

## Survey of Contemporary Ranking Algorithms

Satish Muppidi<sup>1</sup>

Assistant Professor, Department of IT,  
GMR Institute of Technology (GMRIT),  
GMR Nagar, Rajam-532127, Srikakulam,  
Andhra Pradesh, India.  
E-mail: [msatishmtech@gmail.com](mailto:msatishmtech@gmail.com)

Venkata Naveen Koraganji<sup>2</sup>

B.Tech, GMR Institute of Technology (GMRIT),  
GMR Nagar, Rajam-532127, Srikakulam,  
Andhra Pradesh, India.  
E-mail: [naveen4262@gmail.com](mailto:naveen4262@gmail.com)

### Abstract

In the present developing computational world ranking has become a common feature of modern retrieval system. Ranking algorithm is nothing but suppose if we give any query then the relevant and non-relevant data will be displayed so, it is used for reducing the user overhead by displaying only related data first, which satisfies the user given query for retrieving an information. Uses of ranking algorithm is to retrieve an candidate rank details very fast, and ranking algorithm will also be used to save the users time to get the details quickly. There are different types of ranking algorithms like: Page rank algorithm, Hits algorithm, Distance algorithm, weighted page rank algorithm, Ontology page rank algorithm etc. This paper reviews the mentioned algorithms in terms of efficiency and popularity.

**Keywords:** Page rank, Hits, ontology, weighted, Distance, Information Retrieval Systems

### Introduction

In the present developing world, the users' growth rapidly increases in browsing of websites. Comparing to past now-a-days the web grows by roughly a million electronic pages. In such a way the ranking algorithm becomes very important in retrieving data which is relevant to user given query. The main use of this ranking algorithm is to save the user time and to retrieve user wanted data very fastly. In this we use several algorithms like Page Rank Algorithm, HITS Algorithm [2], Distance Algorithm [6], Ontology Page Ranking Algorithm [4]. This paper reviews the mentioned algorithms in terms of efficiency and popularity.

This paper divided into 5 sections. Second section reviews basics of ranking algorithms. Majorly section 3 discussed various ranking algorithms that cover pager ranking, HITS etc. algorithms. Section 4 compares the ranking algorithms and finally section 5 concludes the paper.

### Basics of Ranking Algorithms in Bygone Days

The main aim of the ranking algorithm is when a user gives any query then as a general approach all relevant and non-relevant data will be displayed [8]. Improved ranking algorithms reduces the overhead of user, when user gives any query both relevant and non-relevant data will be displayed but relevant data is displayed first and next the non-relevant data, so that users time saved and user can perform his search operation in an efficient way.

Now-a-Days in the present developing retrieval process the computation is very high among users so the user obviously select which is retrieved efficiently in a short time of period in such a the ranking algorithm is very much used. This paper elaborated some of Algorithms like: Page ranking, hits, ontology based. Various page raking algorithms and their authors are listed in the table 1.

**Table 1:** Various Ranking Algorithms

Algorithm Name	Author	Year
HITS(hyperlink-induced Topic search Algorithm)	Jon Kleinberg	1996
Page Ranking Algorithm	Larrypage and Sergey	1996
Improved HITS Algorithm	Longzhuang Li, Yi Shang, Wei Zhang	2002
Weighted Page Rank Algorithm	Wenpu Xing and Ali Ghorbani	2004
Distance Rank Algorithm	Ali Mohammad Zare Bidoki, Nasser Yazdan	2007
Relation Based Algorithm	F Lamberti	2009
New Page Ranking Algorithm	Hema Dubey, Prof. B. N. Roy	2010
Rank Based Finger Algorithm	J. Machaj, R. Piche and P. Brida	2011
Ontology Ranking Algorithm on Semantic Web	K. Sridevi and Dr. R. Umarani	2013

### Various Ranking Algorithms

Ranking algorithm workflow is shown in Figure 1. In this model documents are taken by the indexer for indexing process. From the index database, top retrievals are created based on various number of user queries. This process generates ranking model which will be frequently updated by learning algorithm so that efficient results can be displayed. Ranking algorithm was improved based on the Optimized Normalization Technique [1], further it was improved based on Ontology ranking algorithm [4] and also based on Finger printing algorithm [3] and also improved based on weighted page rank algorithm [7]. Floris Geerts et al. [5] proposed a popular relational link-based model which was as a basis by various researchers, because it was based on the link analysis methods for analyzing relational databases.

