

Relative study of different Page Ranking Algorithm

Huma Mehtab
 Computer Science and Engg.
 Integral University, Lucknow, India
 huma.mehtab@gmail.com

Shahid Siddiqui
 Computer Science and Engg.
 Integral University, Lucknow, India
 siddiquisahil@gmail.com

Abstract—Search engines play a very important role in web, to retrieve the relevant documents among the large variety of sites. However, it retrieves a lot of variety of documents, that are all relevant to your search topics. To retrieve the foremost purposeful documents associated with search topics, ranking rule is employed in data retrieval technique. one amongst the problems in data mining is ranking the retrieved document. In information retrieval the ranking is one amongst the sensible issues. This paper includes numerous Page Ranking algorithms, page segmentation algorithms and compares those algorithms used for data Retrieval. various Page Rank based mostly algorithms like Page Rank (PR), Weighted Page Rank (WPR), Weight Page Content Rank (WPCR), hyperlink induced Topic choice (HITS), Distance Rank, Eigen Rumor, Distance Rank Time Rank, Tag Rank, relational based mostly Page Rank and question Dependent Ranking algorithms are mentioned and compared.

Keywords—Information Retrieval, web page Ranking, search engine, web mining, page segmentations.

I. INTRODUCTION

DATA mining is to extract or mine information from lots of knowledge known as information Discovery in Databases (KDD), that is that the results of data technology natural that is that the results of information technology natural evolution. In recent years, the information mining technology made great attention among the data trade, that is developing quickly. data processing is AN inter-discipline subject, influenced by multiple disciplines, as well as database system, statistics, machine learning, knowledge analysis, etc. At present, in line with the various sorts of mining information and mining the various objects, several data mining methods and special tools are available. several analysis fields like database, data analysis, machine learning, conjointly benefited lots from the data mining. data Retrieval could be a technique employed in data mining for looking in huge databases to retrieve connected documents. information Retrieval (IR) is that the science of sorting out information among relational databases, documents, text, multimedia system files, and therefore the World Wide web. several users are affianced within the IR field particularly reference librarians, governmental agents, skilled researchers, political analysts, and market forecasters. applications of IR are various however not restricted to extraction of knowledge from huge documents, spam filtering, probing in digital libraries, data filtering, object extraction from pictures, automatic summarization, document

classification and clustering, and web searching. Google's PageRank algorithm is that the one amongst the known algorithms in web search. With the increasing variety of sites [10] and users on the online, the quantity of queries submitted to the search engines are growing quickly day by day. Therefore, the search engines has to be a lot of economical in its process way and its output. web mining techniques are utilized by the search engines to extract applicable documents from the online database documents and supply the necessary and required information to the manipulators. The search engines become terribly successful and well-liked if they use efficient ranking mechanisms. currently recently it's very successful because of its PageRank algorithm. Page ranking algorithms [6] are utilized by the search engines to present the search results by considering the importance, reputation, and content score and web mining techniques to organize them in line with the user interest. Some ranking algorithms rely solely on the link structure of the documents i.e. their quality scores (web structure mining), whereas others search for the particular content within the documents (web content mining), whereas some use a permutation of each i.e. they use content of the document similarly because the link structure to assign a rank worth for an explicit document [5]. If the search results don't seem to be displayed in line with the user interest then the search engine can set its fame. that the ranking algorithms become vital. The sample design of a search engine is shown in Fig. 1.

