Dynamic Personalized Ranking of Facets for Exploratory Search

Esraa Ali
ADAPT Centre
School of Computer Science and Statistics,
Trinity College Dublin
Dublin, Ireland
esraa.ali@adaptcentre.ie

ABSTRACT

Faceted Search Systems (FSS) have gained prominence in research as one of the exploratory search approaches that support complex search tasks. They provide facets to educate users about the information space and allow them to refine their search query and navigate back and forth between resources on a single results page. When the information available in the collection being searched across increases, so does the number of associated facets. This can make it impractical to display all of the facets at once.

To tackle this problem, FSS employ methods for facet ranking. Ranking methods can be based on the information structure, the textual queries issued by the user, or the usage logs. Such methods reflect neither the importance of the facets nor the user interests. I focus on the problem of ranking facets from knowledge bases (KB) and Linked Open Data (LOD). KB have the advantage of containing high quality structured data. With the increasing size and complexity of LOD datasets, the task of deciding which facets should be manifest to the user, and in which order, becomes more difficult.

Facets are usually extracted from entity attributes. In reality, entities belong to several categories, as a result they inherit attributes from their parent categories. Determining which attributes best characterize a set of entities is not a trivial task. For example, the attributes related to Barack Obama as a U.S. president are likely to be more important than those related to him being an author.

In addition, the relevance of attributes or facets can be subjective. It varies not only from one person to another but also for the same person from one situation to another. Users' knowledge, interests and therefore search needs evolve with time. Maintaining an updated user profile and reflecting that in the search results adds complexity to the problem. Moreover, the idea of personalizing exploratory search can be challenging and tricky, since personalization in IR (specifically precision-oriented search engines) implicitly implies narrowing and focusing the information space to retrieve the most relevant results according to the users' interests and desires. On the contrary, exploratory search systems are typically recall-oriented and they favor covering as much from the information space as possible. They also encourage diversifying the user knowledge to help them learn and discover the unknown.

The generation of a ranked list of facets should be a dynamic process for a number of reasons. First of all, manually setting up

facets is a time consuming task which relies upon domain experts. Second, it is not practical on large, multi-domain datasets. Even one-off automatic facet generation and ranking might not be suitable for data that changes and grows over time. Lastly, the relevance of facets can be user, query and context dependant.

I am proposing a personalized approach to the dynamic ranking of facets. The approach combines different sources of information to recommend the most relevant facets. The first source is the knowledge-base from which the facets are originally generated. The second is facets generated from the top-ranked documents in a search system. The user search query is submitted to a general search engine and the top ranked documents are used to add context to the ranking process. Finally, the third source uses a user interests profile, which is collected from social media and the user's behavior in the system. These sources contribute to the final ranking score to reflect the importance of facets without ignoring user interests.

My proposed research aims to answer the following **research questions**: **RQ1**: To what extent does the addition of features from retrieved search results from a general web search improve the computation of facet relevance?. **RQ2:** What is the most effective method to incorporate personal interests and user usage data into the ranking process?. **RQ3:** Does personalising facet ranking have a measurable impact upon the user search experience?

The **primary area of contribution** of this work is the proposal of a personalised ranking model for facets generated from a knowledge base. This model also incorporates the user's interests, and features from the top-ranked search results using the same query submitted to a general web search engine, into the ranking process. The area of contribution is novel and not explored by earlier research. Another potential contribution is the development of an evaluation approach and a benchmark dataset to validate such a faceted search system.

CCS CONCEPTS

• **Information systems** → *Information retrieval diversity*; *Rank aggregation*; **Personalization**;

KEYWORDS

Exploratory Search, Personalization, Faceted Search, Knowledge Bases