Abstract

Information Retrieval deals with searching and retrieving information within the documents and it also searches the online databases and internet. Web crawler is defined as a program or software which traverses the Web and downloads web documents in a methodical, automated manner. Based on the type of knowledge, web crawler is usually divided in three types of crawling techniques: General Purpose Crawling, Focused crawling and Distributed Crawling. In this paper, the applicability of Web Crawler in the field of web search and a review on Web Crawler to different problem domains in web search is discussed.
Web Crawler: A Review

- Infoseek, Mar. 2008. URL www.infoseek.co.jp
Web Crawler: A Review


- Rajashree Shettar, Dr. Shobha G, &quot;Web Crawler On Client Machine&quot;,
- Ashutosh Dixit and Dr. A. K. Sharma, &quot;A Mathematical Model for Crawler Revisit Frequency&quot;,

- Shruti Sharma, A. K. Sharma and J. P. Gupta &quot;A Novel Architecture of a Parallel Web Crawler&quot;,
- Alex Goh Kwang Leng, Ravi Kumar P, Ashutosh Kumar Singh and Rajendra Kumar Dash &quot;PyBot: An Algorithm for Web Crawling&quot;,
IEEE 2011
- Song Zheng, &quot;Genetic and Ant Algorithms Based Focused Crawler Design&quot;,
- Lili Yana, Zhanji Guia, Wencai Dub and Qingju Guoa &quot;An Improved PageRank Method based on Genetic Algorithm for Web Search&quot;,
Procedia Engineering, pp. 2983-2987, Elsevier 2011
- Bahador Saket and Farnaz Behrang &quot;A New Crawling Method Based on AntNet Genetic and Routing Algorithms&quot;,
PROCEEDINGS OF ICETECT, pp. 926-932, 2011
pp. 562-566, IEEE 2012

3 / 5

Index Terms

Computer Science Information Sciences

Keywords

WWW Web Crawler Crawling techniques Web Crawler Survey Search engine Parallel Crawler