

SQL Query Recommendations

Publicly available databases become popular nowadays. They provide numerous ways of how to access the underlying data, including formulating queries directly to a database on SQL. This way, though the most flexible, rises the problem of formulating user needs properly: they could not have sufficient experience in (1) the domain or (2) SQL. For those users providing recommendations in a form of SQL requests is beneficial.

To give a query recommendation one needs first to find similar behavior among the requests of the others. In order to do so the notion of query similarity is required. A similarity function is based on a query representation that defines which meaningful features of an SQL request needs to be compared. There are a few query representations and corresponding query representations existing:

- The feature-based (FB) representation focuses on the query structure: which attributes, tables, UDFs and filtering conditions (without values) a query consists of.
- The witness-based (WB) representation relies on the result of a query to a database.
- The access area-based (AAB) representation captures the area of a data space that a user is interested in.

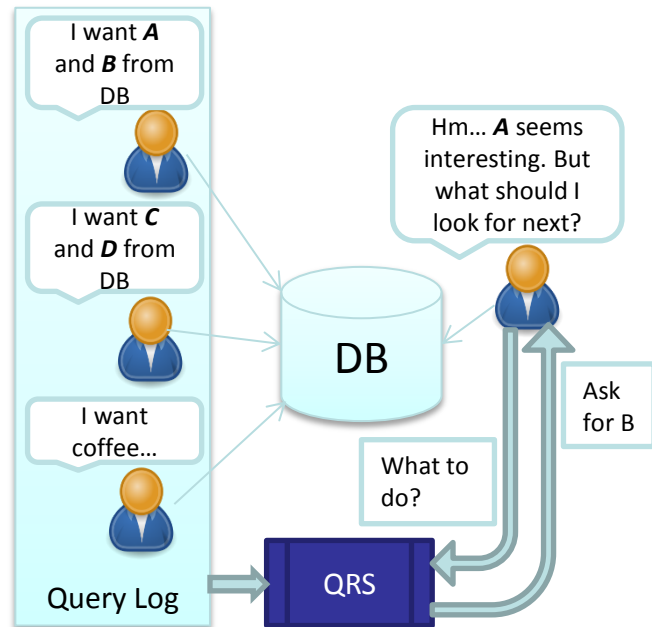
The master thesis implies the following tasks:

- Review of classical approaches of giving recommendations;
- Study a real-world data **query log from the SkyServer project**: identifying patterns of user behavior within it;
- Implement of query recommendation system for listed query representations and corresponding similarity functions;
- Conduct the experiments and evaluate the results.

In this work you will learn about the approaches of giving recommendations, apply them on the area of SQL requests. You will gain skills in development and evaluation of recommendation systems in the domain of SQL queries.

You are:

- able to provide your own ideas;
- familiar with Java, at least you learn quickly;
- has strong knowledge in databases, including query optimization techniques;
- starting with your master thesis not earlier than October, 2017.



Contact

Natalia Arzamasova natalia.arzamasova@kit.edu Room: 362

Am Fasanengarten 5 76131 Karlsruhe

Building: 50.34