Abstract

As the web is escalating day by day, so the most concerned issue for the users would be how to collect the useful information and to find their genuine information effectively and quickly. With the tremendous growth of information available to end users through the web, search engines play a vital role in retrieving and organizing relevant data for various purposes. The ranking of the web pages for the web search engine is one of the significant problems at present. This leads to the important attention to the research community. In this paper, a page rank mechanism called Hybrid Page Rank Algorithm is proposed which is based on both content and link structure of the web pages. This algorithm is used to find more relevant information according to user's query. This paper also presents the comparison between SimRank Algorithm and the Hybrid Page Rank Algorithm.

References

- Companion slides for the text by Dr. M. H. Dunham, "Data Mining: Introductory and Advanced Topics", Prentice Hall, 2002.
- Jaroslav Pokorny, Jozef Smizansky, "Page Content Rank: An Approach to the Web Content Mining".
- Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR'apos;04), 2004 IEEE.
- Zhao, C., Zhang, Z., Li, H., Xie, X., "A Search Result Ranking Algorithm Based
on Web Pages and Tags Clustering"; Published in IEEE, Print ISBN No: 978-1-4244-8728-8, 2011, pp. 609-614.


- Hyperlink Analysis: Techniques and Applications Prasanna Desikan, Jaideep Srivastava, Vipin Kumar, and Pang-Ning Tan, Department of Computer Science, University of Minnesota, Minneapolis, MN, USA {desikan, srivastava, kumar, ptan} @cs.umn.edu.


- Alta Vista Search Engine; http:// www. altavista. com


Index Terms

Computer Science
Algorithms

Keywords

WWW; Data mining; Web mining; Search engine; Page ranking