GENERATING QUERY FACETS USING KNOWLEDGE BASES

R.Bharathi,
Assistant Professor,
PG & Research Department of Computer Science,
Sengunthar Arts and Science College,
Tiruchengode,Tamilnadu, India.

L.Senthilkumar,
Research Scholar,
PG & Research Department of Computer Science,
Sengunthar Arts and ScienceCollege,
Tiruchengode,Tamilnadu, India.

Abstract: A query facet is a significant list of information nuggets that explains an underlying aspect of a query. Existing algorithms mine facets of a query by extracting frequent lists contained in top search results. The coverage of facets and facet items mined by this kind of methods might be limited, because only a small number of search results are used. In order to solve this problem, we propose mining query facets by using knowledge bases which contain high-quality structured data. Specifically, we first generate facets based on the properties of the entities which are contained in Freebase and correspond to the query. Second, we mine initial query facets from search results, then expanding them by finding similar entities from Freebase. Experimental results show that our proposed method can significantly improve the coverage of facet items over the state-of-the-art algorithms.

Keywords: Data mining, Query Facet Mining, QD Miner algorithm, clusters, Query Search

1. INTRODUCTION

Data mining is the process of extracting out valid and unknown information from large databases and use it to make difficult decisions in business (Gregory, 2000).Data mining or data analysis with complex and large datasets brings the wealth of research and knowledge in machine learning and statistics for the task of discovering new sets of knowledge in large databases. Over the past three decades, large amounts of difficult data's of business are stored electronically and this volume will continue to increase in future. In order to manage huge volumes of data, the techniques of data mining are also becoming sophisticated and advanced, day by day.

Smyth, Mannila and Hand (2001) have defined that "progress in digital acquisition storage technology has resulted in the development of vast database. This has happened in all areas of human attempt from the mundane to the exotic. Little wonder then that attention has development in the possibility of tapping these data, of demanding from them information that might be of value to the proprietor of the database. The regulation concerned with this assignment has become recognized as data mining. Defining a scientific discipline is always a contentious task; researchers often disagree about the exact range and limits of their field of study. Bearing this in and tolerant that others might oppose about the details, they shall accept as their functioning definition of data mining". Data mining is the analysis of observational data sets to get unsuspected relationships and to sum up the data in novel ways that are both understandable and useful to the data proprietor.

The relationships and summaries derived throughout a data mining exercise are habitually referred to as patterns or models. Illustrations include linear equations, graphs, tree structures, clusters, rules and recurrent patterns in time series. The description over refers to observational data as opposed to investigational data. Data mining classically deals with data that have already been composed for various reasons other than the data mining analysis. This means that the objectives of the data mining implement play on role in the data collection plan. This is single way in which data mining is at variance from much of statistics in which data are frequently composed by using well-organized strategies to reply particular questions (Tan, Kumar and Steinbach, 2006).

II. RELATED WORK

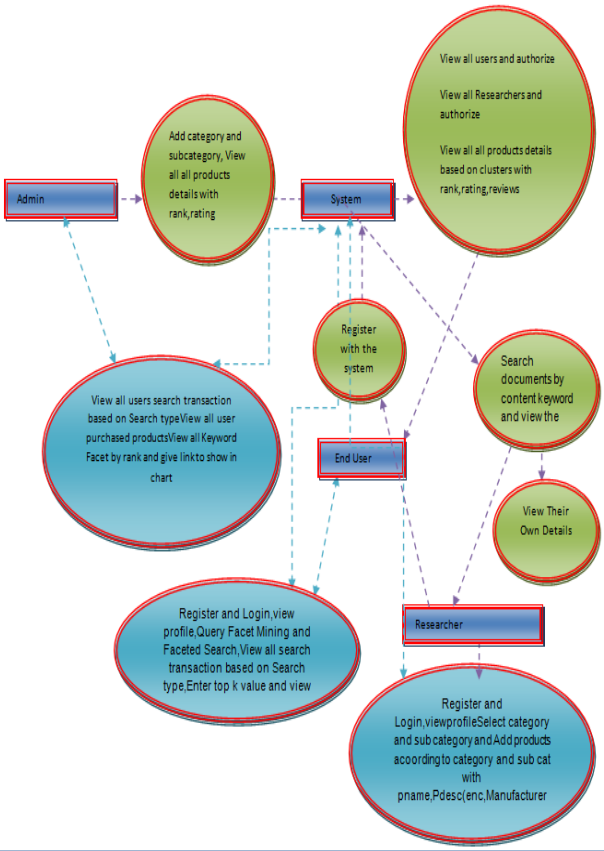
Query Facet Mining and Faceted Search : `Query facets summarize a query in different aspects. They may help users quickly understand important aspects of the query and help them explore information. Dou et al. first introduced this problem and proposed QDMiner algorithm