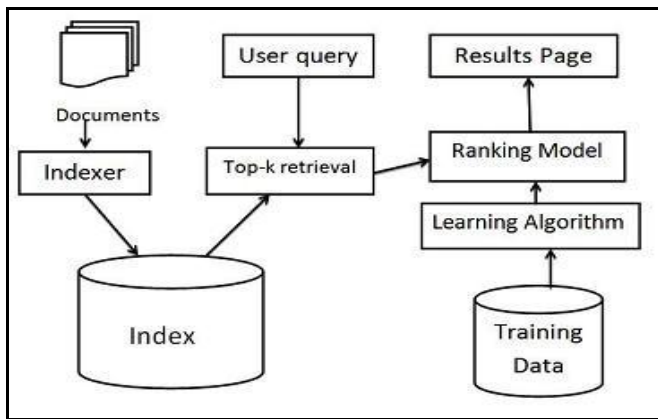


Figure 1: General Process of Ranking Algorithm

### A. Page Rank Algorithm

Page Rank Algorithm was invented by Larry Page and Sergey in 1996. With the rapid growth of Internet in the past History, Google became one of the most popularly used searching processes by users all over the world. The success behind Google is its quality of search results as compared to other searching Techniques. This quality of search results is based on PageRank, Having a good understand algorithm to rank web pages. Page Rank algorithm is the most commonly used algorithm for ranking the various pages. Figure 2 shows the process of page ranking algorithm model.

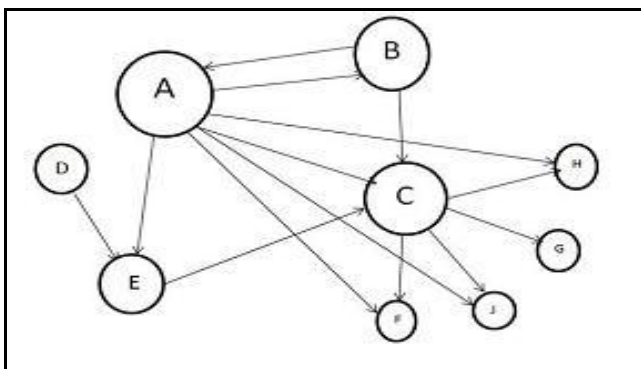


Figure 2: Page Ranking Algorithm

The working ability of the Page Rank algorithm depends upon the back links that how much a web page contains. All the ranks of the back links is larger than the page then it is provided with a more rank. The Page Rank is Given By:  
 $PR(A) = (1-d) + d (PR(T_1) / C(T) + \dots + PR(T_n) / C(T_n))$ .

### B. HITS Algorithm

HITS algorithm developed by Jon Kleinberg, he made use of the links of the web network in order to discover and rank pages relevant to user given query for retrieving data [2]. HITS algorithm is link structure algorithm. This algorithm will rank the pages by exceeding as in-links and out-links of the web pages. There are many several applications of hits such as Find web communities, finding related information, Hits algorithm is derived as Hubs and Authorities shown in figure 3.

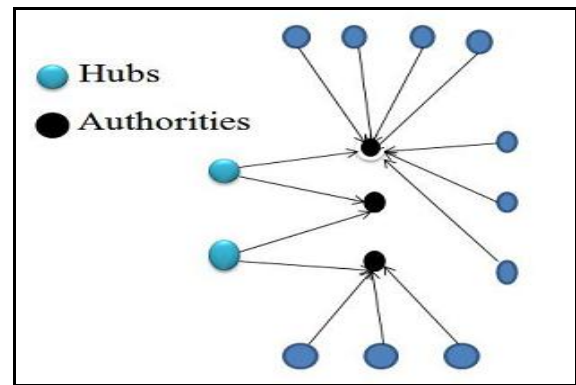


Figure 3: Hubs and Authorities

Now initialize all weights to 1 and Repeat this until convergence

- 1) O operation: By using this O operation hubs will collect the weight of authorities

$$h_i = \sum_{j:i \rightarrow j} a_j$$

- 2) I operation: By using this I operation authorities collect the weight of the hubs.

$$a_i = \sum_{j:j \rightarrow i} h_j$$

- 3) Normalize weights are taken under some norm

### C. Distance Algorithm

The Distance Algorithm was introduced by Ali Mohammad Zareh Bidoki, Nasser Yazdan in the year 2007 [6]. The Main Aim of this algorithm is to find the Distance between the pages and to give the ranking based on the distance. This algorithm is based on the distance between any pages. The ranking is given on the basis of the shortest logarithmic distance between two pages.

The ranking is given on the basis of the shortest logarithmic distance between two pages. Suppose let us take one small example: let us take distance between two points i.e...distance between point i and j was less than the distance between i and k that means ( $d_{ij} < d_{ik}$ ) then the probability that a random surfer from i reach to j will be more than k.

$$d_{ij} = \sum_{s \in \text{path}(i,j)} \log \alpha(s) = - \sum_{s \in \text{path}(i,j)} \log \frac{1}{\alpha(s)} = - \log \prod_{s \in \text{path}(i,j)} \frac{1}{\alpha(s)}$$

Now suppose if  $d_{ij}$  shows the distance between two points i and j from above equation, then  $d_j$  denotes the average distance of page j, It can be expressed as following:

$$d_j = \frac{\sum_{i=1}^N d_{ij}}{N}$$

But the main problem of the average distance is its complexity.

