KB-Enabled Query Recommendation for Long-Tail Queries

Zhipeng Huang*, Yudian Zheng*, Bogdan Cautis$, Reynold Cheng*
*The University of Hong Kong, $Huawei Noah’s Ark Lab

Query Recommendation

Recommend queries to a search engine user who issues query q.

**Long-tail queries:** queries that are seldomly issued by users.

![Frequency Distribution of Queries](chart)

**Example:**
- q₁: “akira kurosawa influence george lucas”
- q₂: “hidden fortress star wars comparison”

Existing Methods

**Query Flow Graph (QFG):**
A directed graph of queries, with weights on edges modeling the probability of co-occurrence of the two queries.

To recommend queries for q:
1. Find the queries q’ that have largest weight w(q, q’).

**Term-Query Flow Graph (TQGraph):**
A directed graph of queries of terms.

To recommend queries for q:
1. Find all the terms within q as `term(q)`;
2. For each t in `term(q)`, we perform a Personalized PageRank (PPR) from t.
3. Aggregate the results of all PPR, and recommend the queries with largest weights.

**Weaknesses:** Rely totally on query logs! Cannot handle long-tail queries well.

Knowledge Base

**KB:** a set of facts (s, t, r), meaning that entity s and entity t has relationship r. We model a KB as a directed graph G_{KB} of entities, with different types of edges.

Meta Path [3]

**Meta Path:** a sequence of node types and edge types.

\[ P_1: \text{Person} \rightarrow \text{Person} \]

When an entity is queried by a user, we may use a meta path to retrieve related entities to the users!

For example, if “barack obama” is queried, we can use the meta path P above, to recommend “michelle obama”, we can also use meta path P₂ to recommend “hillary clinton”.

The problem is: how to set the weight to different meta paths.

KB-QRec: QFG + KB

KB-QRec: a quadruple \((G_{qf}, K, t_{EQ}, P)\), where \(t_{EQ}\) denotes the transition probability from entities to queries, and \(P\) is the set of meta path we learn from query logs with corresponding weights.

Learn P:
1. \[ t_{EQ}(e,e') = 1 - \prod_{(e,e') \in EQ} (1 - t_{EQ}(e \rightarrow e')) \]
2. Find the shortest path between e and e’ in K.
3. Accumulate the weights to the corresponding meta path.

**To recommend queries (three steps):**
1. Entity linking;
2. Entity expansion;
3. Query searching.

q₁: “akira kurosawa influence george lucas”

Experiments

![Experiments](chart)

References