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 initialquery facets, and then use Freebase to expand these facets to getsimilar items. We propose two different ways to expand afacet. First, we try to assign each facet to a suitable propertyof the entities corresponding to the query, and add the targetentities of the property to enrich the facet. We denote thismethod as property based facet expansion. For example, for the query ―Michael Jackson,‖ an initial facet mined byQDMiner is comprised of his compositions ―you rock myworld,‖ ―butterflies,‖ ―man in the mirror,‖ ―thriller,‖ ―cry,‖etc. We find that entity Michael Jackson has a property NominatedWorks which covers most items in this facet, hencewe could use other target entities of this property to enrichthis initial facet. When no relevant entities or propertiesare retrieved, we use the second way –type based facetexpansion. We find a common type that covers most facetitems. For example, ―Kristen Stewart,‖―Robert Pattinson,‖―Taylor Lautner,‖ etc. is a facet of query ―Eclipse.‖ The besttype covering most facet items in Freebase is Celebrities.

(3) **The facet candidates** constructed by facet generationand expansion are further merged, because there might beduplicate items within these candidates. We then re-weightthe final facets by checking the occurrence of the facet itemswithin 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 leveragethe advantages of both knowledge bases and search resultsto generate high-quality query facets, hence QDMKB hashigh potential to outperform the state-of-the-art algorithmswhich solely use search results for facet mining.

2. Existing System:

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

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

<table>), text patterns, andrepeat 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 aknowledge base as a complementary data source to improvethe quality of query facets. Knowledge bases contain highqualitystructured information such as entities and theirproperties and are especially useful when the query isrelated to an entity. We propose using both knowledge basesand search results to mine query facets in this paper. Thereason why we don't abandon search results is that searchresults reflect user intent and provide abundant context forfacet generation and expansion. Our target is to improvethe recall of facet and facet items by utilizing entities andtheir properties contained in knowledge bases, and at thesame time, make sure that the accuracy of facet itemsare not harmed too much. Our approach consists of two methods which are **facet generation** and **facet expansion**.

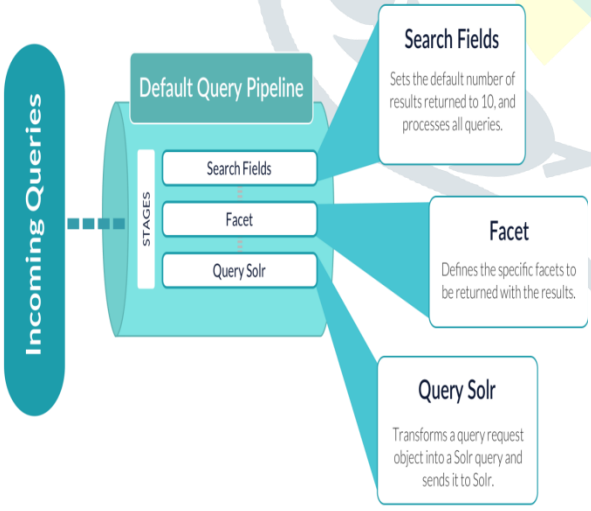
- In facet generation, we directly use properties of entitiescorresponding to a query as its facet candidates. In facetexpansion, we expand initial facets mined by traditionalalgorithms such as QDMiner to find more similar itemscontained in a knowledge base such as Freebase1. The facetsconstructed by the two

methods are further merged andranked 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:



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:

1. H. Shum, "Bing dialog model: Intent knowledge and user interaction", 2010, [online] Available: http://research.microsoft.com/en-us/um/redmond/events/fs2010/presentations/Shum_Bing_Dialog_Model_RFS_71210.pdf.

2. J. Niccolai, "Yahoo vows death to the '10 blue links'", 2009, [online] Available: http://www.pcworld.com/businesscenter/article/165214/yahoo_vows_death_to_the_10_blue_links.html.

3. R. Browne, "No longer just '10 blue links'", 2007, [online] Available: http://blogoscoped.com/archive/2007–06-28-n28.html.

4. Z. Dou, S. Hu, Y. Luo, R. Song, J.-R. Wen, "Finding dimensions for queries", *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 1311-1320, 2011.

5. W. Kong, J. Allan, "Extending faceted search to the general Web", *Proc. 23rd ACM Int. Conf. Inf. Knowl. Manage.*, pp. 839-848, 2014.

6. W. Kong, J. Allan, "Extracting query facets from search results", *Proc. 36th Int.*

ACM SIGIR Conf. Res. Develop. Inf. Retrieval, pp. 93-102, 2013.

7. Z. Dou, Z. Jiang, S. Hu, J. Wen, R. Song, "Automatically mining facets for queries from their search results", *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 2, pp. 385-397, Feb. 2016.

8. O. Ben-Yitzhak et al., "Beyond basic faceted search", *Proc. Int. Conf. Web Search Data Mining*, pp. 33-44, 2008.

9. M. Diao, S. Mukherjea, N. Rajput, K. Srivastava, "Faceted search and browsing of audio content on spoken Web", *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 1029-1038, 2010.

10. D. Dash, J. Rao, N. Megiddo, A. Ailamaki, G. Lohman, "Dynamic faceted search for discovery-driven analysis", *Proc. 17th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 3-12, 2008.

11. C. Li, N. Yan, S. B. Roy, L. Lisham, G. Das, "Facetedpedia: Dynamic generation of query-dependent faceted interfaces for wikipedia", *Proc. 19th Int. Conf. World Wide Web*, pp. 651-660, 2010.

12. W. Dakka, P. G. Ipeirotis, "Automatic extraction of useful facet hierarchies from text databases", *Proc. IEEE 24th Int. Conf. Data*, pp. 466-475, 2008.

13. W. Dakka, R. Dayal, P. G. Ipeirotis, "Automatic discovery of useful facet terms", *Proc. ACM SIGIR Workshop Faceted Search*, pp. 18-22, 2006.

14. W. Dakka, P. G. Ipeirotis, K. R. Wood, "Automatic construction of multifaceted browsing interfaces", *Proc. 14th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 768-775, 2005.

15. K. Latha, K. R. Veni, R. Rajaram, "AFGF: An automatic facet generation framework for document retrieval", *Proc. Int. Conf. Advances Comput. Eng.*, pp. 110-114, 2010.

16. E. Stoica, M. A. Hearst, "Automating creation of hierarchical faceted metadata structures", *Proc. Human Language Technol. Annu.Conf. North Amer. Chapter Assoc. Comput.Linguistics*, pp. 244-251, 2007.

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**

18.S. Basu Roy, H. Wang, G. Das, U. Nambiar, M. Mohania, "Minimum-effort driven dynamic faceted search in structured databases", *Proc. 17th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 13-22, 2008.