Abstract

Extensible Markup Language (XML) is a simple text format which was designed to describe data using custom tags. The use of custom tags makes XML extremely flexible and enables it to not only describe structured data like information from a table of relational database but also semi-structured data. An XML document is self-describing which has made it a standard means of data exchange between applications and for use in configuration files of enterprise applications. The increasing preference to store and transmit data in the XML format has led to a need for searching these data stores for information. Query languages like Xpath and XQuery are used to retrieve information from xml document. But these query languages are complex for non expert user to learn. Keyword search allows such user to retrieve information without knowledge of complex query language. In this paper we proposed an algorithm for relevance ranking of nodes which retrieved as result by considering keyword ambiguity and intension of user.

References

- S. Boag, D. Chamberlin, and M. F. Fernandez. XQuery 1.0: An XML query language.
 Keyword Search on XML Repository with Relevance Ranking

- Ziyang Liu, Jeffrey Walker, Yi Chen “XSeek: A Semantic XML Search Engine Using Keywords” xseek.asu.edu/xseekdemo.pdf
- Arash Termehchy “Keyword and Natural Language Query Processing for Semi-Structured ata Sources” , IDAR 2009.
- Yi Chen, Wei Wang, Ziyang Liu, Xuemin Lin “Keyword Search on Structured and Semi-Structured Data”, SIGMOD’09.

Index Terms

Computer Science Information Retrieval

Keywords
XML Natural Processing Keyword Search Relevance Ranking