iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VIII **Month of publication:** August 2017

DOI: http://doi.org/10.22214/ijraset.2017.8335

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue VIII, August 2017- Available at www.ijraset.com

A Hybrid Page Ranking Algorithm for Organic Search Results

M. Usha[1], Dr. N. Nagadeepa[2]
[1]*Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, Tamilnadu, India*
[2]*Principal, Karur Velalar College of Arts and Science for Women, Karur, Tamilnadu, India*

Abstract: ***Page Ranking techniques are in use by search engines to rank the web pages in order of their importance and relevance to user query. Many algorithms for page ranking either focus on link-structure of the page or similarity of the page with the given query to calculate Page Rank values of the pages. These conventional algorithms never consider user interaction time while page ranking. Generally, Users only look at the first page of the SERP. They do not go beyond the first page. By considering these factors along with the existing page ranking algorithms, a new hybrid page rank algorithm is proposed. This work combines, Content Similarity with synonym based approach, Link Structure and user interaction time (number of clicks and time spent on web page) of the page, to calculate page ranking of the page. Using proposed work, user can get the most relevant web pages in the top of the result, without going deeper into the organic search engine results and this work will improve the order of the URL listed by the search engine.***

Keywords: Page Rank, Organic Search Result, Synonym Based Approach, Content Mining, Structure Mining, Usage Mining.

I. INTRODUCTION

The World Wide Web (WWW) is trendy and interactive intermediary to telecast in turn these days. It is a massive, contrary diverse, dynamic and mostly formless data warehouse. As on today WWW is the prevalent information depository for awareness indication. The subsequent challenges in Web Mining are:

A. Web is vast.
B. Web pages are moderately ordered.
C. Web information stands to be multiplicity in meaning.
D. Degree of quality of the in sequence extracted.
E. Winding up of knowledge from information extracted.

It is predictable that WWW has persistent by about 2000% since its evolution and is replication in size every six to ten keywords with the catalog proceeds the URLs of the pages to the months. With the swift augmentation of WWW and the user's stipulate on knowledge, it is becoming more difficult to deal with the information on WWW and satisfy the user desires. Therefore, the users are in search of improved information repossession techniques and tools to position, extract, and filter and locate the essential information. Most of the users use information reclamation tools akin to search engines to find information from the WWW. There are tens and hundreds of search engines obtainable but some are popular like Google, Yahoo, Bing etc., because of their swarming and ranking methodologies. The search engines download, index and store up hundreds of millions of web pages. They response tens of millions of queries every day. So Web mining and ranking mechanism becomes very significant for effective information retrieval. [13]

Web mining is the process of using data mining methods and algorithms to mine information directly from the Web by extorting it from Web documents and services, Web content, hyperlinks and server logs. The goal of Web mining is to look for patterns in Web data by collecting and analyzing information in order to gain insight into trends, the industry and users in general. [11]

1) Categories of Web mining

a) Web content mining:

This is the process of mining useful information from the contents of Web pages and Web documents, which are mostly text, images and audio/video files. Techniques used in this discipline have been heavily drawn from natural language processing (NLP) and information retrieval. [11]

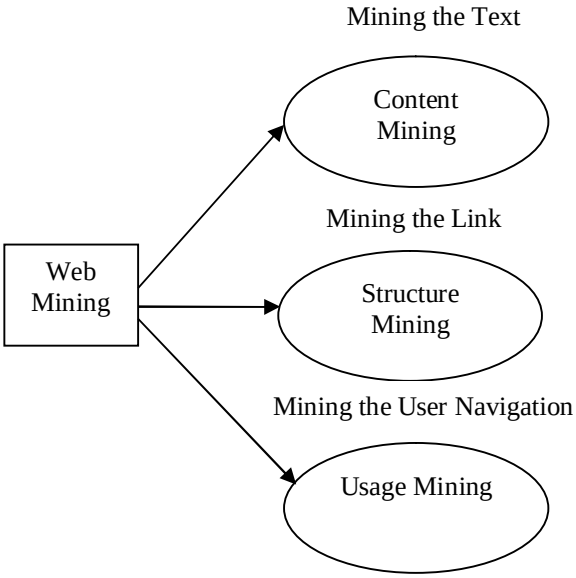


Fig.1 Web Mining Categories

b) *Web structure mining*: This is the process of analyzing the nodes and connection structure of a website through the use of graph theory. There are two things that can be obtained from this: the structure of a website in terms of how it is connected to other sites and the document structure of the website itself, as to how each page is connected. [11]

c) *Web usage mining*: This is the process of extracting patterns and information from server logs to gain insight on user activity including where the users are from, how many clicked what item on the site and the types of activities being done on the site. [11]

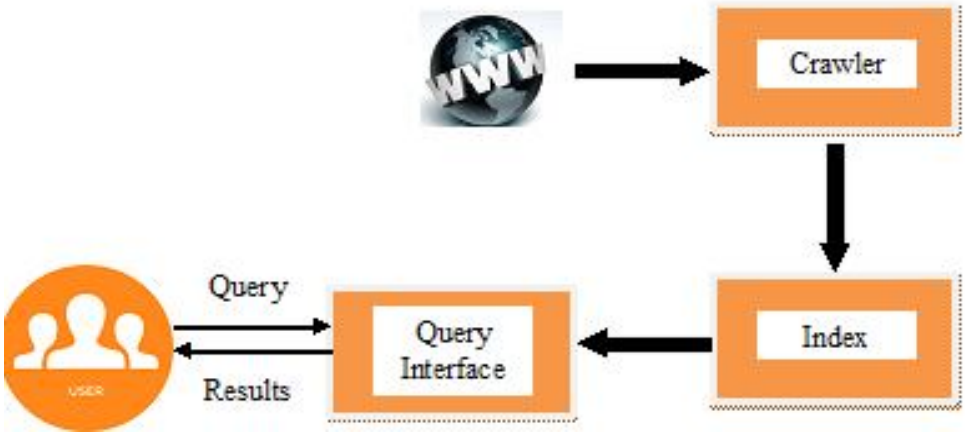


Fig. 2. Architecture of Search Engine

The architecture of a search engine consists of Crawler, Indexer and Ranking mechanism. The crawler traverses the web and downloads the web pages also known as spider or robot. The downloaded pages are then given to a indexing module which parses the web pages and constructs index based on the keywords in the web pages. An index is maintained using the keywords contained in the web pages.

