Abstract

Internet is actively used for the exchange of information. People upload the web pages and updating the new web pages very frequently. There is a frequent change in the content of the web page hence it become necessary to develop an efficient system which could detect these changes efficiently and in the minimum browsing time. So as to achieve this we compare the old web page and the new web page. Changes in a web page can be detected with the use of various algorithms. Various tools and services are also available which can be used to detect these changes. In this paper a new algorithm for the structural as well as content change detection has been proposed and described. For better results tree has been designed for the corresponding web pages. The proposed change detection algorithm is based on assigning hash value to each leaf node and tag value to the non leaf nodes. Bottom up approach has been used for assignment. The level of each node has been used to find hash values and modification in a node. It has been shown with the help of suitable examples that the proposed algorithm extracts the changes very efficiently from the various web pages.

References
An Efficient Algorithm for Web Page Change Detection


- Leonardi E. , Sourav Bhownick S. , "Detecting Content Changes on Ordered XML Documents Using Relational Databases".


Index Terms

Computer Science Information Sciences

Keywords

Web Page Change Detection Tag Of Node Tree Matching Hash Value Change Monitoring