

# Approaches to Rank Technical Blog Users : A Survey

<sup>1</sup>Amruta Jadhav, <sup>2</sup>Sharvari Tamane\*

<sup>\*1</sup>Department of Computer Science, Jawaharlal Nehru Engineering College, Aurangabad (M.S.) India

<sup>2</sup>Department of Information Technology, Jawaharlal Nehru Engineering College, Aurangabad (M.S.) India

**Abstract** – The unprecedented growth of the Internet is leading to its widespread use among the online communities for various purposes. As web is expanding day by day, people are relying on search engines to explore the web. Research shows that, low barrier to publish content, anonymity of content uploader, exposure to millions of users and a potential of a very quick and widespread diffusion of message, due to this, various platforms on the Internet such as online discussion forums or blogs have become one of most important parts to promote ideologies in a sophisticated manner.

In such a scenario it is the duty of blog service provider to provide proper, relevant and quality information to the web user against their query. It is a challenge for web service provider to provide proper, relevant and quality information to the web user by using the blog page contents.

There are some works in this field to rank users based on their behavioral features. Whereas, this paper is our first step of research to detect technical blog users using web page ranking algorithms. Since this is a survey paper, it deals with literature study, analysis and comparison of blog posts ranking algorithms based on various parameters to find out their advantages and limitations for the ranking of the blog users, possibly technical blog users. Based on the analysis of different ranking algorithms, a comparative study is done to find out their relative strengths and limitations to find out the further scope of research in technical blog users ranking.

**Keywords** – Web content mining, ranking, ranking algorithms, blog, discussion forums

## I. INTRODUCTION & MOTIVATION

Over the past decade, blogs and online discussion forums has emerged into a dynamic form of world-wide interpersonal communication. As the volume of information on the internet is increasing day by day there is a challenge for blogging service providers to provide proper and relevant information to users.

Due to simplicity of navigation, low publication barriers (users only need to have a valid account on website) and anonymity of content uploader (liberty to upload any content without revealing their real identity) have led users to misuse blogging services in several ways by uploading spam or irrelevant data [1].

Dynamically increasing unstructured or semi-structured information on the blogging websites lead a great challenge for both the users, who are seeking for efficiently valuable information and for the technical people, who are offering service to an individual user, covered in the billions of blog posts. To triumph over these problems, data mining techniques must be useful on the www. Nowadays, most of the people rely on web search engines to find and retrieve information [29]. The

enormous growth, assorted, dynamic and unstructured nature of web makes internet awfully difficult in searching and retrieving relevant information and in presenting query results. Every day search engines are giving response to millions of queries. An efficient ranking of query words has a major role in efficient searching for queried content. There are various challenges associated with the ranking of blog posts such that some blog posts are made only for navigation purpose and some of the posts do not possess the quality of self descriptiveness. Over past 10 years, since 2005 several approaches, techniques, algorithms and tools have been proposed for ranking of web pages and to bring solutions to detect blog users based on their behavioral features.

The aim of the study presented in this paper is to conduct a systematic literature survey to analyze the currently important algorithms for ranking of blog users to find out their relative strengths, limitations and provide a future direction for the research in the field of efficient algorithm for ranking of the web pages [13][14]. The remaining part of this paper is organized as follows: Brief about web mining is summarized in Section II. Section III introduces the term ‘ranking’ and various aspects related to ranking. Various ranking algorithms are discussed in section IV. A tabular summary is presented in section V, which summarizes the techniques, advantages and limitations of some of the important ranking algorithms. Based on the literature analysis, a comparison of some of various ranking algorithms is presented in section VI and a conclusion is given in section VII.

## II. WEB MINING

Web mining is making use of data mining techniques to automatically discover and extract information from Web documents or services (Etzioni, 1996, CACM 39). It consists of following tasks,

1) Resource finding: It involves the task of retrieving intended web documents. It is the method by which we extract the data from online or offline text resources accessible on web.

2) Information selection and pre-processing: It involves the automatic selection and pre-processing of specific information from retrieved web resources.

3) Transformation: This process transforms the original retrieved data into information. The transformation could be rejuvenation of stop words, arising or it may be aimed for obtaining the desired representation such as finding phrases in training mass.

4) Generalization: It automatically determines general patterns at individual web sites as well as across multiple sites. Machine learning and data mining techniques are helpful in generalization.

