Measuring Web Search Quality Without Click Data

In order to seamlessly provide good results, web search engine operators need to be able to tell the quality of search results without asking users. Commonly, this is achieved by invisibly examining user behaviour.

Examining user behaviour to determine result quality is called **implicit feedback**. A change is made to the way the search engine works for a number of users, and their behaviour recorded. Changes in behaviour are then used to implicitly determine changes in search result quality.

**Click data** (information about the individual results that users have clicked on) is a well studied and effective source for implicit feedback. However, in our studies we observe that nearly two thirds of web search queries do not any have clicks. For this reason, it is important to investigate other implicit feedback measures.

We conducted a study with 140 users. Each user completed six search tasks, three with a commercial search engine, and three with an intentionally degraded version of the same search engine. User behaviour was heavily logged. We then examined differences between user behaviour on the normal and degraded engines.

By far the best new indicator is **scrolling**, shown to the right.

The plot to the left shows using mean number of clicks over all queries as an implicit feature.

The mean clicks per query from the normal engine was subtracted from the mean clicks per query from the degraded engine (Y axis). If a user is below the X axis, there are more clicks on the normal engine.

This correctly predicts result quality 72% of cases, as predicted by other work.

The plot to the right shows using mean number of scrolls over all queries as an implicit feature.

This correctly predicts result quality for 69% of users, as accurately as clicks. Users scroll more when results are bad, possibly because they’re examining the results more closely.

This is an exciting discovery, as scrolling can be used in all searches (rather than 1/3 of searches with click data), and is relatively unstudied.

Timothy Jones (Student)
David Hawking (Supervisor)
Ramesh Sankaranarayana (Supervisor)