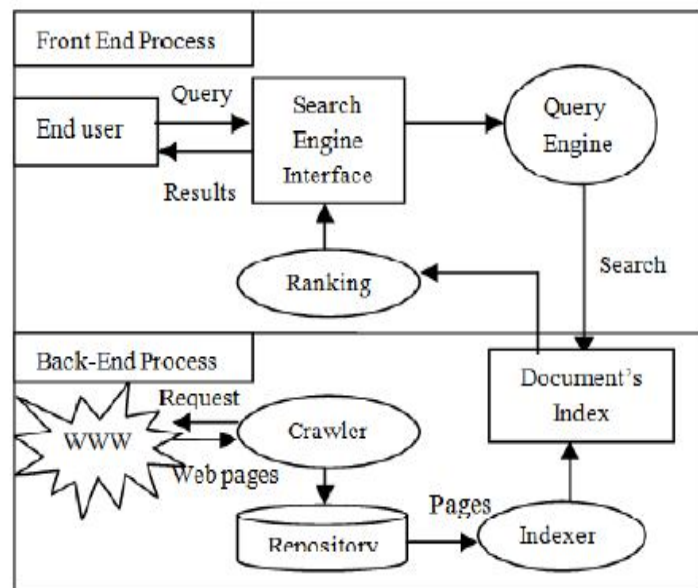


Fig. 1 Architecture of search engine



There are a unit 3 very important elements in a very program called Crawler, skilled worker and Ranking mechanism. The Crawler is additionally referred to as as a automaton or spider that navigates the net and downloads the net pages. The downloaded pages area unit being transferred to Associate in Nursing assortment module that parses the net pages and erect the index supported the keywords in individual pages. Associate in Nursing alphabetical index is generally sustaining victimization the keywords. once a question is being drifted by a user, it means that the question transferred in terms of keywords on the interface of an enquiry engine, the question mainframe section examine the question keywords with the index and precedes the URLs of the pages to the user. however before presenting the pages to the shopper, a ranking mechanism is completed by the search engines to gift the foremost relevant pages at the highest and fewer vital ones at the substructure. It makes the search outcomes routing easier for the user.

2. RANKING TECHNIQUE

Web mining [1] is that the mechanism to classify the net pages and net users by taking into thought the contents of the page and behavior of net user within the past. Associate in Nursing application data} mining technique could be a net mining that is employed impromptu to search out and retrieve information from the globe Wide net (WWW). per analysis targets, net mining [13] is formed of 3 basic branches i.e. web page mining (WCM), net structure mining (WSM) and net usage mining (WUM).

A. Web Page Mining (WCM)

Web Content Mining [18] is that the progression of extracting helpful data from the contents of net credentials. the net credentials might consists of text, audio, video, image or structured records like tables and lists. Mining are often purposeful on the net documents similarly the results pages intentional from an enquiry engine. There area unit atomic number 83 approaches in content mining referred to as agent primarily based} approach and information based approach. The agent based mostly technique concentrate on looking out applicable data victimization the distinctiveness of a specific domain to interpret and organize the collected data. The information approach is employed for go back to the semi-structure knowledge from the net.

B. Web Usage Mining (WUM)

Web Usage Mining is that the technique of shipping out helpful data from the secondary knowledge sequent from the interactions of the user whereas surfing on the net. It extracts knowledge accumulated in server referrer logs, access logs, agent logs, user profile and Meta knowledge client-side cookies.

C. Web Structure Mining (WSM)

The aim of the net Structure Mining is to get the structural abstract concerning the net web site and web content. It tries to work out the link structure of the hyperlinks at the bury

document level. Basic underpinning on the topology of the hyperlinks, net Structure mining [8] can classify the net pages and spawn the knowledge like similarity and relationship between totally different internet sites. this kind of mining are often administrated at the document level (intra-page) or at the link level (inter-page). it's necessary to understand the net system for data Retrieval. The 3 classes of net mining delineate and its own appliance areas as well as web site improvement, business intelligence, net personalization, web site modification, usage characterization and ranking of pages, classification etc.

3. PAGE RANKING ALGORITHMS

The page ranking algorithms [38] area unit usually employed by search engines to search out additional necessary pages. totally different Page Rank based mostly algorithms [11] like Page Rank (PR), Weighted Page Rank (WPR), Weight Page Content Rank (WPCR), link elicited Topic choice (HITS), Distance Rank, Manfred Eigen Rumor, Distance Rank Time Rank, Tag Rank, relative based mostly Page Rank and question Dependent Ranking algorithms.

