Abstract

With the tremendous growth of the Internet, many web pages are available online. Search engines use web crawlers to collect these web pages from web for the purpose of storage and indexing. Most of the web pages are autonomous and are updated independently of the users that access the sources. As the web pages are updated autonomously, users are unaware of how often the sources change. An incremental crawler visits the web repeatedly after a specific interval for updating its collection. Users can be benefitted by knowing the page importance based upon the page accessing frequency. In this paper the page relevancy is finding out by a novel mechanism and a novel architecture is being proposed.

References

Volume 8– No. 7, October 2010.


Index Terms

Computer Science Information Sciences
Keywords