More Than Relevance: High Utility Query Recommendation By Mining Users’ Search Behaviors

Xiaofei Zhu, Jiafeng Guo, Xueqi Cheng, Yanyan Lan
Institute of Computing Technology, Chinese Academy of Sciences
zhuxiaofei@software.ict.ac.cn, {guojiafeng, cxq, lanyanyan}@ict.ac.cn

1. MOTIVATION

- Relevant query recommendation: Providing alternative queries similar to a user’s initial query.
- Assumption: relevant query ≠ relevant results, satisfy users needs.

2. KEY IDEA

- How to infer query utility?
- Key Idea: Through user’s search behaviors

- A typical search session

3. QUERY UTILITY MODEL

- Relevant Utility α control the probability of the attractiveness
- Posterior Utility β control the probability of users’ satisfaction

4. EVALUATION METRICS

- QRR (Query Relevant Ratio)
  \[ QRR(q) = \frac{R(q)}{N(q)} \]
  Measuring the probability that a user finds (clicks) relevant results when she uses query \( q \) for her search task.

- MRD (Mean Relevant Document)
  \[ MRD(q) = \frac{D(q)}{N(q)} \]
  Measuring the average number of relevant results a user finds (clicks) when she uses query \( q \) for her search task.

5. EXPERIMENTAL RESULTS

- Contribution
  - Recommend high utility queries rather than only relevant queries: to directly toward the ultimate goal of query recommendation.
  - A novel dynamic Bayesian network (i.e., QUM) to mine query utility from users’ reformulation and click behaviors.
  - Introduce two evaluation metrics for utility based recommendation
  - Evaluate the performance on a real query log and show the effectiveness

- Future work
  - Extend our utility model to capture the specific clicked URLs for finer modeling.
  - Extend our utility model to capture the query level utility.

The performance improvements are significant
(t-test, p-value < 0.05)