5) Analysis: It involves the validation and interpretation of the mined patterns. It has an important role in pattern mining. A human has an important role in information on knowledge discovery process on web.

Web mining is made of three branches i.e. web content mining (WCM), web structure mining (WSM) and web usage mining (WUM). WCM is responsible for exploring the proper and relevant information from the contents of web. WSM is used to find out the relation between different web pages by processing the structure of web. WUM is responsible for recording the user profile and user behavior inside the log file of the web. The WCM mainly concentrates on the structure of the document whereas WSM explore the structure of the link inside the hyperlink between different documents and classify the pages of web. The number of out links i.e. links from a page and the number of in link i.e. links to a page are very important parameter in the area of web mining. The popularity of the web page is generally measured by the fact that a particular page should be referred by large number of other pages and the importance of web pages may be adjudged by a large number of out links contained by a page. So WSM becomes a very important area to be researched in the field of web mining [13][14..

### III. RANKING

A ranking is simply defined as an association, between a set of items on repository or web pages such that, for any two items, the first is either 'ranked higher than', 'ranked lower than' or 'ranked equal to' the second. In mathematics, this is known as a weak order or total pre-order of objects. It is not necessarily a total order of objects because two different objects or items can have the equivalent ranking. The rankings as we know are totally in ordered manner.

To bring down detailed measures to an arrangement of more than two things in a successive order, rankings make it possible to measure complex of heterogeneous information according to certain specific criterion. For example, the Internet search engine like; Google, may rank the pages it according to their relevance, making it conceivable for the end user rapidly to select the pages they are likely wish to retrieve. The ranking method is more effective for end-users is that all the similar terms or keyword in the query are used for retrieval, which result is ranked based on co-occurrence of query terms, as modified by statistical term-weighting. The ranking approach also works well for phrases or for the complex queries that may be difficult for end-users. For example, "human factors, human emotions, performance of the system in medical, military, university, organization data-space" is difficult for end-users to express in Boolean logic.

### A. Ranking Algorithms

Ranking algorithms [26] form an essential part of any search engine and a large amount of research has been done on them because they determine the quality of a search engine from the user's perspective. A manual ranking scheme could have been sufficient, if the Web had been composed of a few hundred pages, as was the case during the initial years of the Web with search engines [27]. However, with the dynamic explosion of information over web it could no longer be practical to rank millions of pages and automated means had to be developed in the form of ranking algorithms. The paper gives an overview of the various ranking algorithms that can be used to enhance the search experience of the technical users over the blogs or online discussion forums.

### B. Need for Ranking

There are billions of web pages, blog posts on the web and it is more than likely that when a user enters a word to be searched for there will be thousands of results containing that word. It is obviously impractical for the user to visit all of these pages. Thus, one of the goals of a search engine is to provide the user with results that are most likely to be beneficial to him/her in least possible amount of response time.

When the search engines return the result of a user query, only a predetermined number of documents are presented to the user. Thus, it is very important that the most relevant documents are included in the result and are prioritized in the display. This important task is performed by the ranking function. A ranking function that prioritizes the documents most relevant to a user will satisfy the user.

## IV. RANKING ALGORITHMS

### 1) HITS Algorithm

HITS algorithm ranks the web document by dealing in links and out links of the web document. In this algorithm if a web document is pointed by many hyperlinks then it is named as authority and if the document point to various hyperlinks then it is named as HUB.

HITS algorithm is a link based algorithm. In HITS [12] algorithm, web document's ranking is determined by analyzing their textual contents against a given query. After collection of the web documents, the HITS algorithm focus on the structure of the web only, neglecting their textual contents. Primarily HITS algorithm has some cons which are given below. (i) High rank value is given to some well-known website that is not highly applicable to the given query. (ii) Drift of the topic occurs when the hub has multiple topics as equivalent weights are given to all of the out links of a hub page.

To minimize the problem of the original HITS algorithm, a clever algorithm is proposed by allusion [11]. Clever algorithm is the adjustment of standard original HITS algorithm. This algorithm provides a weight value to every link depending on the terms of queries and endpoints of the link. An anchor tag is combined to decide the

weights to the link and a large hub is broken down into smaller parts so that every hub page is intensified only on one topic. Another limitation of standard HITS algorithm is that it assumes equal weights to all the links pointing to a webpage and it fails to identify the facts that some links may be more important than the other. To resolve this problem, a probabilistic analogue of the HITS (PHITS) algorithm is proposed by Allusion [10]. A probabilistic clarification of relationship of term document is provided by PHITS. It is able to identify authoritative document as claimed by the author. PHITS gives better results as compared to original HITS algorithm. Other difference between PHITS and standard HITS is that PHITS can estimate the probabilities of authorities compared to standard HITS algorithm, which can provide only the scalar magnitude of authority [7].