Set Expansion : set expansion, which focuses on expanding a partial set of "seed" objects into a more complete set

Query Recommendation : Query-based Summarization Text summarization technique is a fundamental research area. Different approaches of query-based summarization include graph-based, language model-based and machine learning. Although both summarization and query facets aim to conclude valuable information from retrieved documents, they vary in the way of presentation. Summarization uses a flat list of sentences, while query facet mining digs deeper to generate multiple lists of semantically related terms. Organizing information as lists is a moreintuitive way to guide user's browsing.

IJCRCST © 2019 | All Rights Reserved

III. ARCHITECTURE DIAGRAM



IV. PROBLEM STATEMENT

A. EXISTING SYSTEM

Query facets summarize a query in different aspects. They may help users quickly understand important aspects of the query and help them explore information. Dou et al. first introduced this problem and proposed QD Miner algorithm.
QD Miner first extracts frequent lists in top search results using predefined patterns, and then weights each list and group them into final facets. Similar to QD Miner, Kong and Allan developed supervised approaches, namely QF-I and QFJ, to mine query facets.

Facet item candidates are extracted from frequent lists which are obtained in a similar way as QD Miner. Then two Bayesian models are learned to estimate how likely a candidate is a facet item and how likely two candidates belong to the same facet. All the existing works are based on top search results; hence the quality of final facets might be limited. If some words or phrases don't appear in a list within top search results, they have no opportunity to be facet items.

A. DISADVANTAGES

- There is no efficient techniques to generate Query facets like Query Based Summarization
- There is no option to Provide Protection on User Query facets data .

B. PROPOSED SYSTEM

In the proposed system, the system proposes 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. The system also proposes 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.

The system's 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.

C. ADVANTAGES

- Automatic Query Recommendation.
- The system is more efficient due to Query Based Summarization.

V. MODULES

A. Admin module

In this module, admin server has to login with valid username and password. After login successful he can do some operations such as View all users and authorize, View all Researchers and authorize and Add category and subcategory, View all all products details based on clusters with rank,rating,reviews and View all products details with rank,rating,reviews, View all users search transaction based on Search type and View all user purchased products, View all Keyword Facet by rank and give link to show in chart, View product's rank in Chart

B. User module

In this module, User should register before searching the Website contents. After registration successful the user can login by using valid user name and password. After Login successful the user will do some operations Manage Accounts and Query Facet Mining and Faceted Search

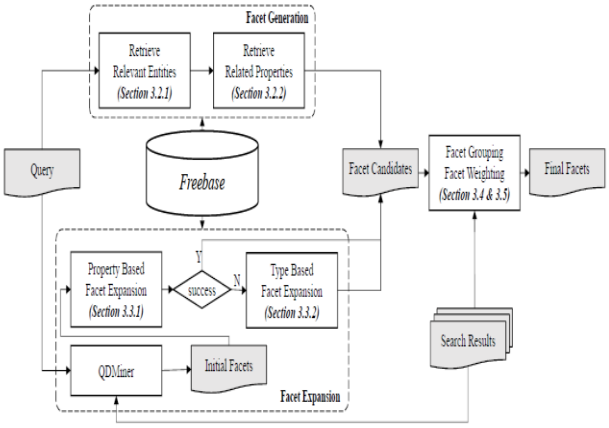
- **Single Facet Query Search :** enter your keyword or keyword +cat, or keyword +cat +Sub Cat, each based on pdesc Cat +Sub cat and display products Facet details and view details(inc rank),review, Rate,, view similar products facets, purchase
- **Multiple Facet Query Search :** Select Category and Sub category and display products Facet details and view facet details (inc rank),review, Rate, view similar products facets, purchase
- **Question Type Query Search :** Type some question and match any one word with cat or sub cat or pdesc and display products Facet details and view details(incrank),review ,Rate, view similar products facets,purchase View all search transaction based on Search type and Enter top k value and view searched Keyword facets.

