Abstract

Today, almost 60% of the web is dynamic in nature. Web crawler visits web pages repeatedly to check whether the web page changed or not in order to maintain the index of search engine up-to-date, to provide fresh information to users. This consumes a lot of network bandwidth and time. A lot of mechanisms have been reported in the literature. In this paper, a novel algorithm of web page change detection is proposed. The proposed algorithm is based on detecting web page change based on hybrid i.e. it detects both structural and content changes.

References

A Novel Algorithm of Web Page Change Detection

Page change detection based on Segmentation. JKSU.

Index Terms

Computer Science
Information Sciences

Keywords

Content changes, Structural changes, Web page, Web Crawler, Change detection.