### 2) Eigen Rumor Algorithm

As the number of blogging sites is increasing regularly, there is a challenge for service provider to provide good blogs to the users. Though page rank and HITS algorithms shows potential in providing the rank value to the blogs but some limitations occurs, if these two algorithms are applied directly to the blogs. The rank scores of blog entries as determined 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. To resolve these restrictions, an Eigen Rumor algorithm [7] is proposed for ranking the blogs. This algorithm provides a rank score to every blog by weighting the scores of the hub and authority of the bloggers depending on the calculation of Eigen vector. This algorithm enables a higher score to be assigned to the blog entries submitted by a good blogger but not yet linked to by any other blogs based on acceptance of the blogger's prior work.

### 3) Weighted Links Rank Algorithm

Web page developers give more importance to some links using different HTML tags, because some Web resources are more significant than others. Hence, a link ranking technique that gives different weights to links may improve over uniform weight links [8]. This algorithm provides weight value to the link based on three parameters i.e. length of the anchor text, tag in which the link is contained and relative position in the page. Simulation results show that the results of the search engine are enhanced using weighted links. The length of anchor text seems to be the best attributes in this algorithm. Relative position, which reveal that physical position does not always in synchronism with logical position is not so result oriented. Future work in this algorithm allows, tuning of the weight factor of every term for further advancement.

### 4) Weighted Page Content Rank Algorithm

Weighted Page Content Rank Algorithm (WPCR) is a anticipated page ranking algorithm which is used to give a sorted order to the web pages returned by a search engine in response to a user query. WPCR has a numerical value based on which the web pages are given an order. This algorithm handles web structure mining as well as web

content mining techniques. The Web structure mining is used to calculate the weight of the web page and web content mining is used to find how much relevant a web page is? Here importance means the popularity of the page i.e. how many pages are pointing to that page or how many pages are referred by this particular page [4]. It can be calculated by depending on the number of in links and out links of the page. Here relevancy means matching of the page with the user fired query. If a page is maximally matched to the query, that becomes more relevant page than other.

### 5) Tag Rank Algorithm

An innovative algorithm named as Tag Rank [6] is proposed by Shen Jie, Chen, Zhang Hui, Sun Rong-Shuang, Zhu Yan and He Kun. It deals with ranking the web page based on social annotations. This algorithm calculates the heat of the tags by using time factor of the new data source tag and the annotations behavior of the web users. This algorithm provides a better authentication method for ranking the web pages. The results of this algorithm are very precise and this algorithm index new information resources in a better way. Future work in this track can be to utilize co occurrence factor of the tag to determine weight of the tag. This algorithm can also be improved by using semantic relationship among the co-occurrence tags.

### 6) Query Dependent Ranking Algorithm

A query dependent ranking algorithm for search engine have been presented by Lian- Wang Lee, Jung- Yi Jiang, ChunDer Wu and Shie-Jue Lee [5], where 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 consequent document. Whenever a query arises, then documents are extracted and ranked depending on the rank scores intended by the ranking model. The ranking form 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.

### 7) Correlation Ranking Algorithm

In addition to relevance ranking, this algorithm also detects redundant documents. Removal of these redundant documents improves the quality of search results by providing unique relevant information. Normalized discounted cumulative gain method is used for evaluating this ranking algorithm [13]. Correlation analysis is used to find the related documents from the input document set of some particular category.

### 8) Distance Rank Algorithm

Ali Mohammad Zareh Bidoki and Nasser Yazdani [18] has proposed an intelligent ranking algorithm named as distance rank, which is based on reinforcement learning algorithm. In this algorithm, the distance between pages is considered as a punishment factor. In this algorithm the ranking is done on the basis of the shortest logarithmic

distance between two pages and ranked according to them. The Advantage of this algorithm is that it can find pages with high quality and more quickly with the use of distance based solution. The Limitation of this algorithm is that the crawler should perform a large calculation to calculate the distance vector, if new page is inserted between the two pages.

