in the judicial domain", Information Processing & Management, Volume: 40, Issue: 5, 2004, pp. 791-805.
- [5] J. Freyne, B. Smyth, M. Coyle, E. Balfe, P. Briggs, "Further Experiments on Collaborative Ranking in Community-Based Web Search", Artificial Intelligence Review, Volume: 21, Issue: 3-4, 2004, pp. 229-252.
- [6] B. Dom, I. Eiron, A. Cozzi, Y. Zhang, "Graph-based ranking algorithms for e-mail expertise analysis", the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery, 2003, pp. 42-48.
- [7] H. Li, Z. Li, W. Lee, D. L. Lee, "A probabilistic topic-based ranking framework for location-sensitive domain information retrieval", 32nd international ACM SIGIR conference on Research and development in information retrieval, Boston, USA, 2009, pp. 331-338.
- [8] A. Klementiev, D. Roth, k. Small, I. Titov, "Unsupervised Rank Aggregation with Domain-Specific Expertise", 21<sup>st</sup> International Joint Conference on Artificial Intelligence (IJCAI'09), San Francisco, USA, 2009, pp. 1101-1106.
- [9] S. K. Guha, A. Kundu, S. Bhadra, R. Dattagupta, "Dynamic Web-page Ranking of Search Engine using Importance Increasing Factor", International Journal of Digital Content Technology and its Applications (JDCTA), Volume 7, Number 12, 2013, pp. 28-39.
- [10] S. K. Guha, A. Kundu, R. Dattagupta, "Introducing Session Relevancy Inspection in Web Page", Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012), Delhi, India, Springer, May 25-27, 2012, pp. 181-192.
- [11] N. Bansal, P. Singh, "Improved Web Page Ranking Algorithm Using Semantic Similarity and HITS Algorithm", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 4, 2014, pp. 346-348.
- [12] P. Rani, S. Singh, "An Offline SEO (Search Engine Optimization) Based Algorithm to Calculate Web Page Rank According to Different Parameters", International Journal of Computers & Technology, Volume 9, Number 1, 2013.
- [13] N. Batra, A. Kumar, D. Singh, P. N. Rajotia, "Content Based Hidden web Ranking Algorithm (CHWRA)", IEEE International Advance Computing Conference (IACC), 2014, pp. 586-589.
- [14] A. Jain, R. Sharma, G. Dixit, V. Tomar, "Page Ranking Algorithm in Web Mining, Limitations of Existing methods and a New Method for Indexing Web Pages", International Conference on Communication Systems and Network Technologies (CSNT), 2013, pp. 640-645.
- [15] M. Sote, S. R. Pande, "Application of Page Ranking Algorithm in Web Mining", International Conference on Advances in Engineering & Technology, 2014, pp. 47-51.
- [16] G. Chen, F. Keren, A. Loza, W. Qiang, L. Jia, Y. Jie, "PageRank Tracker: From Ranking to Tracking", IEEE Transaction on Cybernetics, Volume 44, Issue 6, 2014, pp. 882-893.
- [17] J. Bar-Ilan, M. Levene, "The hw-rank: an h-index variant for ranking web pages", Scientometrics, Volume 102, Issue 3, Springer, 2015, pp. 2247-2253.
- [18] V. Jindal, S. Bawa, S. Batra, "A review of ranking approaches for semantic search on Web", Information Processing & Management, Volume 50, Issue 2, Elsevier, 2014, pp. 416-425.
- [19] M. Sirisha, S. R. Sanku, "Adapting a Ranking Model for Domain-Specific Search", International Journal of Scientific and Research Publications, Volume 3, Issue 11, 2013, pp. 1-5.
- [20] M. S. Aktas, M. A. Nacar, F. Menczer, "Personalizing PageRank Based on Domain Profiles", 10<sup>th</sup> ACM SIGKDD Conference, Seattle, Wasington, USA, 2004, pp. 83-90.
- [21] Maheswari, K. C. S. Reddy, S. V. A. Rao, "Design and Implementation of Ranking Adaption Algorithm for Domain Specific Search", International Journal of Computer Trends and Technologies (IJCTT), Volume 4, Issue 8, 2013, pp.2830-2833.
- [22] R. Shettar, R. Bhuptani, "A Vertical Search Engine – Based on Domain Classifier", International Journal of Computer Science and Security, Volume 2, Issue 4, 2008, pp. 18-27.
- [23] F. Lamberti, A. Sanna, C. Demartini, "A Relation-Based Page Rank Algorithm for Semantic Web Search Engines", IEEE Transactions on Knowledge and Data Engineering, Volume 21, Issue 1, 2008, pp. 123-136.
- [24] G. Kollias, E. Gallopoulos, A. Grama, "Surfing the Network for Ranking by Multidamping", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 9, 2014, pp. 2323-2336.
- [25] V. Chauhan, A. Jaiswal, J. Khan, "Web Page Ranking Using Machine Learning Approach", Fifth International Conference on Advanced Computing & Communication Technologies (ACCT), IEEE Conference Publications, 2015, Rohtak, India, pp. 575-580.
- [26] S. Gupta, N. Duhan, P. Bansal, J. Sidhu, "Page ranking algorithm in online digital libraries: A survey", 3<sup>rd</sup> International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), IEEE Conference Publications, 2014, Noida, India, pp. 1-6