Ramanan, Prakash

Worst-case optimal algorithm for XPath evaluation over XML streams. (English)

Zbl 1192.68187

Summary: We consider the XPath evaluation problem: Evaluate an XPath query $Q$ on a streaming XML document $D$; i.e., determine the set $Q(D)$ of document elements selected by $Q$. We mainly consider Conjunctive XPath queries that involve only the child and descendant axes. Previously known in-memory algorithms for this problem use $O(|D|)$ space and $O(|Q||D|)$ time. Several previously known algorithms for the streaming version use $\Omega(d^n)$ space and $\Omega(d^n|D|)$ time in the worst case; $d$ denotes the depth of $D$, and $n$ denotes the number of location steps in $Q$. Their exponential space requirement could well exceed the $O(|D|)$ space used by the in-memory algorithms. We present an efficient algorithm that uses $O(d|Q| + nc)$ space and $O((|Q| + dn)|D|)$ time in the worst case; $c$ denotes the maximum number of elements of $D$ that can be candidates for output, at any one instant. For some worst case $Q$ and $D$, the memory space used by our algorithm matches our lower bound proved in a different paper; so, our algorithm uses optimal memory space in the worst case.

MSC:

68P05 Data structures
68P15 Database theory

Keywords:
XML; xpath; query evaluation; stream processing

Software:
XPath

Full Text: DOI

References:


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.