User types a query using different keywords on the interface of a search engine provided after which the query processor component matches the query keywords with the index and giving back the URLs of the pages related to the user. The pages are first passed through a ranking mechanism done by the search engines that show the most relevant or high ranked pages at the top and less relevant/ low ranked pages at the bottom [9]

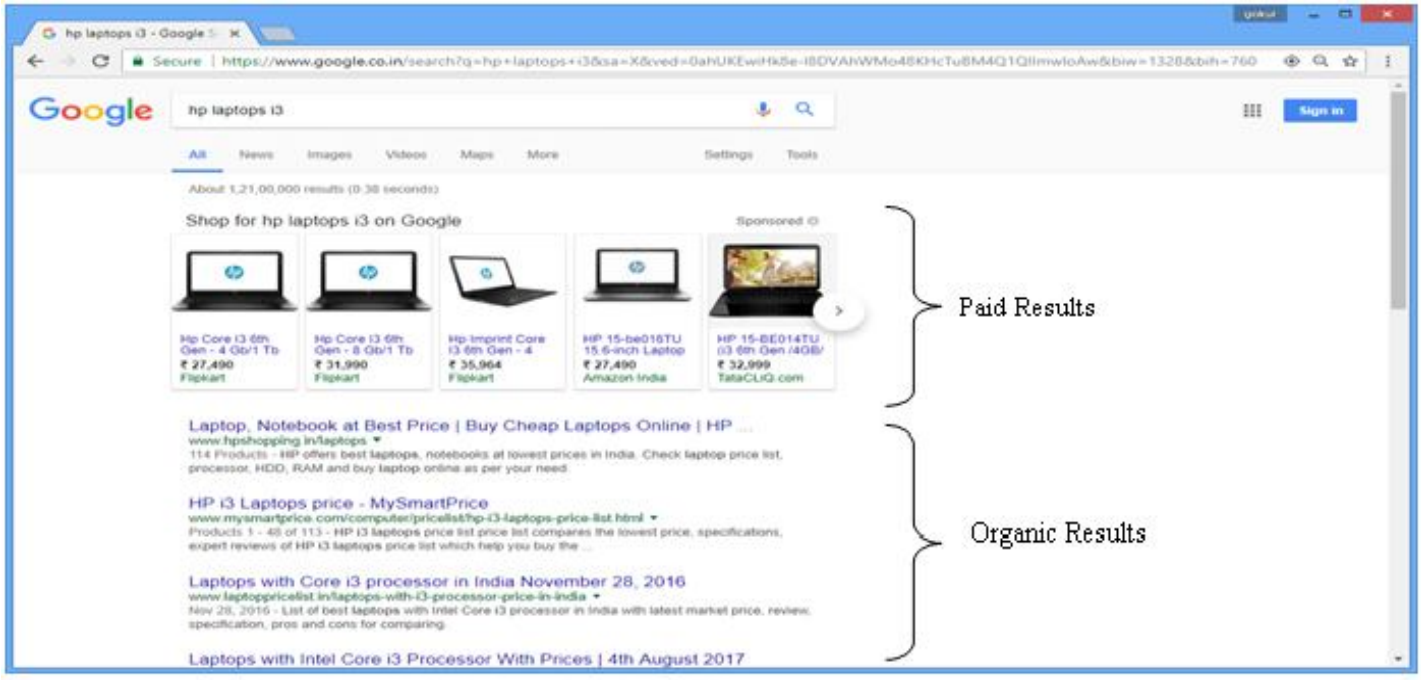


Fig.3 A Google's result page with Paid and Organic Results

After this process the pages result are shown to the user. When a search engine returns its search results, it gives two types: organic and paid. Organic search results are the Web page listings that most closely match the user's search query based on relevance. Also called "natural" search results, ranking high in the organic results is what SEO is all about.Paid results are basically advertisements — the Web site owners have paid to have their Web pages display for certain keywords, so these listings show up when someone runs a search query containing those keywords.

On a search results page, the user can tell paid results from organic ones because search engines set apart the paid listings, putting them above or to the right of the organic results, or giving them a shaded background, border lines, or other visual clues. Fig.3 shows the difference between paid listings and organic results. [12]. The typical Web user might not realize they're looking at apples and oranges when they get their search results. Knowing the difference enables a searcher to make a better informed decision about the relevancy of a result. Additionally, because the paid results are advertising, they may actually be more useful to a shopping searcher than a researcher (as search engines favor research results).

A PageRank is a mathematical algorithm based on the web-graph, created by all WWW pages as nodes and hyperlinks as edges. The rank value indicates importance of a particular page. The PageRank of a page is defined recursively and depends on the number and PageRank metric of all pages that link to it. But in this computation if new page is arises or added then the whole computation is required to perform again for accurate results. [14] To evaluate the PageRank for web a hybrid PageRank algorithm is proposed. A new kind of Page rank algorithm is developed which combines content, link and user navigation.

The rest of the paper is organized as follows. The Related work and Back Ground is covered in Section II. Section III covered the proposed page ranking algorithm. Section IV discusses the Experimental Set-up and Design information. Section V summarizes the results and Section VI includes conclusion.

