An Approach for Generating Query Facets using Knowledge Bases

B.Divya, Guided by -Prof.D.Anji Reddy
1.M.Tech Scholar, Department of CSE, Vaageswari College of Engineering Karimnagar, Telangana, India – divya.bonagiri95@gmail.com.
2.Associate Professor, Department of CSE, Vaageswari College of Engineering Karimnagar, Telangana, India – anjireddy.knr@gmail.com.

ABSTRACT:

In this article, we discuss how to use a high-quality structured data base of knowledge to generate search queries. However, previous studies have shown that many users are not satisfied with this traditional search result page [1], [2], [3]. Users often need to look at many documents and see how to summarize the information they seek, especially when they want to understand topics covering different aspects. It usually takes a lot of time and makes the user difficult. Automatic summary of search results can help users save time by understanding queries without browsing across multiple pages. Using a search engine, users can quickly find web pages containing the information they want by sending queries and receiving search results that consist of "ten blue links".

Index Terms - Query Facets, Knowledge Bases, Query Dimensions

1. INTRODUCTION:

Mining query facets (or query dimensions) is an emerging approach to solve the problem above. Pentax is a Japanese camera brand. Its query facets cover aspects about related camera brands, Pentax’s SLR cameras, Pentax’s small digital cameras, and different kinds of optical devices. These query facets help users learn about the topic—Pentax—and at the same time, users can further narrow down their information needs based on these facets.

The facets constructed by the two methods are further merged and ranked to generate final query facets. More specifically,

(1) Facet Generation: We propose directly mining query facet candidates from Freebase. Given a query, we first retrieve relevant entities from Freebase, then obtain
all the properties of these entities. For example, for the query—Beijing subway, we first retrieve entity Beijing Subway and its properties.

(2) **Facet Expansion**: We use QDMiner to mine initial query facets, and then use Freebase to expand these facets to get similar items. We propose two different ways to expand a facet. First, we try to assign each facet to a suitable property of the entities corresponding to the query, and add the target entities of the property to enrich the facet. We denote this method as property based facet expansion. For example, for the query —Michael Jackson, an initial facet mined by QDMiner is comprised of his compositions —you rock my world, butterflies, man in the mirror, thriller, cry, etc. We find that entity Michael Jackson has a property Nominated Works which covers most items in this facet, hence we could use other target entities of this property to enrich this initial facet. When no relevant entities or properties are retrieved, we use the second way —type based facet expansion. We find a common type that covers most facet items. For example, —Kristen Stewart, Robert Pattinson, Taylor Lautner, etc. is a facet of query —Eclipse. The best type covering most facet items in Freebase is Celebrities.

(3) **The facet candidates** constructed by facet generation and expansion are further merged, because there might be duplicate items within these candidates. We then re-weight the final facets by checking the occurrence of the facet items within top search results. We denote the solution above which generates new facets and expand existing facets using Freebase with QDMKB in this paper. Please note that we actually leverage the advantages of both knowledge bases and search results to generate high-quality query facets, hence QDMKB has high potential to outperform the state-of-the-art algorithms which solely use search results for facet mining.

2. **Existing System:**

Existing query facet mining algorithms mainly rely on the top search results from search engines.

- Dou et al. first introduced the concept of query dimensions, which is the same concept as query facet discussed in this paper. They proposed QDMiner, a system that can automatically mine query facets by aggregating frequent lists contained in the results. The lists are extracted by HTML tags (like <select> and
<table>

>, text patterns, and repeat content blocks contained in web pages.

- Kong et al. proposed two supervised methods, namely QF-I and QF-J, to mine query facets from the results.
- In all these existing solutions, facet items are extracted from the top search results from a search engine (e.g., top 100 search results from Bing.com). More specifically, facet items are extracted from the lists contained in the results.

3. DISADVANTAGES OF EXISTING SYSTEM:

- Many users are not satisfied with this kind of conventional search result pages.
- This usually takes a lot of time and troubles the users.
- The problem is that the coverage of facets mined using this kind of methods might be limited, because some useful words or phrases might not appear in a list within the search results used and they have no opportunity to be mined.

**PROPOSED SYSTEM:**

We propose leveraging a knowledge base as a complementary data source to improve the quality of query facets. Knowledge bases contain high-quality structured information such as entities and their properties and are especially useful when the query is related to an entity. We propose using both knowledge bases and search results to mine query facets in this paper. The reason why we don’t abandon search results is that search results reflect user intent and provide abundant context for facet generation and expansion.

Our target is to improve the recall of facet and facet items by utilizing entities and their properties contained in knowledge bases, and at the same time, make sure that the accuracy of facet items are not harmed too much. Our approach consists of two methods which are facet generation and facet expansion.

- In facet generation, we directly use properties of entities corresponding to a query as its facet candidates. In facet expansion, we expand initial facets mined by traditional algorithms such as QDMiner to find more similar items contained in a knowledge base such as Freebase1.

The facets constructed by the two
methods are further merged and ranked to generate final query facets.

**ADVANTAGES OF PROPOSED SYSTEM:**

- Experimental results show that our proposed method QDMKB significantly outperforms all state-of-the-art methods including QDMiner, QF-I, and QF-J.
- It yields significantly higher recall of facet items.

**SYSTEM ARCHITECTURE:**

![Diagram of system architecture]

**CONCLUSION**

In this article, we discuss how to use a high-quality structured data base of knowledge to generate search queries. However, previous studies have shown that many users are not satisfied with this traditional search result

**REFERENCE:**

17. Zhengbao Jiang, Zhicheng Dou, Member, IEEE, and Ji-Rong Wen, Senior Member, IEEE, “Generating Query Facets using Knowledge Bases”, IEEE Transactions on Knowledge and Data Engineering 2017.s