Ranking and Mapping of Hotels & Restaurants using Approximation Algorithm

Abhijeet Vijay Javlekar[1] Pratik Hemant Dabholkar[2] Madhav Manjunath Kantak[3] Jayesh Vijay Dhawase[4]
[1,2,3,4]Student
[1,2,3,4]Department of Information Technology
[1,2,3,4]Atharva College Of Engineering Mumbai, Maharashtra, India

Abstract— Keyword queries on databases provide easy access to data, but often suffer from low ranking quality, i.e., low precision and/or recall, as shown in recent benchmarks. It would be useful to identify queries that are likely to have low ranking quality to improve the user satisfaction. For instance, the system may suggest to the user alternative queries for such hard queries. In this paper, we analyze the characteristics of hard queries and propose a novel framework to measure the degree of difficulty for a keyword query over a database, considering both the structure and the content of the database and the query results. We evaluate our query difficulty prediction model against one effectiveness benchmarks for popular keyword search ranking methods. Our empirical results show that our model predicts the hard queries with high accuracy. Further, we present a suite of optimizations to minimize the incurred time overhead. We also provide mapping methodology for hotels and restaurants which will be used for getting directions with the help of Google Maps.

Key words: Query performance, query effectiveness, keyword query, robustness, databases,Google Maps

I. INTRODUCTION

Keyword query interfaces (KQIs) for databases have attracted much attention in the last decade due to their flexibility and ease of use in searching and exploring the data. Since any entity in a data set that contains the query keywords is a potential answer, keyword queries typically have many possible answers. KQIs must identify the information needs behind keyword queries and rank the answers so that the desired answers appear at the top of the list. Unless otherwise noted, we refer to keyword query as query in the remainder of this paper. Databases contain entities, and entities contain attributes that take attribute values. Some of the difficulties of answering a query are as follows: First, unlike queries in languages like SQL, users do not normally specify the desired schema element(s) for each query term. For instance, query Q1: Burger King does not specify if the user is interested in food whose title is Burger King or Foods distributed by the Burger King Company. Thus, a KQI must find the desired attributes associated with each term in the query. Second, the schema of the output is not specified, i.e., users do not give enough information to single out exactly their desired entities. For example, Q1 may return foods or company. Google maps for giving directions to users so that they can go to their desired hotel or restaurants. Prediction of Hotels and Restaurants can also be done by predicting the mood of the user. User will import the state of mood he is in and according to the input given the system will give results.

A. Aim And Objectives

Aim: To predict difficult keyword queries efficiently over database and to evaluate our query prediction model for popular keyword search ranking method.
Objectives:

1) To predict hard queries.
2) High Accuracy.
3) Query Effectiveness.
4) Robustness.
5) User Friendly Interface.

B. Problem Statement

There have been collaborative efforts to provide standard benchmarks and evaluation platforms for key-word search methods over databases. The results indicate that even with structured data, finding the desired answers to keyword queries is still a hard task more interestingly, looking closer to the ranking quality of the best per-forming methods Suffer from low ranking quality and perform very poorly on a subset of queries. The proposed system uses approximation algorithm which provides high ranking quality and helps to predict hard keyword queries and gives best results.

C. Scope

The scope of the proposed system is that it is an online database that provides information related to Restaurants (which includes foods, cafes, drinks, deserts, buffets, rank, quality, cost) and Hotels (which include 2 star, 3 star, 4 star, 5 star and 7star hotels).

The proposed systems also provide mapping constraint (google maps) which will give directions to reach any hotels and restaurants from any location.

1) Admin level - Here the admin has the access, read, write, manipulation, deletion permissions of all the profiles of users. The Files shared are stored in a Database located in the admin PC which can be directly controlled by the administrator.
2) User level- They have the read, write and access permissions to their own profile. Users can search for information about any hotels and restaurants as per the requirements.