II. LITERATURE SURVEY

A. Anamika Rajput and Sushil Kumar Chaturvedi

Proposed a new recommender system for user rating.A synonyms based ranking of the websites has been proposed. When the user searches for a keyword, that keyword is searched and also the related synonyms. Based on its search, the websites are displayed and user is asked to give rating. The user rating is used to upgrade or degrade the page rank in the database. The map reduce algorithm has been used to differentiate and reduce the data retrieved. The system created a new recommender system for page ranking up gradation in data mining field. [1]

B. Nagappan.V.K and Dr. P. Elango

Has devised a new algorithm called Agent based Weighted PageRank algorithm. This algorithm is intended at improving the order of the pages in the result list. [2]

C. Nandnee jain and Upendra Dwivedi

Have proposed an enhanced page ranking algorithm on the basis of response time of each web page. The idea is to compute access time of each web page and then compute average response time on the basis of which ranking of web pages can be predicted. A Comparative analysis is done on Google page ranking and proposed algorithm over various iterations and computational time and goodness factor in which the proposed work performs better as compared to the existing page ranking techniques.[3]

D. Lissa Rodrigues and Shree Jaswal

Presented an advanced approach which considers web content mining and web structure mining towards ranking of web pages to give relevancy of user's query. The weights of in links, outlinks of a page is calculated. Page level Keyword Counter is increased every time when a page is hit, whenever a particular web page contains the input keyword. The count output by page level keyword counter is keyword sum or the repetition of the keyword on a specific page. The score dependent on content as well as link structure.[4]

E. W. Xing and A. Ghorbani has described Weighted Page Rank algorithm (WPR) which works on inlinks and outlinks of web pages and gives rank score depending upon farne of the pages. Weight of inlink Win(V,U) and weight of outlink Wout(v,u) are passed down to record the farne ofthe webpages. The weights to the links are assigned on the basis of the popularity, thus popularity is more of that webpage which has the sum of number of inlinked pages and outlinked pages more. Surfer jamming and theme drift are drawbacks of weighted page rank algorithm [5]

F. R. Joshi, V. Kumar Gupta has devised hybrid algorithm depending on web content mining and web structure mining. Page rank value of web pages are calculated and analysis of some algorithms that makes use of link structure or web structure and some algorithms which uses web content mining was made. This paper compares the existing algorithm with the calculated webpages ' rank of proposed algorithm [6].

G. S. Goel and S. Yadav have expressed Page level keywords where keywords are found in specific pages from a website. Page level keywords are used to measure search engine results. The results retrieved from search engine contain large number of pages. Also search engines are classified on educational queries where page level keywords are used [7].

H. A. Jain.et.al have addressed various algorithms for link analysis such as Page Rank (PR), Hyperlink-Induced Topic Search (HITS), Weighted Page Rank (WPR) and CLEVER and also comparative study among these algorithms is carried out [8].

III. STATEMENT OF PROBLEM

As discussed in Section 2, due to the insufficiency prevalent in current Page Ranking algorithms the users are not able to acquire accurate and precise result with respect to their search query. They have to write related queries multiple times to get information that matches their interest and need. Link structure based techniques solely rely on link structure to rank pages. Which implies greater the links to the page larger the rank of page rank. Therefore the outcomes obtainedare autonomous of the keywords as specified in user's question.

However, different people have different interests, knowledge and choices therefore a single rank may not suit the requirements of an individual user. Also current search engines donot rank the web pages on the knowledge of past relevance of the page with respect to a particular query and on user response. Therefore due to lack of any preference, taxonomy and improvement in result-sets search engines are not able to provide accurate and precise information and users have to extract the most important information by cultivating result-sets manually. All the existing algorithms may provide satisfactory performance in some cases but many times the user may not get the relevant information.

The problem we all face when we search a topic in the web using a search engine like Google is that we are presented with millions of search results. It is not practically feasible to visit all these millions of web pages to find the required information.It consumes high data and time. When we visit few initial links shown in the search results, we may not get the relevant information. Therefore, we all feel the requirement of a mechanism so that we can get the relevant information according to the query posted by us.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue VIII, August 2017- Available at www.ijraset.com

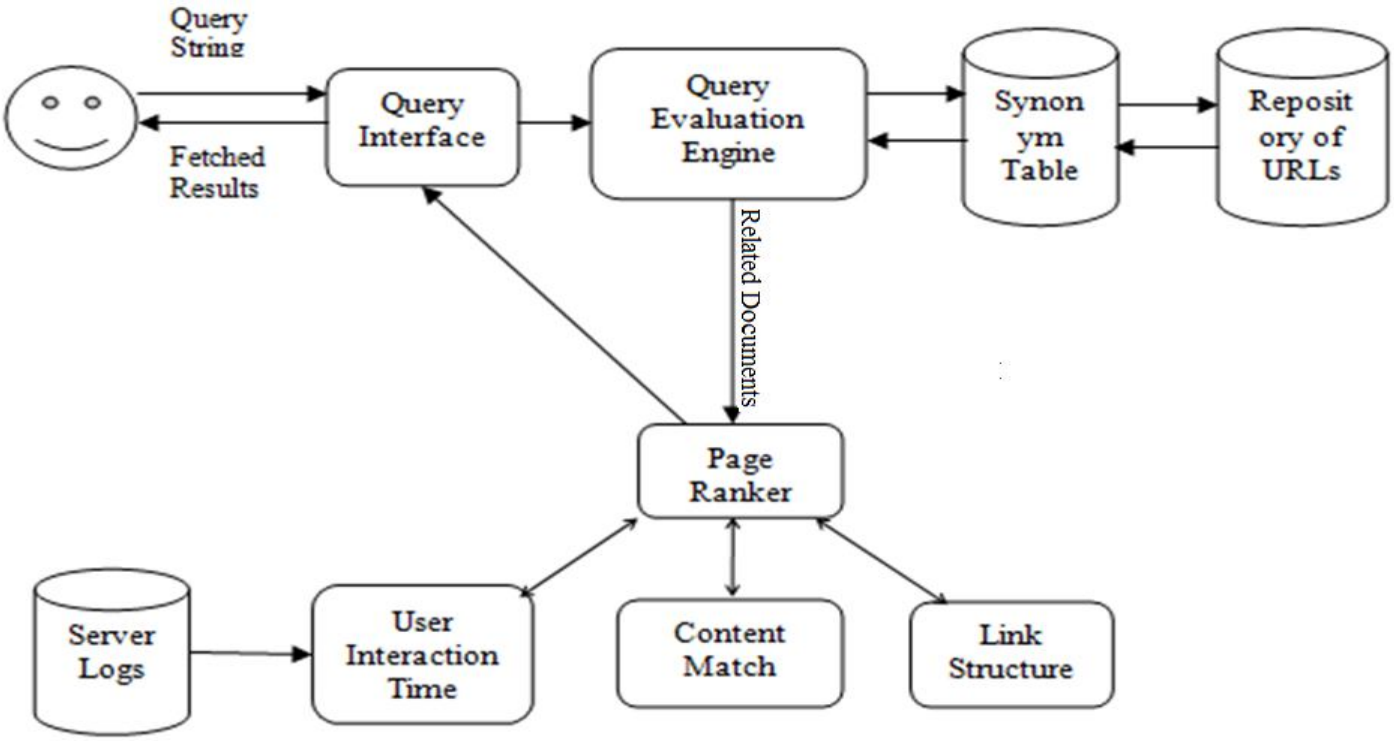
The major problem that we feel with all these algorithms is that none of them include the synonym based approach. There is a need for infering the meaning of the query and indexing should be based on that.

