Abstract

As we know that web is a collection of huge amount of data, it is not very easy to find relevant information. To find the desired data, user visits different web pages. Most Web users typically use a Web browser to navigate a Web site. They start with the home page or a Web page found through a search engine or linked from another Web site, and then follow the hyperlinks they think relevant in the starting page and the subsequent pages, until they have found the desired information in one or more pages. The aim of this work is to study the different characteristics of various ranking algorithms. Here the factors affecting the ranking of pages of a website are considered and it has been studied that how the popularity of a site can be raised and how spam pages can be tracked. Firstly the importance of different characteristics responsible for Page Ranking are determined. Then by taking this information into consideration a technique is developed that successfully distinguishes spam pages from licit pages.

References

- Nidhi Grover, Ritika Wason, “Comparative Analysis Of Pagerank AndHITS
Analysis of Web Pages through Link Structure

Algorithms", (IJERT), ISSN: 2278-0181, Vol. 1 Issue 8, October -2012
- Nan Ma, Jiancheng Guan, Yi Zhao, &quot;Bringing PageRank to the citation analysis&quot;, Information Processing and Management 44 (2008) 800–810.
- Ji-Rong Wen, &quot;Enhancing Web Search through Web Structure Mining&quot; 2009, IGI Global.
- Z. Gyongyi and H. Garcia-Molina, &quot;Web spam taxonomy&quot;, First International Workshop on Adversarial Information Retrieval on the Web, 2005
- Stefano Leonardi,Carlos Castillo,Debora Donato and Ricardo Baeza-Yates, &quot;LinkBased Characterization and Detection of Web Spam&quot;, AIRWEB&apos;06, August 10, 2006, Seattle, Washington, USA.

Index Terms

Computer Science
Information Sciences

Keywords
PageRank, Inbound Links, Outbound Links, Spam Page