A. Page Ranking Algorithms

In the ranking algorithms [4] the usage of net is drastically will increase day by day. The program is incredibly helpful to retrieve the relevant documents from net simply. The ranking algorithms [12] area unit important as a result of the search results aren't per their user wants then the program loss their quality. Google is that the famed program tool in page ranking formula. Some ranking algorithms rely solely on the recognition score i.e. net structure mining and web page mining. The PageRank values area unit calculated supported the quantity of pages that time to a page. adaptational strategies for the Computation of PageRank formula [2], [10] is employed to hurry up the computation of PageRank is sort of half-hour. Filter-Based adaptational PageRank and changed adaptational PageRank formula is employed to reducing the redundant computation. The page rank take into account solely back link to come to a decision the page score. within the below equation the variable d could be a damping issue [5] values are often set between zero and one. $PR(A)$ is that the PageRank of page A, T_1, \dots, T_n is all pages that link to page A, $PR(T_i)$ is that the PageRank of page T_i , $Q(T_i)$ is that the range of pages to that T_i links to $PR(T_i)/Q(T_i)$ is PageRank of T_i distributing to any or all pages that T_i links to, $(1-d)$ is to create up for a few pages that don't have any out-links to avoid losing some page ranks [40]. $PR(A)$ is the incoming link to page A and $C(T_1)$ is that the outgoing link from page $PR(T_1)$. The PageRank of a page A is given below

$$PR(A) = (1 - d) + (PR(T_1)/C(T_1) + \dots + PR(T_n)/C(T_n)) \quad (1)$$

An example of back link is shown in Fig. 2, N is the back link of M & Q and M & Q are the back links of O.

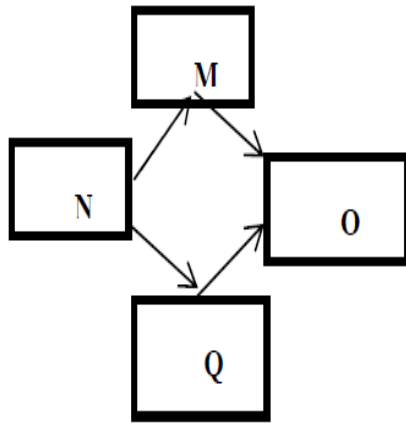


Fig. 2 Back links Example

B. Implementation of Page Rank Algorithm

The following steps explain the method for implementing Page Rank Algorithm [7]:

Step 1. Initialize the rank value of each page by $1/n$. where n is total no. of pages to be ranked. Suppose we represent these n pages by an Array of n elements. Then $A[i] = 1/n$ where $0 \leq i < n$

Step 2. Take some value of damping factor such that $0 < d < 1$. e.g. 0.15, 0.85 etc.

Step 3. Repeat for each node i such that $0 \leq i < n$. Let PR be an Array of n element which represent PageRank for each web page.

$PR[i] \leftarrow 1-d$

For all pages Q such that Q Links to PR[i] do

$PR[i] \leftarrow PR[i] + d * A[Q]/Q_n$

where $Q_n =$ no. of outgoing edges of Q

Step 4. Update the values of A $A[i]= PR[i]$ for $0 \leq i < n$

Repeat from step 3 until the rank value converges i.e. values of two consecutive iterations match.

The advantages of page rank are less query time, Less susceptibility to localized links, more efficiency and feasibility.

C. Weighted Page Rank

Weighted Page Rank algorithm (WPR) [14] is the extension of Page Rank algorithm. In the WPR contains both in link and out link, the link is assigned based upon the page rank priority.

$$W^{in}(m, n) = \frac{I_n}{\sum_{P \in R(m)} I_P} \tag{2}$$

where I_n and I_p represent the number of in links of pages n and page p , respectively. $R(m)$ denotes the reference page list of page m . The weight of link (m,n) calculated based on the number of out links of page n and the number of out links of all reference page of page m .

$$W^{out}(m, n) = \frac{O_n}{\sum_{P \in R(m)} O_n} \tag{3}$$

where O_n and O_p represent the number of outlinks of the page n and page p , respectively. $R(m)$ denotes the reference page list of page m . Modification of page rank formula is given:

$$WPR(n) = (1 - d) + d \sum_{m \in B} WPR(m) w^{in}(m, n) w^{out}(m, n) \tag{4}$$

The values of $WPR(A)$, $WPR(B)$, $WPR(C)$ and $WPR(D)$ are shown in equations respectively. The relation between these are $WPR(A) > WPR(B) > WPR(D) > WPR(C)$.