To overcome these deficiencies a hybrid Page ranking mechanism based on combination of web structure (link-structure of web page), web usage (past usage pattern of web page and user preference based on domain profile and type of page) and web content mining (how closely query term match the contents of web page) is being proposed that aims to provide best web pages in conjunction with user needs and interest.

IV. PROPOSED WORK

In this paper, a Hybrid web Page Rank Algorithm is proposed which is based on content, link structure and usage of the web pages and also it searches for the given keyword as well as for synonyms of the keyword. In order to optimize the results of a search engine by retrieving the more relevant pages on top of the search result list, we have to mine the text, hyperlinks and user navigation of the web pages. Through this proposed algorithm users can view their most relevant pages in the top of the result. This algorithm tries to improve the order of the pages displayed to the user.

In this system, when the user searches for a keyword, that keyword is searched and also the related synonyms will be searched.The architecture of the proposed algorithm is given in Fig 4:



A. *Proposed Methodology*

1) Collection of the Training Data Set
2) The training Data set consists of the URLs, Keywords & Domain of the web page.
3) Construct the Synonym table for each keyword in the training data set
4) Designing the User Interface to enter the search query string
5) Design query evaluation engine to parse the query and to fetch the matched documents from repository
6) Apply the proposed work on the retrieved URLs
7) Display the result
8) Update the server logs

The major modules of the proposed system are:

1) Create DataSet
2) User Query Interface
3) Query Evaluation Engine
4) Page Ranker
5) Server Logs

TABLE I
NOTATIONS USED

Notations Used	Description
UR j	User
QRY j	Query given by the user
LBR	Link Based Rank
CBR	Content Based Rank
AvgTime	Average User Interaction Time
kwd k	Keyword in the given query
PRV	Page Rank Value

A. Create DataSet Module

1) First dataset is created which contains webpage name (URL), Domain & keywords that relate to the web page.
2) Second dataset is created which contains the synonyms of each keyword in the first dataset. So, search will be based on the keyword given by the user as well as the synonyms of that word.

TABLE II
URL TABLE

URL	Domain	Keywords
http://www.w3schools.com/html/html_intro.asp	HTML	HTML Introduction, HTML Tags, HTML Page Structure
http://www.collegesinsouthindia.com	Institution	College, Institution, South India Colleges
https://thehomeschoolscientist.com	Science	Kids Projects, Assignments
http://www.kidspot.com.au	Science	Experiments, Activity, School

TABLE III
SYNONYM TABLE

Keywords	Synonyms
Technique	Method, Procedure, System, Process, Performance
College	Institution, Education, Learning, University, Academy
Projects	Experiments, Activities, Research

B. User Query Interface

An interface is provided to the user to enter the query and to display the result. This user interface contains, Textbox to enter the query string and a search button. User can select the domain of the query term to improve the search. If the User does not know about the domain , he no needs to select this.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue VIII, August 2017- Available at www.ijraset.com

C. Query Evaluation Engine Module

Once the user enters the query, the query evaluation engine first remove the stop words from the query string and retrieves the keywords or terms. It will extract the synonyms of the keyword from the synonym table. Then it makes search for the URL of the keyword and its synonyms as well in the first training data set. Next fetches the URLs that are related to the keyword and sends the fetched result to the Page Ranker.

D. Remove Stop Words

Commonly used words that are excluded from searches to help index and parse web pages faster. Some examples of **stop words** are: "a", "and", "but", "how", "or", and "what." For example, if you are to search for "What is a motherboard?", the engine will only look for the term "motherboard."

The algorithm is implemented as below given steps.

Step 1: The given query string is tokenized and individual words are stored in array.

Step 2: First word is read from the array.

Step 3: That word is compared to stop word list using sequential search technique.

Step 4: If it matches, the word in array is removed , and the comparison will be discontinued.

Step 5: After removal of stopword completely, next word is read from arrau and again algorithm follows step 3. The algorithm runs continuously until all the words in the array are compared.

Step 6: Resultant text free from stopwords is considered for stemming process.

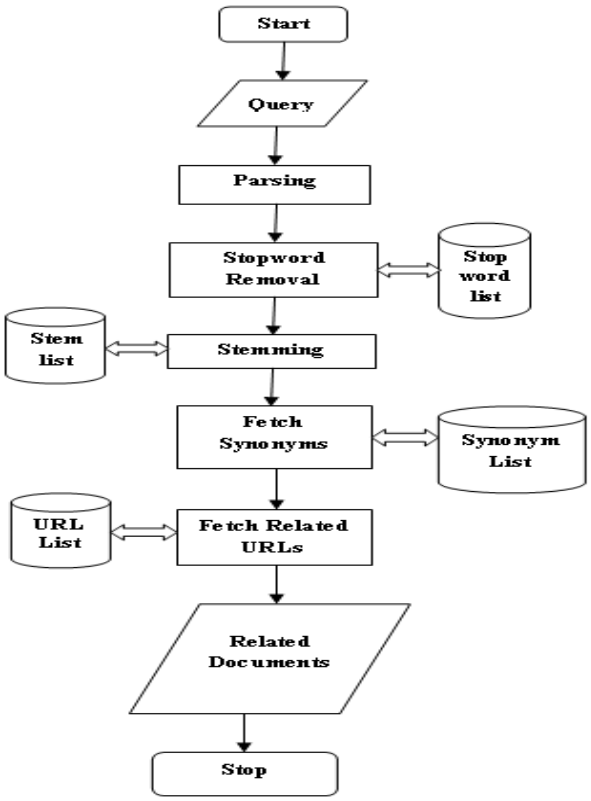


Fig 5. Flowchart of Query Evaluation Engine Task

E. Page Ranker Module

Page ranker has to order the fetched URLs based on the relevancy, link structure and user interaction time of the web page. After ranking the URLs, top k results will be sent to the user query interface.