### D. Ontology Page Ranking Algorithms

Now-a-Days Web is the representation of knowledge this consists of a huge amount of data. Ontologies provide an

efficient way to reduce the amount of information overload by defining the structure of a specific domain and enabling easier access to the information.

The use of ontology has been increased, similar to the Web search. Ontology process can prove its ability only when the retrieval involves with highly relevant information based on the users given query [4].

#### E. Rank Based Fingerprinting Algorithm for Indoor Positioning:

This algorithm was proposed by J. Machaj, R. Piche and P. Brida in the year 2011 [3]. The main Aim of this algorithm is its performance same for any receiver, without the need of calibration of RSS (Received Signal Strength) values.

Results of this process shown are that the algorithm can achieve better accuracy than some commonly used fingerprinting algorithms. In future more testing may be done, and the influence of AP placement and the number of

APs on localization accuracy may have a chance of investigated. Similarity measures, such as Webber et al., will also be studied by using this finger printing process.

#### Comparison Between Various Ranking Algorithms

Among all the ranking algorithms, page rank algorithm got more popular because of its wide usage. As mentioned earlier several algorithms are coming up, most of them based upon the page ranking algorithm only. Finding the differences between various algorithms is also a mundane task, but just this paper made an attempt on this. Differences made by comparing main technique, methodology, quality of result, relevancy, importance and algorithms wide usage. Table 2 shows these differences.

**Table 2:** Comparison between Various Ranking Algorithms

Name	Hits Algorithm	Page Rank Algorithm	Distance Algorithm	Ontology Ranking Algorithm
<b>Main Technique</b>	Both Web Structure and Web Content Mining	Web Structure Mining	To Reduce the Distance Between any two Distances	A way to reduce the amount of overload data by defining the Structure for a Specific domain
<b>Methodology</b>	It will Compute the HUBS and authority of relevant pages	This will compute for score at the time of indexing of pages	Minimizes punishment or distance so that a page with less distance to have a higher rank	Increases the scope of the knowledge regarding searching and makes the user to view the relevant needed data query on the top most
<b>Quality Of Result</b>	Larger than Page Rank Algorithm	Medium	Higher than Page Rank Algorithm	High
<b>Relevancy</b>	More	Less	More	High
<b>Importance</b>	Moderate	Highly Important because Backlinks are considered	Minimizes the distance so page will be given Higher Rank with less Distance	Semantic web are for searching in an efficient manner
<b>Uses</b>	Used to Observe the Important and behavior of web Pages	Used for Ranking Various Pages	To Reduce the complexity between two Distances	To Reduce the Overload of Data

#### Conclusion

This paper made an attempt to review various recent ranking algorithms. After considering several features together it was considered still page ranking algorithm has its own flavor and basis for most of the ranking algorithms. As semantic web is considered as next generation web, ontologies will play a crucial role for searching the contents on the web. So, ontology based ranking algorithms may rule the world. This paper majorly compared HITS, page rank, distance model and ontology based ranking algorithms.

#### References

- [1] Hema Dubey, Prof. B. N. Roy; "An Improved Page Rank Algorithm based on Optimized Normalization Technique", International Journal of Computer Science and Information Technologies, vol. 2, no. 5, pp.2183-2188, 2011.
- [2] Xiuxia Tian, Yajun Du, Wen Song, Wenjun Liu, Yingmin Xie; "Improvements of HITS Algorithm Based on Triadic Closure", Journal of Information & Computational Science, pp.1861-1868; April, 2014.

- [3] J. Machaj, R. Piché and P. Brida; “Rank Based Fingerprinting Algorithm for Indoor Positioning”, International Conference on Indoor Positioning and Indoor Navigation, Guimaraes, Sept 23, 2011.
- [4] K. Sridevi and Dr. R. Umarani; “Ontology Ranking Algorithms on Semantic Web: A Review”, International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, pp: 3471-3476; September, 2013.
- [5] Floris Geerts Heikki Mannila Evimaria Terzi; “Relational link-based ranking”, Proceedings of the 30th VLDB Conference Toronto, Canada, pp: 552-563; Sept 3, 2004.
- [6] Ali Mohammad Zareh Bidoki , Nasser Yazdani; “DistanceRank: An intelligent ranking algorithm for web pages”, Information Processing & Management, vol.44, no. 2, pp. 877-892; June, 2007.
- [7] NeelamTyagi, Simple Sharma; “Weighted Page Rank Algorithm Based on Number of Visits of Links of Web Page”, International Journal of Soft Computing and Engineering, vol-2, no.3, pp. 441-446; July, 2012.
- [8] Gerald J.Kowalski, mark T.Maybury; “User Search Techniques”, Information Storage and Retrieval Systems, Kluwer Academic Publisher, pp.174-175, 2000.