This results shows that the Weighted PageRank order is different from PageRank.

D. Weighted Page Content Rank

Weighted page content algorithm (WPCA) [46] is the modification of original page rank algorithm. It is used to give the sorted order to the web page. WPCR [35] is assigned numerical value based on which the web pages are given an order. This algorithm employs both web structure mining and web content mining techniques. Web structure mining is used to analyse the popularity of the page and web content mining is used to find the page relevancy. The calculation is based on the in links and out links of the page [47]. For example Google Web search receive 34,000 queries per second (2 million per minute; 121 million per hour; 3 billion per day; 88 billion per month) for most queries, there exist thousands of documents containing some or all of the terms in the query.

Algorithm WPCR

Input:

Query text Q

Set of pages $\{P_i\}$ Google (Q)

Output:

New $\{P_i\}$

Relevance calculation

Find $f(P_i) = \{$ number of frequency of logical combination of Q $\}$

Find content weight factor $CWF(P_i) = GPA(f(P_i))$

Reorder and return the new $\{P_i\}$

The proposed algorithm SWPCR design new methods to calculate the relevance of a page based on two factors:

- 1) Find $f(P_i) =$ the frequency of logical combination of query text, the number of times that term appears in page P_i .
- 2) Find content weight factor $CWF(P_i) = GPA(f(P_i))$ that is consider the core of SWPCR proposed algorithm based on: Given a matrix with $m \times n$; $n =$ number of words in a given query, each column contains the frequency of n words $f(n)$ in each of the given pages; $m =$ number of pages

E. HITS Algorithm

HITS (Hyper-link Induced Topic Search) algorithm [1] is used to ranks the web page by processing in links and out links. In this algorithm [9] a web page is named as authority and hub, if the web page is pointed by many hyperlinks it is named as authority, and if the page is pointed to various hyperlinks and a web page is named as HUB. Hubs are the pages that act as resource lists. Authorities are the pages having main contents.

A good hub page is a page which is pointing to many authoritative pages on that content and a good authority page is a page which is pointed by many good hub pages on the same content. A page may be a good hub and a good authority at the same time. It uses an iterative algorithm for computing the hub and authority weights. The HITS algorithm [49] gives WWW as directed graph $G(V,E)$, where V is a set of vertices representing pages and E is set of edges corresponds to link. It has two steps first one is sampling step and second one is iterative step. Sampling Step a set of relevant pages for the given query are collected, iterative step Hubs and Authorities are found using the output of sampling step. Following expressions are used to calculate the weight of Hub (H_p) and the weight of Authority (A_p).

$$HUB(H_p) = \sum_{q \in I_p} A_q$$

$$AUTHORITY(A_p) = \sum_{q \in B_p} H_q$$

Here Hub Score of a page is (H_q) and authority score of page is (A_q). $I(p)$ is set of reference pages of page p and $B(p)$ is set of referrer pages of page p . The weight of authority pages is proportional to the weights of hub pages that link to the authority page. Another one is, hub weight of the page is proportional to the weights of authority pages that hub links.

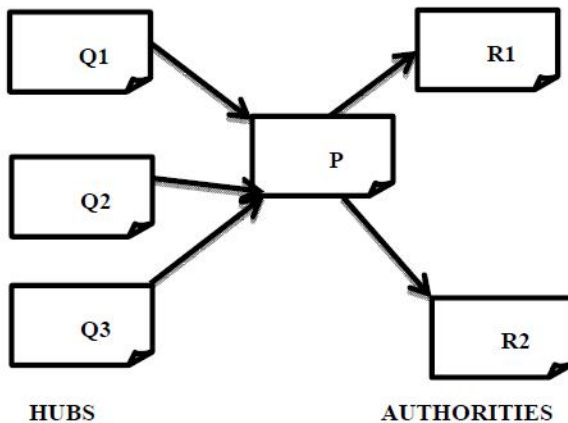


Fig. 3 Hubs and authorities

Hubs and authorities are calculated by way of:

Advantage of HITS:

HITS [11] scores due to its ability to rank pages according to the query string, resulting in relevant authority and hub pages. The ranking may also be combined with other information retrieval based rankings. HITS [9] is sensitive to user query (as compared to PageRank). Important pages are obtained on basis of calculated authority and hubs value. HITS is a general algorithm for calculating authority and hubs in order to rank the retrieved data. HITS induces Web graph by finding set of pages with a search on a given query string. Results demonstrate that HITS calculates authority nodes and hubness correctly.

F. Distance Rank Algorithm

A distance rank algorithm [2] is proposed by Ali Mohammad Zareh Bidoki and Nasser Yazdani. This algorithm is also called as intelligent ranking algorithm based on support learning algorithm. In this algorithm the page is considered as a distance factor and rank the distance between web pages in search engine. The main goal of the ranking algorithm [8] is computed on the basis of shortest logarithmic distance between two pages and ranked according to them. The higher rank is assigned to which page has a smaller distance. The Advantage of this algorithm is being fewer sensitive, it can find pages faster with high quality and more quickly with the use of distance based solution as compared to other algorithms. Distance Rank algorithm implements the PageRank properties. Then, a page has a high rank value if it has more incoming links on a page. Formulas for distance rank algorithms

$$Distance\ n[j] = (1 - a) * Dist_{n-1}[j] + a * \min_i (a * Distance_{n-1}[i] + \log(O[i]), i \in B(j)) \tag{9}$$

$B(j)$ shows list of pages that link to j and $O(i)$ is the number of out links in page i Sort url_queue by distance vector in ascending order.

G. Eigen Rumor Algorithm

The Eigen Rumor algorithm [15] is proposed by Ko Fujimura that ranks each blog entry on basis of weighting the hub and authority scores of the bloggers it is based on eigenvector calculations. So this algorithm enables a high score to be assigned to a blog entry entered by a good blogger but not linked to by any other blogs. In the recent scenario nowadays number of blogging sites is increasing, there is a challenge for web service provider to provide good blogs to the users. Page rank and HITS are providing the rank value to the blogs but some issues arise, if these two algorithms are applied directly to the blogs. The Eigen Rumor algorithm calculates three vectors, i.e., authority vector a , hub vector h , and reputation vector r , from information provisioning matrix P and information evaluation matrix E . These issues are:

1. The number of links to a blog entry is normally very small. As the result, the scores of blog entries are calculated by Page Rank.
2. The rank scores of blog entries as decided by the page rank algorithm is often very low, so it cannot allow blog entries to be provided by rank score according to their importance. So resolve these issues, an Eigen Rumor algorithm is proposed for ranking the blogs. The Eigen Rumor algorithm has connections to PageRank and HITS in that all are based on eigenvector calculation of the adjacency matrix of the links. One important thing is an agent is used to represent an aspect of human being such as a blogger, and an object is used to represent any object such as a blog entity. Using the Eigen Rumor algorithm, the hub and authority scores are calculated as bloggers and the encouragement of a blog entity that does

yet have any in-link entered by the blogger can be computed.

H. Time Rank Algorithm

Time Rank algorithm is used to improving the rank score by using the visit time of the web page measured the visit time of the page after applying original and improved methods of web page rank algorithm to know about the degree of importance to the users. This algorithm consumes the time factor to increase the accuracy of the web page ranking. Due to the methodology used in this algorithm, it can be assumed to be a combination of content and link structure [3]. The results of this algorithm are very satisfactory and in agreement with the applied theory for developing the algorithm.

I. Tag Rank Algorithm

Tag rank algorithm is also known as novel algorithm it is used for ranking the web page based on social annotations. This algorithm calculates the tags by using time factor of the new data source tag and the annotations behavior of the web users. This algorithm [7] provides a better authentication method for ranking the web pages. This algorithm provides very accurate results and this algorithm indexes new information resources in a better way. Future work in this direction can be to utilize co-occurrence factor of the tag to determine weight of the tag and this algorithm can also be improved by using semantic relationship among the co-occurrence tags.