F. Ranking the fetched URLs

To order the web pages retrieved from the repository, PageRankValue (PRV) is calculated by the PageRanker.

G. To calculate rank based on link structure

$$LBR(A) = (1-d) + d\,(PR(T1)/C(T1) + \ldots + PR(Tn)/C(Tn)) \quad \text{--- (1)}$$

where

PR(Ti) is the PageRank of pages Ti which link to page A,

C(Ti) is the number of outbound links on page Ti and is a damping factor which can be set between 0 and 1.

H. To calculate rank based on User Interaction Time

1) Open server log
2) Read average time spent and number of users
3) On client side initially set timer=0
4) Record the total time spent by the user using client side script
5) After receiving the time from client side, server side scripting will update the average time and number of users in the server log.

$$AvgTime = (AvgTime + \text{current timer value}) / (\text{Tot.No.of User}+1) \quad \text{----- (2)}$$

I. To calculate rank based on Content of the Page

Content Weight is based on how many terms in different web page fields match with the Query keywords. The content weight is calculated differently for different types of pages to give more weightage to page characterstics.

$$CBR = 0.4 * HeadText + 0.3*TitleText + 0.3*LinkText + 0.3*ParagraphText \quad \text{----- (3)}$$

where

HeadText= No.of Query keywords that appear in Web Page HTag / Total number of terms in HeadingText

TitleText =No.of Query keywords that appear in Title tag of web page/Total number of terms in Title tag

LinkText= No.of Query keywords that appear in Link tag of web page / Total number of terms in Link Text

ParagraphText= No.of Query keywords that appear in P tag of web page / Total number of terms in P

Final page rank value can be calculated by using,

$$PRV = 0.3*CBR*0.2*LBR*0.2*AvgTime \quad \text{----- (4)}$$

J. Server Logs Module

It is used to store the average time spent on each web page by the users. It will be updated every time the user visits the page.

TABLE IV

AVG. TIME SPENT BY USERS

URL	Keywords	Avg. Time Spent (in hrs)	Total No.of Users
http://www.w3schools.com/html/html_intro.asp	HTML Introduction, HTML Tags, HTML Page Structure	10.5	13
http://www.collegesinsouthindia.com	College, Institution, South India Colleges	5.4	15
https://thehomeschoolscientist.com	Kids Projects, Assignments	12.55	23
http://www.kidspot.com.au	Experiments, Activity, School	7.3	17

V. EXPERIMENTAL SET-UP & DESIGN

Implementation of the proposed work is done in ASP.Net as the Front end development tool and SQLServer as a Back end database management system. Popularity of a web page is calculated by considering its linked structure. The whole page is parsed to extract the Links of the page. The extracted Links are then stored in the suitable table in the database. When a web page is accessed a script is loaded on the client side from web server. Script is used to check for the click event to occur. When a click event occurs, a message is send to web server with information of current web page and hyperlink. On server side a data base of log file is used to store the web page id, hyperlinks of that page and number of clicks on hyperlinks. Hit count value is increased by one every time a hit occurs on the hyperlink. The database or log files will accessed by crawler at the time of crawling. This hit count information is stored in search engine's database and is used to calculate the rank value of different web pages or documents.

In order to test the page ranking module a small user study was conducted to compare the performance of proposed page ranking mechanism (based on Page Structure, PageContent and Page usage mining) and Google ranking (based on Page Structure). The database consisted of a collection of 100 web pages related students science projects and a total of 15 queries in were selected for the study. After submitting the query the top n results were presented of both the ranking methods on the single screen separated by frames. The user interface was designed to be easy to use and less prone to evaluation errors.

Query1: projects for kids

Fig 5 refers the Google's search result page for the given query. After the user submits the query, it displays the web pages which are only related to the given keyword project. The result page displays the URLs which have the keywords projects and kids.

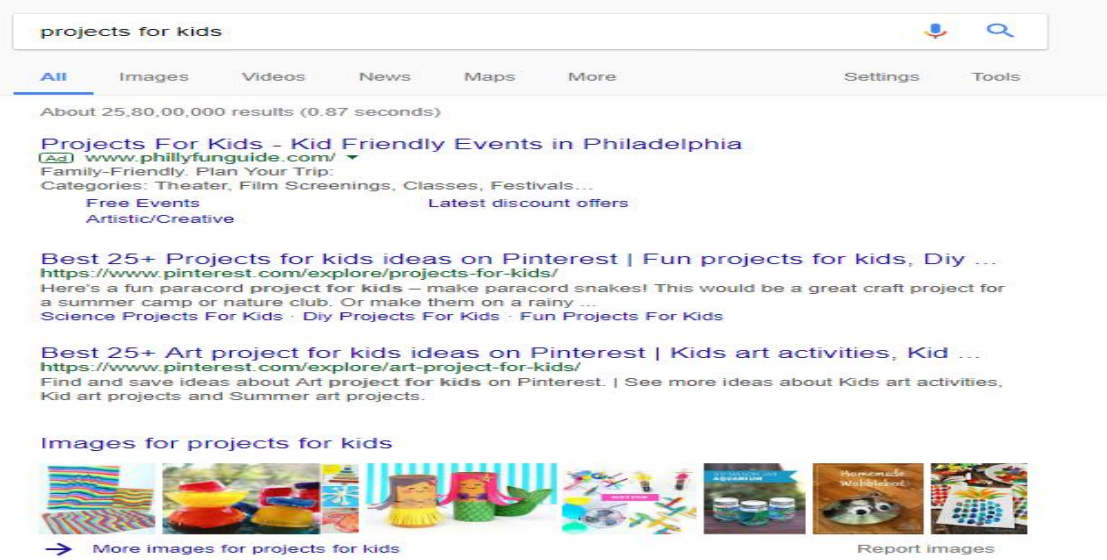


Fig.5. Google's Result Page for Query1

Query2: experiments for kids

Fig 6 refers the Google's search result page for the given query. After the user submits the query, it displays the web pages which are only related to the given keyword experiment. The result page displays the URLs which have the keywords experiments and kids.