#### 9) Time Rank Algorithm

An algorithm named as Time Rank, for improving the rank score by using the visit time of the web page is proposed by H Jiang et al. [19] Authors have 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 utilizes 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. The results of this algorithm are very satisfactory and in agreement with the applied theory for developing the algorithm

#### 10) Relation Based Algorithm

Fabrizio Lamberti, Andrea Sanna and Claudio Demartini [21] proposed a relation based algorithm for the 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 the increased use of scalability into future semantic web repositories.

#### 11) Query Dependent Ranking Algorithm

Lian- Wang Lee, Jung- Yi Jiang, ChunDer Wu and Shie-Jue Lee [22] have presented a query dependent ranking 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.

#### 12) Ranking and Suggestive Algorithm

M Vojnovic et al. [23] have proposed a ranking and suggestive algorithm for popular items based on user feedback. User feedback is measured by using a set of suggested items. Items are selected depending on the preferences of the user. The aim of this technique is to measure the correct ranking of the items based on the actual and unbiased popularity. Proposed algorithm has various techniques for suggesting the search query. This

algorithm can also be used for providing tag suggestion for social tagging system. In this algorithm various techniques for ranking and suggesting popular items are studied and results are provided based on their performance. Results of this algorithm demonstrate that randomized update and light weight rules having no special configurations provide better accuracy.

#### 13) Comparison and Score Based Algorithm

NL Bhamidipati et al. [24] have proposed a more common approach whereby the scoring scheme may be perceived to be dissimilar if they induce identical ranking. In this algorithm a metric has been proposed to compare score vectors and the similarity and dissimilarity are measured on the basis of score fusion. Experimental results are in agreement with the theory applied and results demonstrate the various applications of the metric used in the theory applied for the proposed algorithm.

#### 14) Algorithm for Query Processing in Uncertain Databases

Xiang Lian and Lei Chen [25] have proposed an algorithm for ranked query processing in uncertain databases. Uncertain database management is used in various areas such as tracking of mobile objects and monitoring of sensor data. Due to intrinsic difference between certain and uncertain data. To remove these limitations authors have proposed a novel algorithm. Uncertain database are not exact points and generally occurs within a limited region. Existing algorithms for rank query processing are generally developed for exact or certain data but they cannot be applied directly to uncertain database due to accelerate the probabilistic rank query with monotonic preference functions over the uncertain databases. Authors have proposed two effective techniques named as probabilistic and spatial to reduce the PRank search space. Exhaustive experiment results show that proposed algorithm is very effective and efficient with respect to number of PRank candidates to be refined and wall clock time.

#### 15) Extended Page Rank Algorithm

Tarique Anwar et. al. [2] have proposed an approach to identify a ranked list of radically influential users in Web forums, by formulating a radicalness measure and a variety of collocation-based association measures, and designed an algorithm based on Page Rank to rank the radically influential users. The experimental results on a standard data set are promising that outperforms the existing User Rank algorithm in which the contingency coefficient measure is found as the most promising measure. The result confirms that collocation-based association measures deal with such ranking problem more effectively than textual and temporal similarity based measures.

#### 16) B2Rank Algorithm

Blogs not have so efficient search engines for them. One reason is differences between regular web pages and blog pages and inefficiency of conventional web pages

ranking algorithms for blogs ranking. There are some works in the field but users' behavioral features have not considered yet. M. A. Tayebi et. al. [28] presents a new blogs ranking algorithm called B2Rank based on these features.

V. COMPARISON OF RANKING ALGORITHMS

Based on the literature studied, a comparison of some of various ranking algorithms is shown in Table I & Table II. Comparison is done on the basis of parameters such as main technique used, methodology, input parameter, relevancy, quality of results, importance of algorithms and limitations.

