Enhance Crawler: A Dual-Stage Crawler for Efficiently Harvesting Deep Web Interfaces

International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

Volume 168 - Number 8

Year of Publication: 2017

Authors:
Sujata R. Gutte, Shubhangi S. Gujar

Abstract

Internet is became important part of our day to day life. That’s why due to heavy usage of internet very large amount of diverse data is spread over it and which provide access to search particular data. Very challenging issue for search engine is ‘fetch most relevant data as per users need. So to reduce large amount of time spend on searching most relevant data as per user’s need. We proposed the “enhanced crawler “In this proposed approach framework is divided into two stages. in the first stage , for center pages search engine perform site based searching for getting more accurate result of focus crawler (it avoid to visit large no of pages ) And ranking is used for prioritize highly relevant ones for given input topic. IN second stage of framework, In-site searching is done by extracting most relevant links with an adaptive link ranking. We design link tree data structure to achieve wider coverage of deep website.

References

Enhance Crawler: A Dual-Stage Crawler for Efficiently Harvesting Deep Web Interfaces


Index Terms

Computer Science

Information Sciences

Keywords

Enhance crawler, deep website, ranking, adaptive learning.