When the user gives the keyword project he/she can only view the links related to the keyword 'project'. Similarly when the user gives the keyword experiment he/she can only view the links related to the keyword 'experiment'. But both the URLs in Fig.5 and Fig.6 contain relevant information and paid results are also displayed. The proposed system deals with this problem effectively by using the synonym based approach. When the user gives the keyword either project or experiment, the system will search for the both the keywords. Because it uses synonym based approach. So this system will fetch the web pages which are related to the given keyword and its synonym as well. In addition to the synonym based approach, the proposed system uses the efficient ranking technique which is based on the text, link and usage of a web page. Since it combines all these features in the algorithm, it displays the most relevant web pages in top k results and the order of the results is also improved.

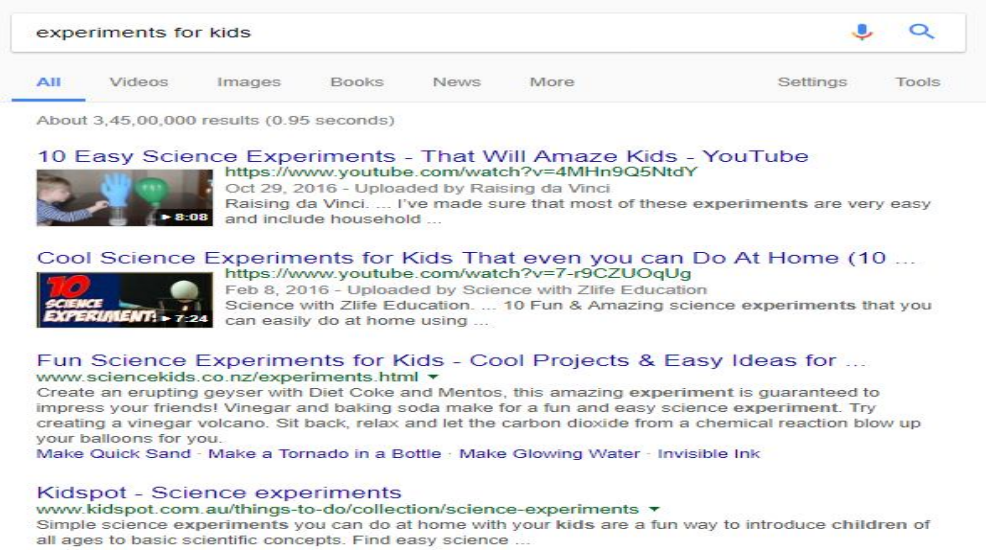


Fig.6. Google's Result Page for Query2

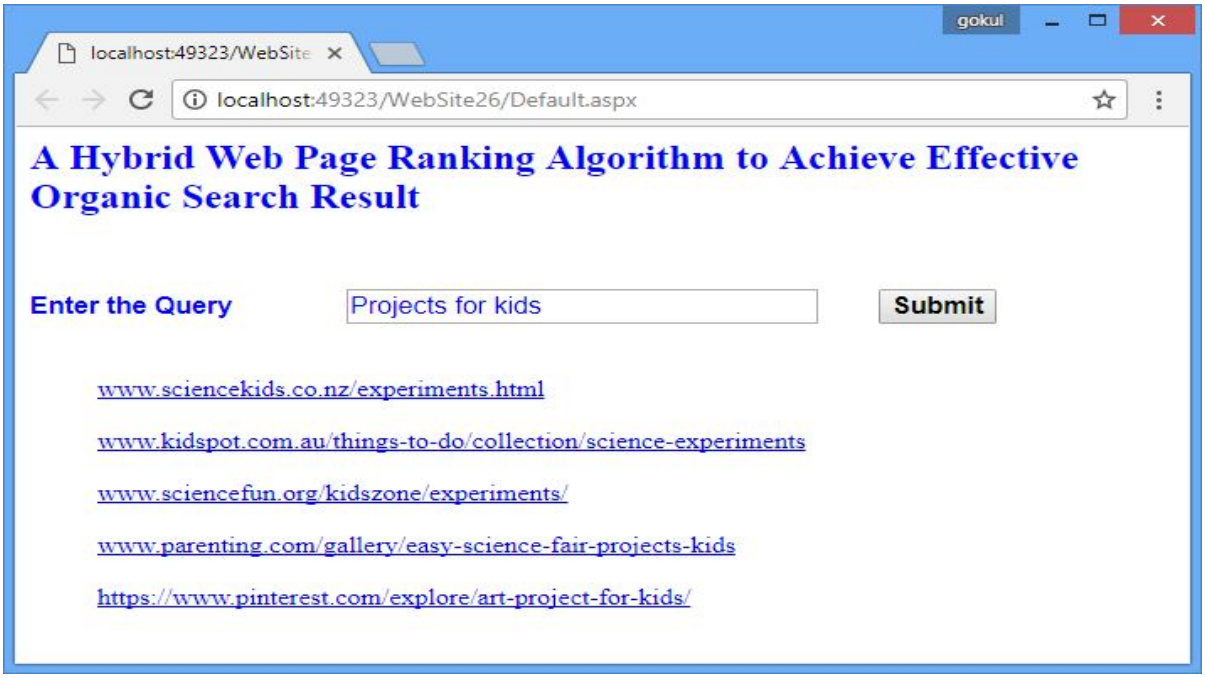


Fig.7. Result Page of the Proposed System for Query1

VI. RESULTS

A total of 15 queries were selected for the study and they were given in both the Google's search engine and the proposed system. For each of the query two result set were obtained and taken for further analysis. URLs in each of the result set marked as "more relevant", "less relevant" and "irrelevant" depending upon whether the selected web page closely matched the query and their interest or closely related to the query and interest or consisted of only series of links and not at all related to the query and interest. A score value of 1, 0.5 and 0 was then assigned to pages marked as more relevant, less relevant, and irrelevant respectively. These scores enabled the calculation of the precision of the ranking mechanisms for each of the search queries by using formula 4.

