

# Comparison of various page Rank Algorithms

Mahesh Joshi

Gujarat University Technological

**Abstract**—Web is expanding day by day and people generally rely on search engine to explore the web. Thus, it has become very important for the sources to give relevant and qualified result. The main aim of this paper is to get the knowledge about various page rank algorithm and to find the optimized result among them. The comparison will be done on the basis of their speed, limitations, benefits and input parameters, efficiency of results.

**Keyword**— Web page ranking , Page range

## I. INTRODUCTION

As the content of information is increasing everyday, it's a challenge to know about them and be updated everywhere and at every time. Fig 1. shows about the working of a search engine. In Front end process, a user enters a query, which passes through search engine interface which in turn is parser by a query parser. Query parser searches the query in index files. In Back end process the WWW, from where the web pages are gained or crawled by a spider and it sends the requested web page to the indexer, which will be stored in index files, and through ranking algorithm, the user will get a result.

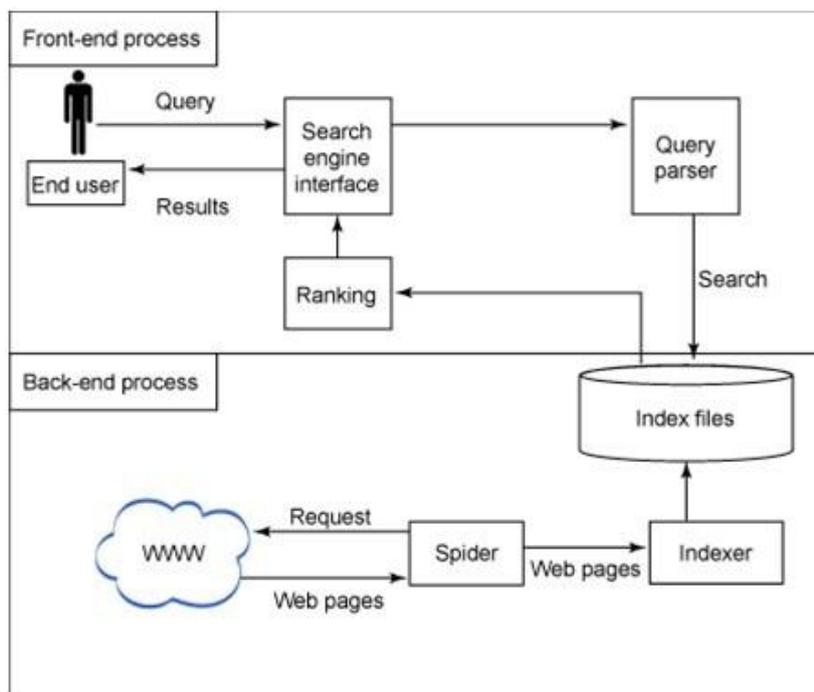


Figure 1-Working of a search engine

## II. MODERN ALGORITHMS

**Google Panda[9]** is a change to Google's search results ranking algorithm that was first released in February 2011. The change aimed to lower the rank of "low-quality sites" or "thin sites", and return higher-quality sites near the top of the search results. Google Panda is a filter that prevents low quality sites and/or pages from ranking well in the search engine result page. The first Panda update,

known as Panda 1.0, was released in February of 2011, affecting nearly 12 percent of all search queries and making it one of the most significant algorithm changes in Google's history. It was intended to weed out low-quality content and reward sites with high-quality, actively updated content.

The next major update was in May of 2014, a rollout known as Panda 4.0, which affected approximately 7.5 percent of all English search queries, adding more features to reward high-quality content and detect instances of low-quality writing.

#### **Panda 4.1**

The rollout began on September 25, 2014, but according to an official Google+ update from Pierre Far, the rollout was a slow, gradual one, planned end in the first week of October. According to Far, the update continues in the tradition of previous Panda updates by identifying low-quality content with more precision. Google has evidently added a few new signals of low quality to its algorithm--but of course, those signals are not public public.