II. LITRATURE SURVEY

Shiwen Cheng, Arash Termehchy, & Vagekus Hristidist al[1] is to predict the effectiveness of keyword queries over databases. It shows the current prediction method for queries over unstructured data sources cannot be effectively used to solve the problem. This paper set forth the principled framework and novel algorithms to measure the degree of the difficulty of a query of a database using the ranking robustness principle .

Deepti S Deshmukh, , Simran Khiani.al[2] this paper has the frame work which shows Post retrieval methods which predict the difficulty of a query with computing its results. The computation the query prediction based on approximation algorithm gives the outcomes of corrupted database in re-ranked form.

Basim Ali Razooqi al[3] : This paper focuses on main problem of retrieving appropriate top results for a keyword query and predicting the difficulty level of the query. In this paper, we analyze the properties of complex queries and measure the degreeof complexity of a keyword query over a database. We measured the degree of the complexity of a query over a database, using the ranking robustness principle. The framework efficiently predicts the effectiveness of a keyword query.

III. PROPOSED SYSTEM

The conceptual model of the system involves creation of a database driven web based application that can be easily deployed on all types of platforms and form factors. As a reason responsive web designing is adopted.
The web application will be divided into the following modules

1) Client side interface
2) Server side server to handle the client request
3) Container to store the business logic
4) Database to store information's of hotels and restaurants

In a nutshell the basic aim of the project is to create a keyword search engine application that will accept multiple keywords from the user. The proposed system will help to generate results and then the admin process the results and gives it to the user. It should promptly detect and recognise keywords from the database system, to create an excel sheet that will mark each keyword's presence or absence. This excel sheet is accessible by the respective admin and should be content editable in case some keywords are missed out or incorrectly identified

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:
The block diagram for this system is given as follows and is self-explanatory.

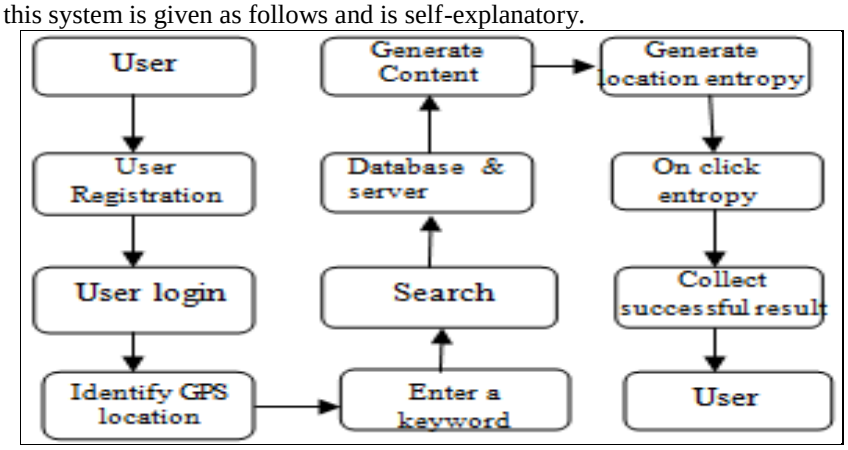

Fig. 1: Proposed System

Working of the system:

1) Admin Login
2) New Hotels and Restaurants Addition
3) Edit / Delete Old Hotels and Restaurants Database
4) Create Training Dataset
5) View & Authenticate Users
6) View Users Comments
7) User Registration
8) User Login

9) Search Hotels and Restaurants& View Result
10) Comment on Post

The technology and algorithms used for achieving this goal are explained next section.

A. Approximation Algorithm

In this section, we propose approximation algorithms to improve the efficiency of SR Algorithm. Our methods are independent of the underlying ranking algorithm. Query-specific Attribute Values Only Approximation (QAO-Approx): QAO-Approx corrupts only the attribute values that match at least one query term. This approximation algorithm leverages the following observations:

- Observation1: The noise in the attribute values that contain query terms dominates the corruption effect.
- Observation 2: The number of attribute values that contain at least one query term is much smaller than the numbers of all attribute values in each entity.
- Observation 3: Given that only the Top-K result entities are corrupted, the global DB statistics do not change much.