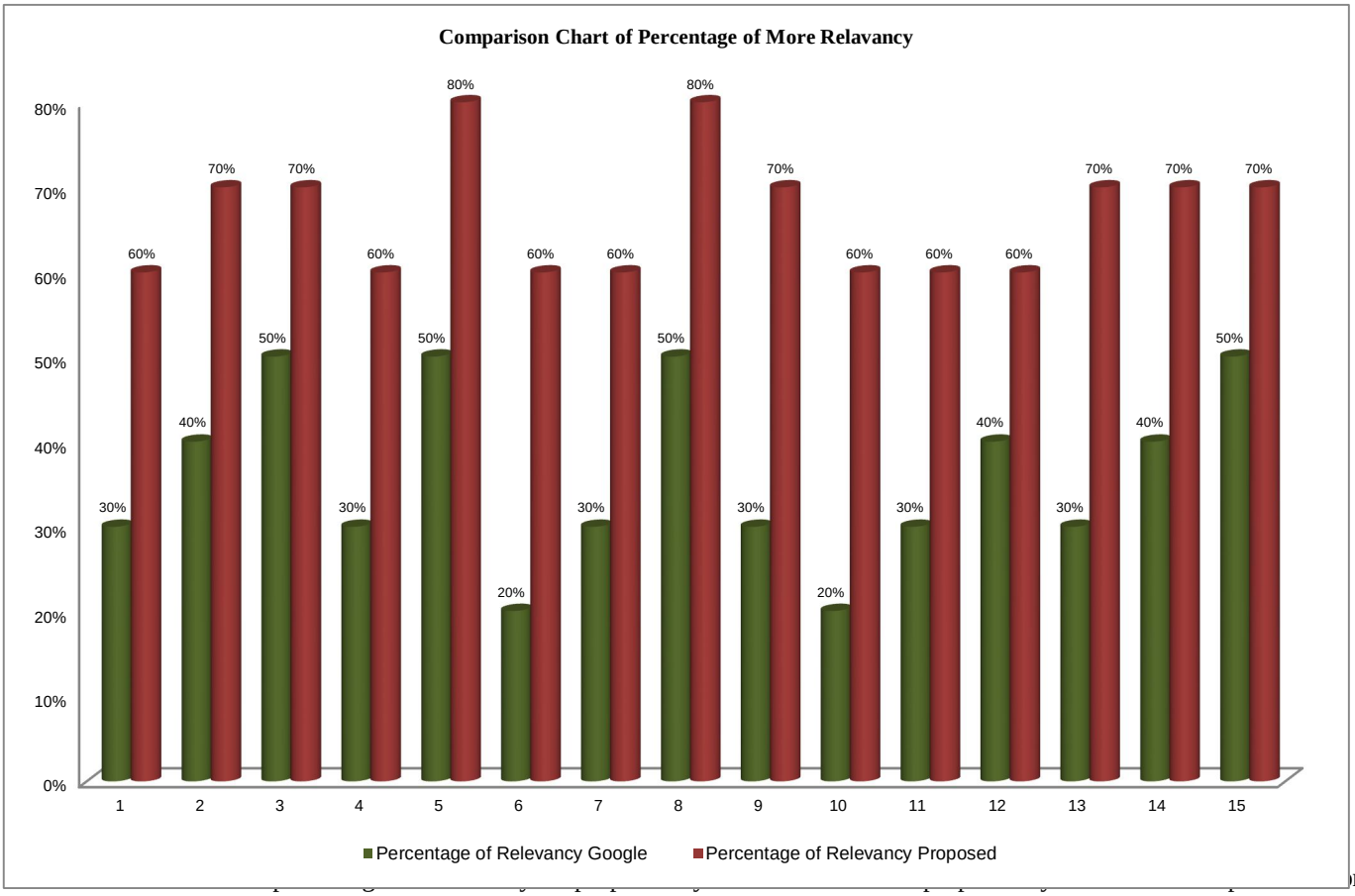
$$\text{Percentage of Relevancy} = \frac{\text{No.of relevant Pages}}{\text{Total No. of Pages}} \times 100$$

TABLE V.

COMPARISON OF PPERCENTAGE OF RELEVANCY FOR PROPOSED PAGE RANKING METHOD WITH GOOGLE PAGE RANKING

Query	More Relevant Pages		Less Relevant Pages		Irrelevant Pages		Percentage of Relevancy	
	G	P	G	P	G	P	G	P
1	3	6	4	4	3	0	30%	60%
2	4	7	2	3	4	0	40%	70%
3	5	7	3	3	2	0	50%	70%
4	3	6	4	4	3	0	30%	60%
5	5	8	2	2	3	0	50%	80%
6	2	6	4	4	4	0	20%	60%
7	3	6	3	4	4	0	30%	60%
8	5	8	2	2	3	0	50%	80%
9	3	7	4	3	3	0	30%	70%
10	4	6	4	4	2	0	20%	60%
11	3	6	4	4	3	0	30%	60%
12	4	6	4	4	2	0	40%	60%
13	3	7	3	3	4	0	30%	70%
14	4	7	3	3	3	0	40%	70%
15	5	7	3	3	2	0	50%	70%

Table V the percentage of relevant pages calculated for top N results retuned by Google Page Ranking method and proposed ranking method for different set of queries.