IJCRCST © 2019 | All Rights Reserved

C. Researcher module

In this module, Researcher should register before searching the Website contents. After registration successful the user can login by using valid user name and password. After Login successful the researcher will do some operations Select category and sub category and Add productsaccording to category and sub cat with pname,Pdesc(enc),Manufacturer name,Price,Nodel no,Color,Add image, View all your products details with rank,rating,reviews,View all purchased your products with total bill and View all Keyword Facet by ran

VI. OVERALL SOLUTION



In this paper, we use Freebase for mining query facets. Freebase is a large knowledge base consisting of data harvested from sources such as Wikipedia, NNDB, FMD, and Music Brainz , as well as individually contributed data from users. Freebase provides a JSON-based HTTP API2 to programmers, and we can use Meta web Query Language (MQL) to customize a query to retrieve data. In this paper, we use Search API and Topic API to retrieve entities, properties, and types. All entities, properties and types in Freebase are presented in italics in this paper. The overall solution, namely QDMKB, is shown in Figure 1. We use both facet generation and facet expansion to construct facet candidates. These candidates are further merged and ranked to generate final facets

VII. CONCLUSION

Existing query facet mining algorithms, including QD Miner, QF-I, and QF-J mainly rely on the top search results from the search engines. The coverage of facets mined using this kind of methods might be limited, because usually only a small number of results are used. We propose leveraging knowledge bases as complementary data sources. We use two methods, namely facet generation and facet expansion, to mine query facets. Facet generation directly uses properties in Freebase as candidates, while facet expansion intends to expand initial facets mined by QD Miner in property based and type-based manners. Experimental results show that our approach is effective, especially for improving the recall of facet items.

VIII. REFERENCES

[1]. H. Shum, "Bing dialog model: Intent, knowledge, and user interaction," http://research.microsoft.com/en-us/um/ red mond/events/fs2010/presentations/Shum_Bing_Dia log_ Model_RFS_71210.pdf.

[2]. J. Niccolai, "Yahoo vows death to the '10 blue links'," http://www.pcworld.com/businesscenter/article/165 214/ yahoo_vows_death_to_the_10_blue_links.html.

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

[4]. Z.Dou,S.Hu,Y.Luo,R.Song,andJ.-R.Wen,"Findingdimensions for queries," in Proceedings of CIKM'11, 2011, pp. 1311–1320.

[5]. W. Kong and J. Allan, "Extending faceted search to the general web," in Proceedings of CIKM '14, 2014, pp. 839–848.

[6]. "Extracting query facets from search results," in Proceedings of SIGIR '13, 2013, pp. 93–102.

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

[8]. O. Ben-Yitzhak, N. Golbandi, N. Har'El, R. Lempel, A. Neumann, S. Ofek-Koifman, D. Sheinwald, E. Shekita, B. Sznajder, and S. Yogev, "Beyond basic faceted search," in Proceedings of WSDM '08, 2008, pp. 33–44.

[9]. M. Diao, S. Mukherjea, N. Rajput, and K. Srivastava, "Faceted search and browsing of audio content on spoken web," in Proceedings of CIKM '10, 2010, pp. 1029–1038.

[10]. D. Dash, J. Rao, N. Megiddo, A. Ailamaki, and G. Lohman, "Dynamic faceted search for discovery-driven analysis," in CIKM '08, 2008, pp. 3–12.

[11]. C. Li, N. Yan, S. B. Roy, L. Lisham, and G. Das, "Facetedpedia: dynamic generation of query-dependent faceted interfaces for wikipedia," in Proceedings of WWW '10, 2010, pp. 651–660.

[12]. W. Dakka and P. G. Ipeirotis, "Automatic extraction of useful facet hierarchies from text databases," in Proceedings of ICDE '08, 2008, pp. 466–475.

[13]. W. Dakka, R. Dayal, and P. G. Ipeirotis, "Automatic discovery of useful facet terms," in SIGIR Faceted Search Workshop, 2006, pp. 18–22.

[14]. W.Dakka,P.G.Ipeirotis,andK.R.Wood,"Automaticc onstruction of multifaceted browsing interfaces," in Proceedings of CIKM '05, 2005, pp. 768–775.

[15]. K. Latha, K. R. Veni, and R. Rajaram, "Afgf: An automatic facet generation framework for document retrieval," in Proceedings of ACE '10, 2010, pp. 110–114.