Getahun, Fekade; Chbeir, Richard
RSS query algebra: towards a better news management.  (English) Zbl 1320.68067

Summary: Existing XML query algebras are not fully appropriate to retrieve RSS news items mainly due to three reasons: (1) RSS document is text rich and its content is dependent on the wording and verification of the author, thus semantic-aware operators are needed; (2) news items are dynamic and consequently time oriented retrieval is needed; and (3) a news item may evolve through time, or overlap with other news items and hence identifying relationships between items is needed. In this paper, we aim to solve these issues by providing a dedicated RSS algebra based on semantic-aware operators which are capable of considering RSS characteristics. The provided operators are application domain specific and can be tuned according to the user preferences. We also provide a set of query rewriting and equivalence rules that would be used during query simplification and optimization. In addition and in order to validate our proposal, we develop a prototype called EasyRSSManager that allows a user to formulate RSS query using our operators.

MSC:
68P20  Information storage and retrieval of data

Keywords:
RSS retrieval; algebraic operator; semantic query; query rewriting; similarity function; equivalence rule

Software:
Quilt; TAX; XPath; WordNet; SPARQL; XSEarch

Full Text: DOI

References:

I. Manolescu, D. Florescu, D. Kossmann, F. Xhumari, D. Olteanu, Agora: living with XML and relational, in: Proceedings of the 26th Int. Conf. on Very Large Data Bases (VLDB), Cairo, Egypt, 2000, pp. 623-626.


