Anil Armagan
Fatma Balci
FUAT BASIK

A TAXONOMY OF WEB SEARCH

Classic IR is inherently predicated on users searching for information, but the need behind a web search is often not informational – it might be navigational or transactional. In fact, informational requires constitute less than 50% of web searches.

Aim of the paper

1) Point out the differences between search intents
2) Introduce and analyze a taxonomy of web searches
3) Show how search engines evolved to deal with these web-specific needs

The classical model for information retrieval, augmented for the web

First of all, there is a user who has a task. To complete his task the user needs to reach information. The user should explain his need in verbal form and construct a query in some query language. The query is submitted to a system that selects from a collection documents (we called it as corpus), those documents that match the query as indicated by certain matching rules. A query refinement process might be used to create new queries and/or to refine the results.

A taxonomy of web searches

We classify web queries according to their intent into 3 classes:

1) **Informational:** The purpose of such queries is to find information assumed to be available on the web in a static form. By static form we mean that the target document is not created in response to the user query. Many informational queries are extremely wide, while some are narrow.
2) **Navigational:** The purpose of such queries is to reach a particular site that the user has in mind, either because they visited it in the past or because they assume that such a sita exists. With respect to evaluation, navigational queries have usually one “right” result.
3) **Transactional:** The purpose of such queries is to reach a site where further interaction will happen. This interaction constitutes the transaction defining these queries. Binary judgement might be all have, appropriate or not.

Statistics

1) **User survey:** For the user survey it is used a pop-up window. Queries that are neither transactional, nor navigational, are assumed to be informational. It could not be found a simple question to distinguish between transactional and informational queries. Instead, it is identified some of the most popular transactional queries.
2) **Log analysis:** Queries that were neither transactional, nor navigational, were assumed to be informational in intent.

The evolution of search engines

We identify three stages in the evolution of web search engines:

1) **First Generation:** It supports mostly informational queries. It is very closely to classic IR.
2) **Second Generation:** It supports both informational and navigational queries. It uses link analysis, anchor text, click-through data.

3) **Third Generation:** It is emerging now, attempts to blend data from multiple sources in order to try to answer ‘‘the need behind the query’’. The aim is to support informational, navigational, and transactional queries.

Conclusion

An understanding of this taxonomy is essential to the development of successful web search. Current search engines deal well with informational and navigational queries, but transactional queries are satisfied only indirectly and hence a third generation in search engines is emerging.