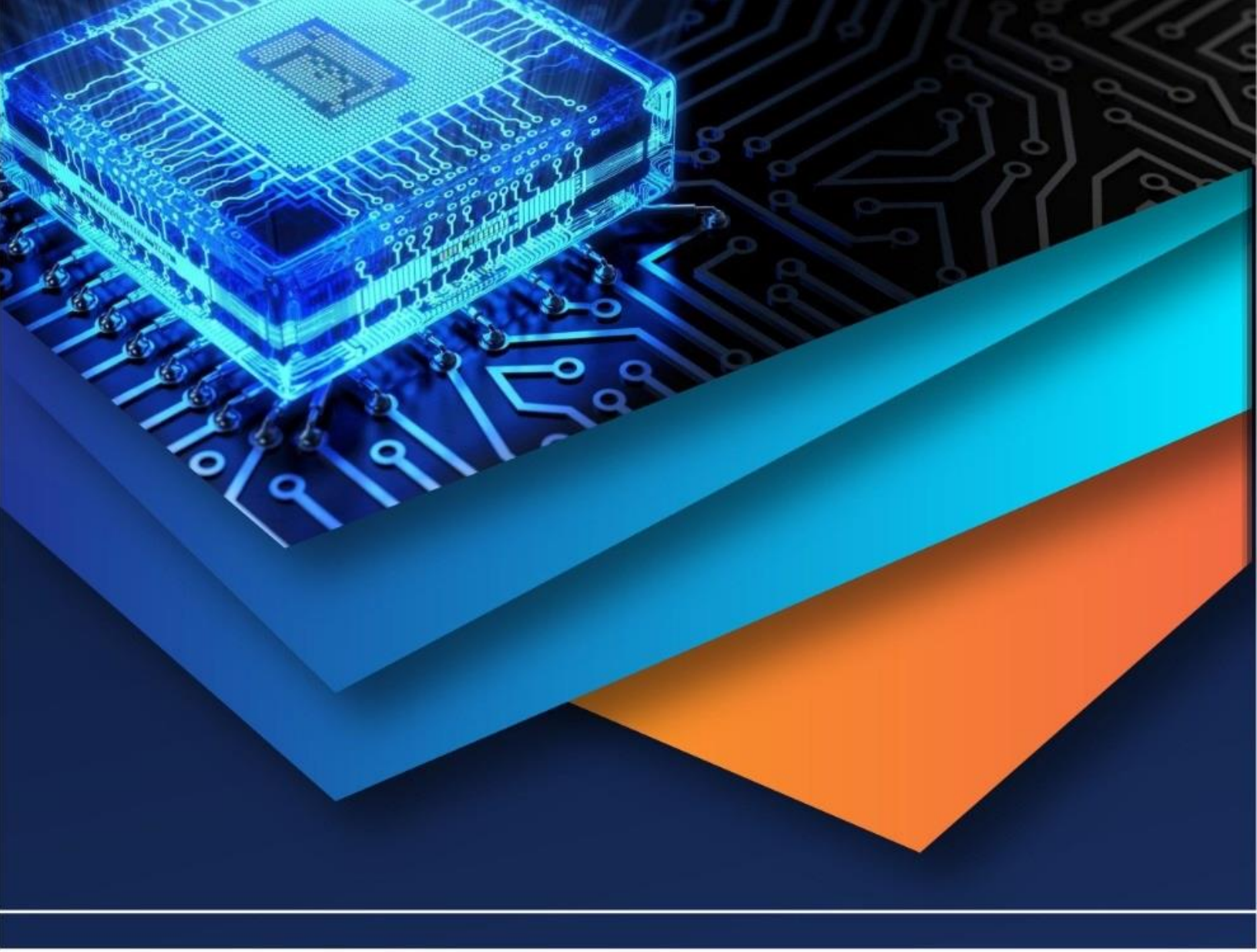
re relevant web pages with respect to the user query. These preliminary results although based on small user study are promising and provide us with a mechanism that can provide quality search results based on user preferences.

VII. CONCLUSION

As many Internet users do not use query syntax language for retrieving results from a search engine, they have to write multiple queries to acquire the information that meets their need and interest. Also due to weaknesses prevalent in page ranking algorithms some important pages may not lie in relative higher rank position. The page rank mechanism proposed in this paper not only considers link structure and page content of the web page but also considers user interaction time for page ranking. The advantage of the proposed mechanism is that user gets the full information within the first few URLs and will not have to go deeper into the search results returned by the search engine. While these preliminary results are not highly significant statically given very small user study, but they are promising. The proposed system seems to provide us with a mechanism that can help retrieve high quality documents with maximized user satisfaction.

REFERENCES

[1] Anamika Rajput ; Sushil Kumar Chaturvedi ,User Rating and Synonyms Based Modified Ranking Technique for Recommender Systems,inInternational Conference on Computational Intelligence and Communication Networks, DOI 10.1109/CICN.2015.196
[2] V.K Nagappan and Dr. P. Elango, “Agent Based Weighted Page Ranking Algorithm for Web Content Information Retrieval”, in IEEE International Conference on Computing and Communications Technologies (ICCCT’15), doi: 78-1-4799-7623-2/15, 2015
[3] Nandnee Jain ; Upendra Dwivedi, Ranking web pages based on user interaction time, in IEEE International Conference on Advances in Computer Engineering and Applications DOI: 10.1109/ICACEA.2015.7164709
[4] Lissa Rodrigues ; Shree Jaswal, An Efficient Page Ranking Approach Based on Hybrid Model, in Advances in Computing and Communication Engineering (ICACCE), 2015 Second International Conference on Advances in Computing and Communication Engineering DOI: 10.1109/ICACCE.2015.57
[5] W.Xing and A.Gorbani, "Weighted Page Rank Aigorithm", Proceedings Of the Second Annual Conference on Communication Networks and Services Research, May 2004, pp.305-314.
[6] S. Goel, S. Yadav, "Search engine evaluation based on page level keyword", Proceedings ofthe 2013 3rd IEEE International AdvanceComputing Conference, IEEE Computer Society (2013), pp. 870- 876.
[7] Ashish Jain, Rajeev Sharma, Gireesh Dixit, Varsha Tomar, "PageRanking A1gorithms in Web Mining, Limitations of Existing methods and a New Method for Indexing Web Pages", 2013 International Conference on Communication Systems and Network Technologies.
[8] Rutusha Joshi , Vinit Kumar Gupta, "Tmproving Pagerank Calculation by using Content Weight",International Journal of Science and Research (lJSR), June 2014
[9] Laxmi Choudharyand Bhawani Shankar Burdak “Role of Ranking Algorithms for Information Retrieval”, 2010
[10] Nikhil Sanyog Choudhary, Himanshu Yadav and Anurag Jain “Message Filtering Techniques in Social Networks over Web Environment- A Survey”, IJETAE, 2014.
[11] https://www.techopedia.com/definition /15634/ web-mining
[12] http://www.dummies.com/web-design-development/ search-engine-optimization/ the-difference-between-organic-and-paid-search-engine - results/
[13] Neelam Tyagi, Simple Sharma, “Comparative study of various Page Ranking Algorithms in Web Structure Mining (WSM), International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-1, June 2012 14
[14] Lili Yan, Yingbin Wei, ZhanjiGui, Yizhuo Chen, “Research on PageRank and Hyperlink-Induced Topic Search in Web Structure Mining”, 978-1-4244-7255-0/11/$26.00 ©2011 IEEE



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY