Abstract

XML is recognized as a standard for data storage and exchange for web applications. This is because it has certain unique features like it is self-describing, extensible and it is stored in the form of text documents. In spite of all these unique features, XML has an inherent limitation of verbosity. Because of the strong presence of XML in database technology and its inherent verbosity, there is an ever-increasing need to design compact storage for XML which can be effectively utilized for efficient indexing and querying of XML. The proposed technique creates a structure index, which is a compact summarization of the XML document, and a data index which groups and stores the contents of all similar paths at one place. Based on this compact storage, a novel query algorithm is proposed which can answer XPath queries very efficiently. This approach dramatically reduces the storage requirement for XML coupled with efficient processing of XPath queries. The implementation of this technique and comparison with other techniques confirm our claim.

References

Efficient Querying of Structure and Contents for XML Documents


- A. Arion, A. Bonifati, G. Costa, S. D'aguzzo, M. Manolescu, and A. Pugliese. "XQueC: Pushing queries to compressed XML data", in Proceedings of the 29th International Conference on Very Large Data Bases (VLDB'03), 2003.


- Radha Senthilkumar, Priyaa Varshinee and A. Kannan. "Designing and Querying a

Index Terms
Computer Science Information Sciences

Keywords
Compact Storage Structure Index Content Index