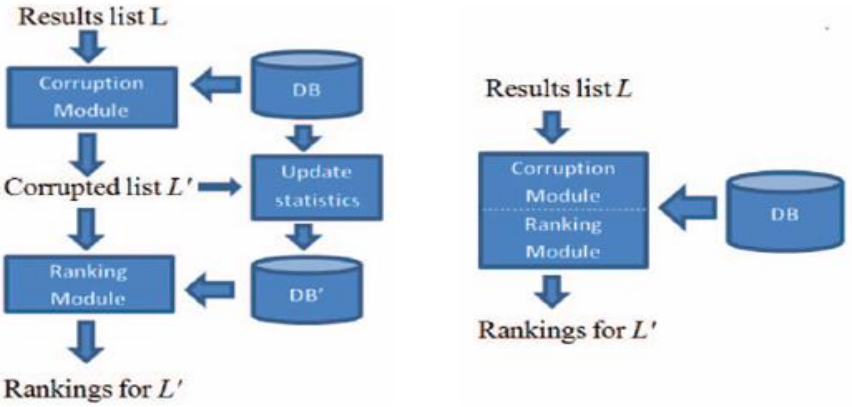

Fig. 3: Approximation algorithm

IV. CONCLUSION

We introduced the novel problem of predicting the effectiveness of keyword queries over DBs. We showed that the current prediction methods for queries over unstructured data sources cannot be effectively used to solve this problem. We set forth a principled framework and proposed novel algorithms to measure the degree of the difficulty of a query over a DB, using the ranking robustness principle. Based on our framework, we propose novel algorithms that efficiently predict the effectiveness of a keyword query. Our extensive experiments show that the algorithms predict the difficulty of a query with relatively low errors and negligibletime overheads.

ACKNOWLEDGMENT

It gives us great pleasure in presenting this project report titled “Ranking and Mapping of Hotels & Restaurants Using Approximation Algorithms” and we wish to express our immense gratitude to the people who provided invaluable knowledge and support in the completion of this project. Their guidance and motivation has helped in making this project a great success. We express our gratitude to our project guide Prof. Sumita Chandak, who provided us with all the guidance and encouragement throughout the project development. We would also like to express our sincere gratitude to the respective Project coordinators. We are eager and glad to express our gratitude to the Head of the Information Technology Dept. Prof. Neelima Pathak, for her approval of this project. We are also thankful to her for providing us the needed assistance, detailed suggestions and also encouragement to do the project. We would like to deeply express our sincere gratitude to our respected principal Prof. Dr. Shrikant Kallurkar and the management of Atharva College of Engineering for providing such an ideal atmosphere to build up this project with well-equipped library with all the utmost necessary reference materials and up to date IT Laboratories. We are extremely thankful to all staff and the management of the college for providing us all the facilities and resources required.

REFERENCES

[1] G V. Hristidis, L. Gravano, and Y. Papakonstantinou, “Efficient Irstylekeyword search over relational databases,” in Proc. 29th VLDB Conf.,Berlin, Germany, 2003, pp. 850–861
[2] Y. Luo, X. Lin, W. Wang, and X. Zhou, “SPARK: Top-k keyword query in relational databases,” in Proc. 2007 ACM SIGMOD,Beijing, China, pp. 115–126.
[3] V. Ganti, Y. He, and D. Xin, “Keyword++: A framework to improvekeyword search over entity databases,” in Proc. VLDB Endowment, Singapore, Sept. 2010, vol. 3, no. 1–2, pp. 711–722.
[4] J. Kim, X. Xue, and B. Croft, “A probabilistic retrieval model for semistructured data,” in Proc. ECIR, Tolouse, France, 2009, pp. 228–239.
[5] N. Sarkas, S. Paparizos, and P. Tsaparas, “Structured annotations of web queries,” in Proc. 2010 ACM SIGMOD Int. Conf. Manage.Data, Indianapolis, IN, USA, pp. 771–782.