

Abstract

Nowadays, XML has shown to be the standard for electronic data interchange on the Internet. Whenever the data size or energy consumption limitations form the bottleneck of an XML based application, these applications can profit from the usage of XML compression. It is desirable that these applications can perform all XML based operations directly on the compressed XML data, to avoid additional computation caused by prior decompression and subsequent compression.

In this thesis, I present three basic approaches to the compression of XML structure as well as two combinations of these approaches that support query evaluation as well as update operations directly on the resulting compressed data. These approaches are amended by a generic, efficient XPath evaluator.

Finally, I have performed extensive performance evaluations to compare the presented compression approaches with other available approaches to XML compression, and I show that the presented compression approaches outperform the other approaches in terms of query evaluation und updates on the compressed data.

When comparing the newly developed approaches to XML compression with each other, it can be seen that each approach shows its strength in different aspects: While one approach reaches a strong compression, another approach shows fast compression and decompression times, whereas the third approach allows efficient query evaluation. None of the presented approaches outperforms the other approaches completely.