J. Relation Based Algorithm

Fabrizio Lamberti, Andrea Sanna and Claudio Demartini proposed a relation based algorithm [3] for ranking the web

page for semantic web search engine. Various search engines are presented for better information extraction by using relations of the semantic web. This algorithm proposes a relation based page rank algorithm for semantic web search engine that depends on information extracted from the queries of the users and annotated resources. Results are very encouraging on the parameter of time complexity and accuracy. Further improvement in this algorithm can be increased the use of scalability into future semantic web repositories.

K. Query Dependent Ranking Algorithm

Lian-Wang Lee, Jung- Yi Jiang, ChunDer Wu and Shie-Jue Lee [4] have presented a query dependent raking algorithm for search engine. In this approach a simple similarity measure algorithm is used to measure the similarities between the queries. A single model for ranking is made for every training query with corresponding document. Whenever a query arises, then documents are extracted and ranked depending on the rank scores calculated by the ranking model. The ranking model in this algorithm is the combination of various models of the similar training queries. Experimental results show that query dependent ranking algorithm is better than other algorithms.

4. COMPARISON OF PAGE RANKING METHODS

On the basis of analysis, a comparison of various page ranking algorithms is done on the basis of some vaults such as main technique use, methodology, key in parameter, complexity, relevancy, quality of results, and limitations. On the basis of parameters we can find the powers and limitations of each algorithm.

Table 1. Comparison of page ranking algorithms

Algorithms	Main Technique	Methodology	Input parameters	Relevancy	Drawbacks
Page Rank	Web Structure Mining	This algorithm computes the score for pages at the time of indexing of the pages	Back Links	Less (this algo. Rank the pages on the indexing time)	Results come at the time of indexing and not at the query time.
Weighted Page Rank	Web Structure Mining	Weight of web page is calculated on the basis of input and outgoing links and on the basis of weight the importance of page is decided.	Back links and Forward links	Less as ranking is based on the calculation of weight of the web page at the time of indexing	Relevancy is ignored
Weighted Page Content	Web Structure Mining, Web	Gives sorted order to the web page	Content, Back links and Forward links	more (it consider the relative	WPCR is a numerical value based



Rank	Content Mining	returned by search engine as a numerical value in response for a new query.		position of the pages)	on which the web pages are given an order.
HITS	Web Structure Mining, Web Content Mining	It computes the hubs and authority of the relevant pages. It relevant as well as important page as the result.	Content, Back links and Forward links	More (this algo. Uses the hyperlinks so according the Hen zinger, 2001 it will give good results and also consider the Content of the page)	Topic drift and Efficiency problem.
Distance Rank	Web Structure Mining	Based on reinforcement learning which consider the logarithmic distance between the pages.	Forward links	Moderate due to the use of the hyperlinks.	If new page inserted between two pages then the crawler should perform a large calculation to calculate the distance vector.
Eigen Rumor	Web Structure Mining	Eigen rumor use the adjacency matrix, which is constructed from agent to object link not page to page link.	Agent/Object	High for Blog so it is mainly used for blog ranking	It is most specifically used for blog ranking not for web page ranking as other ranking like page rank, HITS.
Time Rank	Web Usages Mining	In this algorithm the visiting time is added to the computational score of the original page rank of that page	Original Page Rank and Sever Log	High due to the updation of the original rank according to the visitor time	Important pages are ignored because it increases the rank of those web pages which are opened for long time.
Tag Rank	Web Content	Visitor time is	Popular tags	Less as it	It is



	Mining	used for ranking. Use of sequential clicking for sequence vector calculation with the uses of random surfing model.	and related bookmarks	uses the keyword entered by the user and match with the page title	comparison based approach so it requires more site as input.
Relational Based Page Rank	Web Structure Mining	A semantic search engine would take into account keywords and would return page only if both keywords are present within the page and they are related to the associated concept as described in to the relational note associated with each page.	Keywords	High as it is keyword based algorithm so it only returns the result if the keyword entered by the user match with the page.	In this ranking algorithm every page is to be annotated with respect to some ontology, which is the very tough task.
Query Dependent Ranking	Web content Mining	This paper proposed the construction of the rank model by combining the results of similar type queries.	Training query	High (because the model is constructed from the training queries).	Limited number of characteristics is used to calculate the similarity.

