

UC San Diego

Technical Reports

Title

Efficient Rewriting of XPath Queries Using Query Set Specifications

Permalink

<https://escholarship.org/uc/item/0zm0n95m>

Authors

Cautis, Bogdan

Deutsch, Alin

Onose, Nicola

et al.

Publication Date

2009-03-19

Peer reviewed

Efficient Rewriting of XPath Queries Using Query Set Specifications

Bogdan Cautis Alin Deutsch Nicola Onose
Telecom ParisTech UC San Diego UC San Diego
cautis@telecom-paristech.fr deutsch@cs.ucsd.edu nicola@cs.ucsd.edu

Vasilis Vassalos

Athens Univ. of Economics and Business
vassalos@aueb.gr

Abstract

We study the problem of querying XML data sources that accept only a limited set of queries, such as sources accessible by Web services which can implement very large (potentially infinite) families of XPath queries. To compactly specify such families of queries we adopt the Query Set Specifications [1], a formalism close to context-free grammars.

We say that query Q is *expressible* by the specification \mathcal{P} if it is equivalent to some expansion of \mathcal{P} . Q is *supported* by \mathcal{P} if it has an equivalent rewriting using some finite set of \mathcal{P} 's expansions. We study the complexity of expressibility and support and identify large classes of XPath queries for which there are efficient (PTIME) algorithms. Our study considers both the case in which the XML nodes in the results of the queries lose their original identity and the one in which the source exposes persistent node ids.

References

- [1] M. Petropoulos, A. Deutsch, and Y. Papakonstantinou. The Query Set Specification Language (QSSL). In *WebDB*, pages 99–104, 2003.