#### **PENGUIN**

Google Penguin is a code name for a Google algorithm update that was first announced on April 24, 2012. The update is aimed at decreasing search engine rankings of websites [10]. Before, Penguin, Panda came into existence, but because of its faults, it was not up to date. Although, after that other versions of Panda were released. The main target of Google Penguin is spamdexing (including link bombing). here is a difference between Panda and Penguin. It is true that many site owners find Panda to be much convenient his is named at two lower ranking for low-quality sites. There is only this real difference between Panda and Penguin that helps all website owners to draw and maximize the benefit of achieving higher rankings. Panda searches for low quality websites and it reduces rankings.

#### **Humming bird**

The name humming bird comes from precise and fast. [11] The importance of long tailed keywords are basically because of this algorithm. It gives priority to the contextual meaning of the terms used in a query.

The new humming Bird updates works in two ways:-

- i) Conversational searches instead of traditional keyword searches.
- ii) Displays search content right on the search pages which makes it easier for user to pick out the right website.

Hummingbird represents a huge shift in Google's search algorithm from a focus on keyword searches to question-based, conversational, semantic search. Semantic search seeks to provide the best possible results based on what Google knows about you, including your network of contacts, previous searches and social shares, current trends, use of connecting words, and your geographical location. The same search can yield different results to each user, based on their unique circumstances and likely intent. Humming bird uses knowledge graph, a knowledge base used by Google to enhance its search engine search result with semantic search gathered from variety of sources.

### **III. COMPARISON OF VARIOUS WEB PAGE RANKING ALGORITHMS**

Algorithm	Panda	penguin	Humming Bird
Main technique	Web content mining	Web content mining	Web content mining with deepen semantic search
Methodology	aimed to lower the rank of "low-quality sites" or "thin sites".	aimed at decreasing search engine rankings of websites that violate Google's Webmaster Guidelines.	Conversational search leverages natural language, semantic search, and more to improve the way search queries are parsed.
Input parameter			Ip address, geo-data or location
Quality of results	medium	Medium	High
Importance	Moderate. Syndication, User Engagement, Indexation & Keyword Hoarding	Used in webspaming	Conversational search can be done.
Limitation	webspaming	Gaining and loosing traffic	The knowledge graph does not incorporate many languages.

### III. CONCLUSION

Depending on the type of algorithms used, it can be decided a definite rank For a relevant web page. Algorithm have taken a major role in gathering or releasing the content, information or data After going through comparative study of various algorithms ,an efficient algorithm should pass all the critical phases by keeping in mind standard web technology.

### REFERENCES

- [1] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", In proceedings of the 2rd Annual Conference on Communication Networks & Services Research, PP. 305-314, 2004.
- [2] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web", Technical Report, Stanford Digital Libraries SIDL-WP-1999-0120, 1999.
- [3] C. Ridings and M. Shishigin, "Pagerank Uncovered", Technical Report, 2002.
- [4] Sergey Brin and Larry Page, "The anatomy of a Large-scale Hypertextual Web Search Engine", In Proceedings of the Seventh International World Wide Web Conference, 1998.
- [5] An Empirical Analysis of Page Ranking Algorithms Dr Paras Nath Gupta<sup>1</sup>, Pawan Singh<sup>2</sup>, Punit Kr Singh<sup>3</sup>, Sandeep Chaudhary<sup>4</sup>, Pankaj P Singh<sup>5</sup>, Deepak Sinha International Journal of Scientific & Engineering Research, Volume 3, Issue 12, December-2012 ISSN 2229-5518
- [6] Ko Fujimura, Takafumi Inoue and Masayuki Sugisaki,, "The EigenRumor Algorithm for Ranking Blogs", In WWW 2005 2<sup>nd</sup> Annual Workshop on the Weblogging Ecosystem, 2005.
- [7] Ali Mohammad Zareh Bidoki and Nasser Yazdani, "DistanceRank: An Intelligent Ranking Algorithm for Web Pages", Information Processing and Management, 2007.
- [8] [www.huffingtonpost.com/jayson-demers/your-guide-to-googles-panda 4.1](http://www.huffingtonpost.com/jayson-demers/your-guide-to-googles-panda-4.1)

[9] [en.wikipedia.org/wiki/google-penguin](http://en.wikipedia.org/wiki/google-penguin)

[10] <http://www.globalmediainsight.com/blog/hummingbird-google-update> hummingbird spawns new search canon.