Can We Predict User Intents from Queries?

- Intent Discovery for Web Search -

Katsumi Tanaka

Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo, Kyoto 606-8501, Japan
tanaka@dl.kuis.kyoto-u.ac.jp
http://www.dl.kuis.kyoto-u.ac.jp

Abstract. Although Web search engine technologies have made a great progress in recent years, they are still suffering from the low search performance (precision and recall) because of the following reasons:

(1) Queries for search engines are mostly limited to keywords or short natural language sentences, and

(2) Most search engines use traditional "keyword-in-document" information retrieval models.

Obviously, a user's search intent is not easily expressed by a set of keyword terms. A same keyword-query is formulated and executed by many users, but its search intents (e.g. what information are the "really relevant" answers for the users) are different from users. Also, the traditional "keyword-in-document" IR model assumes that query keywords (and/or related keywords) are contained in the target documents (Web pages). For example, it makes difficult to search for documents (Web pages) whose reputation are specified in user queries.

Search intent discovery is a hot research area in Web search, such as search query classification (informational, navigational and transactional queries), search result diversification, and query recommendation.

In this talk, after a brief survey on the research of search intent discovery and query type classification, we introduce a new framework on search intent discovery and intent-based Web search. In our framework, search-intents are roughly classified into four types: (1) content-related intents (topic-relevance, diversity, comprehensibility, concreteness etc.), (2) task-related intents (search for doing some actions), (3) "social" intents (popularity, typicality, novelty/unexpectedness etc.), and (4) aggregation-based intents (such as retrieving the most expensive Kyoto foods).

Then, we survey our research activities to discover "search-intent types" for user search queries. The proposing methods are based on the usages of ontological knowledge, user behavior data analysis, knowledge extracted from CQA corpus & ads, and "relevance" feedback by intent-based page features.

Keywords: Web search and mining, intent discovery, social data, user behavior analysis.

W. Meng et al. (Eds.): DASFAA 2013, Part I, LNCS 7825, p. 2, 2013.