TABLE I. SUMMARY OF RANKING ALGORITHMS

Algorithm	HITS	Eigen Rumor	Weighted Links Rank Algorithm	Weighted Page Content Rank
<b>Main Technique</b>	Web Content Mining	Web Content Mining	Web Content Mining	Web Content Mining
<b>Methodology</b>	It computes the hubs and authority of the relevant pages. It gives relevant as well as significant page as the result	Eigen rumor use the adjacency matrix method, which is constructed from agent to object link not from page to page link	It gives different weight to web links based on 3 attributes: Relative position in page, tag where link is contained, length of anchor text	It gives sorted order to the web pages returned by a search engine as a numerical value in response to a user query
<b>I/P Parameter</b>	Content, Back and Forward links	Agent/Object	Content, Back and Forward links	Back links, Forward links and content
<b>Relevancy</b>	More (this algorithm uses the hyperlinks so according to Henzinger, 2001 it will give good results and also consider the content of the page)	High for Blogs so it is mainly used for blog ranking.	More (it consider the relative position of the pages )	More relevant to a user given query.
<b>Quality of Results</b>	Less than Page Rank	Higher than Page Rank and HITS	Medium	High
<b>Importance</b>	Moderate Hub & authorities scores are handled.	High for blog ranking.	Not specifically quoted.	It provide important information and relevancy about a given query by using web structure and web content mining
<b>Limitations</b>	Topic drift and efficiency problem	It is most purposely used for blog ranking not for web page ranking as other ranking like page rank, HITS	Relative position was not so effective, representing that the logical position not always matches the physical position	Extra calculations to find the weights of links

TABLE II. SUMMARY OF RANKING ALGORITHMS

Algorithm	Tag Rank	Query Dependent Ranking	Correlation ranking algorithm	Distance Rank	Time Rank
<b>Main Technique</b>	Web Content Mining	Web Content Mining	Web Content Mining	Web Structure Mining	Web Usages Mining
<b>Methodology</b>	Visitor time is used for ranking. Use of sequential clicking for sequence vector calculation with the uses of random surfing model.	This proposed the construction of the rank model by combining the results of similar type queries.	It is based on correlation method. Input datasets are preprocessed and then the term frequency for the common words between documents is computed, then correlation coefficient is computed	Based on reinforcement learning which consider the logarithmic distance between the pages.	In this algorithm the visiting time is added to the computational score of the original page rank of that page.
<b>I/P Parameter</b>	Popular tags and related bookmarks	Training query	Web documents	Forward links	Original Page Rank and Sever Log
<b>Relevancy</b>	Less as it uses the keyword entered by the user and match with the page title	High (because the model is constructed from the training quires)	More	Moderate due to the use of the hyperlinks.	High due to the updation of the original rank according to the visitor time.
<b>Quality of Results</b>	Less	High	The quality of search results obtained through this approach is accurate.	High	Moderate
<b>Importance</b>	High for social site	High, because it gives the results for user's query as well as results for similar type of query	High document based searching	High, It is based on distance between the pages.	High, Consideration of the most recently visited pages .
<b>Limitations</b>	It is comparison based approach so it requires more site as input.	Limited number of characteristics are used to calculate the similarity.	It is Comparison based approach	If new page inserted between two pages then the crawler should perform a large calculation to calculate the distance vector.	Important pages are ignored because it increases the rank of those web pages which are opened for long time.

## VI. CONCLUSION

In the current era, the user always desires to get the best in a petite time. User generally spends a lot of time in sifting through the search results to find the relevant content. The ranking algorithms, which are significance of web mining, play a major role in making the user search navigation easier in the results of a search. Solutions to rank web documents and blog posts on the Internet have recently attracted a lot of research attention.

In this survey paper, we reviewed the state-of-the-art in the area of techniques for ranking web documents and blog posts on the Internet and blogging websites or discussion forums. The paper presented a detailed comparison study of some prevalent ranking algorithms. After going through exhaustive analysis of algorithms for ranking of web pages against the various parameters, it is concluded that these algorithms have limitations particularly in terms of time response, accuracy of results, importance of the outcome and relevancy of results. An efficient ranking algorithm should meet out these challenges efficiently with compatibility with global principles of web technology.

## ACKNOWLEDGMENT

One of the authors, Amruta Jadhav is very much thankful to Dr. Madhuri Joshi and Dr. Vijaya Musande, J.N.E.C. Aurangabad for their fruitful discussion and suggestions.

