Graduiertenkolloquium Angewandte Informatik

Rank-aware, approximate query processing on the Semantic Web

Dipl.-Inform.Wirt. Andreas Wagner

AIFB

Abstract:
The amount of data on the WWW that adheres to Semantic Web standards is rapidly increasing. Search over this huge Web data corpus frequently leads to queries having large result sets. So, in order to discover data elements, which satisfy a given information need, users must rely on ranking techniques to sort results according to their relevance. Unfortunately, processing queries with ranked results over a large data corpus is highly expensive in terms of computation time as well as computation resources. At the same time, applications often face information needs, which do not require complete and exact results.

In this thesis, we face the problem of how to process queries over Web data in an approximate and rank-aware fashion. Aiming at this complex problem, we provide several novel contributions.

More specifically, we introduce a rank-aware join operator for Web data. By means of this join operator, we can process queries with ranked results much more efficiently. That is, our rank-aware join operator focuses on computing the top-ranked query results first, while omitting the remainder of the results.

Additionally, we enable systems to trade off result completeness and accuracy, in favor of query computation time. We provide two contributions for this approximate query processing. On the one hand, we present a novel pipeline of operations, which allows to incrementally compute query results. On the other hand, we introduce a new approximate rank-aware join operator. Our operator allows discarding such intermediate query results, which are not likely to lead to a final top-ranked result.

Furthermore, we present a novel approach for selectivity estimation that is tailored towards the needs of Web data and typical Web queries. That is, our selectivity estimation approach allows the estimation of queries, which match structured as well as unstructured data elements in the Web of data. Such selectivity estimation is crucial for query optimization techniques, which can integrate our approximate/rank-aware join operators in physical query plans.

Termin: Mittwoch, 09. April 2014, 15.45 Uhr

Ort: Englerstraße 11, 76131 Karlsruhe
               Kollegiengebäude am Ehrenhof (Geb. 11.40), 2. OG, Raum 231
               (Hinweise für Besucher: www.aifb.kit.edu/web/Kontakt)

Veranstalter: Institut AIFB, Forschungsgruppe Wissensmanagement

Zu diesem Vortrag lädt das Institut für Angewandte Informatik und Formale Beschreibungsverfahren alle Interessierten herzlich ein.

Andreas Oberweis, Hartmut Schmeck, Detlef Seese, Wolffried Stucky, Rudi Studer (Org.), Stefan Tai

Besucheranschrift:
KIT-Campus Süd
Institut AIFB – Geb. 11.40
Kaiserstr. 89
76133 Karlsruhe

Postanschrift:
KIT-Campus Süd
Institut AIFB – Geb. 11.40
76128 Karlsruhe