5. CONCLUSION

The algorithms that are described above are effective in retrieving the web pages from the search engines. The link analysis algorithms are based on link structure of the documents. The page which has many links has many references can improves retrieval efficiency. In the integrated ranking approach comes under personalized web search. In integrated approach both the content and the link are integrated to improve the retrieval efficiency. Page Segmentation algorithms are used to segment the page as blocks and by separating as blocks the retrieval performance in the web context could be improved. Each and every algorithm has got its own merits and demerits. As per the requirements of a search engine we can utilize the above said algorithms. It helps to enhance the current page rank algorithm used by the Google and these web page ranking algorithms

could be used by several other search engines to improve the retrieval efficiency of the web pages as per the user's query.

REFERENCES

- [1] C. Ding, X. He, P. Husbands, H. Zha, and H. Simon, "Link Analysis: Hubs and Authorities on the World". Technical Report: 47847, 2001. [14] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR '04), IEEE, 2004.
- [2] Ali Mohammad Zareh Bidoki and Nasser Yazdani, "DistanceRank: An Intelligent Ranking Algorithm for Web Pages", Information Processing and Management, 2007.
- [3] Fabrizio Lamberti, Andrea Sanna and Claudio Demartini , "A Relation- Based Page Rank Algorithm for. Semantic Web



Search Engines”, In IEEE Transaction of KDE, Vol. 21, No. 1, Jan 2009.

[4] Lian-Wang Lee, Jung-Yi Jiang, ChunDer Wu, Shie-Jue Lee, "A Query- Dependent Ranking Approach for Search Engines", Second International Workshop on Computer Science and Engineering, Vol. 1, PP. 259-263, 2009..

[5] Su Cheng, Pan YunTao, Yuan JunPeng, Guo Hong, Yu ZhengLu and Hu ZhiYu "PageRank, "HITS and Impact Factor for Journal Ranking", Inproceedings of the 2009 WRI World Congress on Computer Science and Information Engineering – Vol. 06, PP. 285-290, 2009 .

[6] Neelam Duhan ,A.K.Sharma and Komal Kumar Bhatia , Page Ranking Algorithms : In proceedings of the IEEE International Advanced Computing Conference (IACC),2009.

[7] Xiang Lian and Lei Chen , "Ranked Query Processing in Uncertain databases", In IEEE KDE, Vol. 22, No. 3, March 2010.

[8] P Ravi Kumar, and Singh Ashutosh kumar, "Web Structure Mining Exploring Hyperlinks and Algorithms for Information Retrieval", American Journal of applied sciences, 7 (6) 840-845 2010.

[9] Saeko Nomura, Tetsuo Hayamizu, "Analysis and Improvement of HITS Algorithm for Detecting Web Communities". Volume 11-No 08,2011.

[10] Rekha Jain, Dr G.N.Purohit, "Page Ranking Algorithms for Web Mining", International Journal of Computer application, Vol 13, Jan 2011.

[11] Tamanna Bhatia, " Link Analysis Algorithms For Web Mining ", IJCST Vol. 2, Issue 2, June 2011. 43

[12] W.Xing and Ali Ghorbani, "Weighted PageRank Algorithm", Proc. Of the Second Annual Conference on Communication Networks and Services Research, IEEE,2013. 48

[13] A.M. Sote, Dr. S. R. Pande" Application of Page Ranking Algorithm in Web Mining" International Conference on Advances in Engineering & Technology–2014.

[14] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR '04), IEEE, 2004.

[15] Ko Fujimura, Takafumi Inoue and Masayuki Sugisaki,, "TheEigenRumor Algorithm for Ranking Blogs", In WWW 2005 2nd Annual Workshop on the Weblogging Ecosystem, 2005.