## REFERENCES

- [1] Muthiah S, Huang B, Arredondo J, Mares D, Getoor L, Katz G, Ramakrishnan N, "Planned Protest Modeling in News and Social Media"
- [2] Tarique Anwar, Muhammad Abulaish, "Ranking Radically Influential Web Forum Users," IEEE Transactions on Information Forensics and Security, vol. 10, no. 6, pp. 1289-1298, 2015
- [3] G. Poonkuzhali, R. Kishore Kumar, P. Sudhakar, G.V.Uma, K.Sarukesi, "Relevance Ranking and Evaluation of Search Results through Web Content Mining", International Multi Conference of Engineers and Computer Scientists 2012 vol.1,IMECS 2012.
- [4] Pooja Sharma et al., "Weighted Page Content Rank for Ordering Web Search Result", International Journal of Engineering Science and Technology vol. 2, no. 12), pp. 7301-7310, 2010.
- [5] 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.
- [6] Shen Jie, Chen, Zhang Hui, Sun Rong-Shuang, Zhu Yan and HeKun, "TagRank: A New Rank Algorithm for Webpage Based on Social Web" In proceedings of the International Conference on Computer Science and Information Technology, 2008.
- [7] Kos Fujimura, Takafumi Inoue and Masayuki Sugisaki, "The Eigen Rumor Algorithm for Ranking Blogs", 2nd Annual Workshop on the Weblogging Ecosystem, 2005.
- [8] Ricardo Baeza-Yates and Emilio Davis, "Web page ranking using link attributes", 13th international World Wide Web conference on Alternate track papers & posters, PP.328-329,2004.
- [9] Wenpu Xing and Ali Ghorbani, "Weighted Page Rank Algorithm", 2nd Annual Conference on Communication Networks & Services Research, pp. 305-314, 2004.
- [10] D. Cohn and H. Chang, "Learning to Probabilistically Identify Authoritative Documents", 17th International Conference on Machine Learning, pp. 167-174, 2000.
- [11] S. Chakrabarti, B. E. Dom, S. R. Kumar, P. Raghavan S.Rajagopalan, A. Tomkins, D. Gibson, and J. Kleinberg, "Mining the Web's Link Structure", Computer, 32(8), PP.60-67, 1999.
- [12] Jon Kleinberg, "Authoritative Sources in a Hyperlinked Environment", Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 1998.
- [13] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", 2nd Annual Conference on Communication Networks & Services Research, PP. 305-314, 2004.
- [14] Neelam Duhhan, A. K. Sharma and Komal Kumar Bhatia, "Page Ranking Algorithms: A Survey", IEEE International Advanced Computing Conference (IACC), 2009.
- [15] D. Cohn and H. Chang, "Learning to Probabilistically Identify Authoritative Documents", 17th International Conference on Machine Learning, pp. 167-174, 2000.
- [16] Sung Jin Kim and Sang Ho Lee, "An Improved Computation of the PageRank Algorithm", European Conference on Information Retrieval (ECIR), 2002.
- [17] Ricardo Baeza-Yates and Emilio Davis, "Web page ranking using link attributes", 13th International World Wide Web conference on Alternate track papers & posters, PP. 328-329, 2004.
- [18] Ali Mohammad Zareh Bidoki and Nasser Yazdani, "DistanceRank: An Intelligent Ranking Algorithm for Web Pages", Information Processing and Management, 2007.
- [19] H Jiang et al., "TIMERANK: A Method of Improving Ranking Scores by Visited Time", Seventh International Conference on Machine Learning and Cybernetics, 2008.
- [20] Shen Jie, Chen Chen, Zhang Hui, Sun Rong-Shuang, Zhu Yan and He Kun, "TagRank: A New Rank Algorithm for Webpage Based on Social Web", International Conference on Computer Science and Information Technology, 2008.
- [21] Fabrizio Lamberti, Andrea Sanna and Claudio Demartini, "A Relation-Based Page Rank Algorithm for Semantic Web Search Engines", IEEE Transaction of KDE, Vol. 21, No. 1, Jan 2009.
- [22] 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.
- [23] Milan Vojnovic et al., "Ranking and Suggesting Popular Items", IEEE Transaction of KDE, Vol. 21, No. 8, Aug 2009.
- [24] NL Bhamidipati et al., "Comparing Scores Intended for Ranking", IEEE Transactions on Knowledge and Data Engineering, 2009.
- [25] Xiang Lian and Lei Chen, "Ranked Query Processing in Uncertain databases", In IEEE KDE, Vol. 22, No. 3, March 2010.
- [26] "Ranking Algorithm", <http://orion.lcg.ufrj.br/Dr.Dobbs/books/book5/chap14.htm>
- [27] The Google Search Engine: Commercial search engine founded by the originators of PageRank.
- [28] M. A. Tayebi, S. M. Hashemi, A. Mohades, "B2Rank: An Algorithm for Ranking Blogs Based on Behavioral Features", IEEE/WIC/ACM International Conference on Web Intelligence, pp. 104-107, 2007.
- [29] <http://webminingissues.blogspot.in>