A Restaurant Recommendation System Based on Range and Skyline Queries

Abhishek Singh, CS Department
Kent State University
asingh22@kent.edu

Asaad Y. Althoubi, CS Department
Kent State University
aalthoub@kent.edu

Pavan Poudel, CS Department
Kent State University
ppoudel1@kent.edu

Tala Talaei Khoei, CS Department
Kent State University
ttalaeik@kent.edu

I. Introduction

Finding relevant outcomes in a sea of information which refers to as big data have become an ongoing challenge for users. Over the last years, building systems and applications meet the user's requests have become a hot subject among the researchers and developers. It is desirable to offer flexible queries match the user's demands. The preference queries usually play a crucial role in providing queries meet the user's minds. Additionally, the preference queries usually help in finding the rank-ordering of the objects retrieved.

Upon the reasons mentioned above, we have developed in this paper a system offers some flexible queries match the user's demands. The preference queries which are exist in this system rely on complete data. In other words, all attributes (dimensions) for every data object are available (no noisy data).

In this paper, we build a restaurant recommendation system employs the preference queries based on complete information for restaurant recommendations. In this system, the queries are categorized into two types. First, spatial range query in which a user requests a set of restaurants in within a given range (radius). Second, Skyline query in which users request a set of restaurants are not dominated by any others in terms of certain attributes. The Skyline queries are applied on the results retrieved from the spatial range queries.

Effectively visualizing data within spaces of multi dimensionality is considered a challenging task to computer scientists. Our system, in addition to displaying the results corresponding to the user query, presents visualization graphs corresponding to the user queries. This system is able to visualize 2 and 3 dimensional graphs, and these visualizations give users the opportunity to when making adjustments to their query.

II. Definitions of Queries

The spatial range query and Skyline query are shown below in Definiton1 and Definition 2, respectively. The Skyline queries are based on the dominance relationship as stated in Definition 2.

Definition (1): Given a query object q, a distance threshold r and a dataset D , a spatial range query returns a set of objects in D that is at most r distant from the query object q.

Definition (2): Given two complete objects q_1 and q_2, we say that q_1 dominates q_2 if the following two conditions hold:

1. For every dimension i, either $q_1.[i]$ is no worse than $q_2.[i]$ or at least one of them is missing.

2. There is at least one dimension j, in which both $q_1.[j]$ and $q_2.[j]$are observed and $q_1.[j]$ is better than $q_2.[j]$.

Upon above, the Skyline query over a set S of objects retrieves a set of objects, $S' \subseteq S$, that are not dominated by any other objects in S.

III. Research Background

[1] has developed a restaurant recommendation system with preference queries on incomplete information. The system, SI^2P, is capable of recommending desirable restaurants based on preference queries that take the incomplete ratings information into consideration. SI^2P provides the server side based on an extended PostgreSQL database that integrates two types of preference queries, namely, skyline and top-k dominating queries over incomplete data. It also offers the browser-based interface for the users to interact with the system.

It was described in [2] a web-based search application used for locating restaurants in a multidimensional space via interactive space visualization. The major goal of this research is to

reduce the cognitive complexity of query selection, as well as to help the user avoid one of the common pitfalls of traditional search mechanisms: retrieving too many or too few results. The approach constructs a graphical output to summarize how further query adjustments will impact subsequent search results.

IV. The System Architecture

Our system adopts a browser model and is implemented with Java8 runtime, and deals with a real-time data stream by using Google Clients API. The user submits a query through the user interface, and the query then starts processing searching for results. After the query is processed, the results are displayed on Google Maps in the user's browser.

The system offers a user interface to the users for generating queries, viewing the retrieved results. The system, also, shows a Google Map to provide interactions by using Google Maps API. Furthermore, the system provides a mechanism for users to sort their query results based on their preferred critiques. In other words, the system is able to suggest the most dominating restaurants within the selected region in terms of the different attributes, such as distance, price-level, cost, time. The user interface is shown in Figure 1 below:

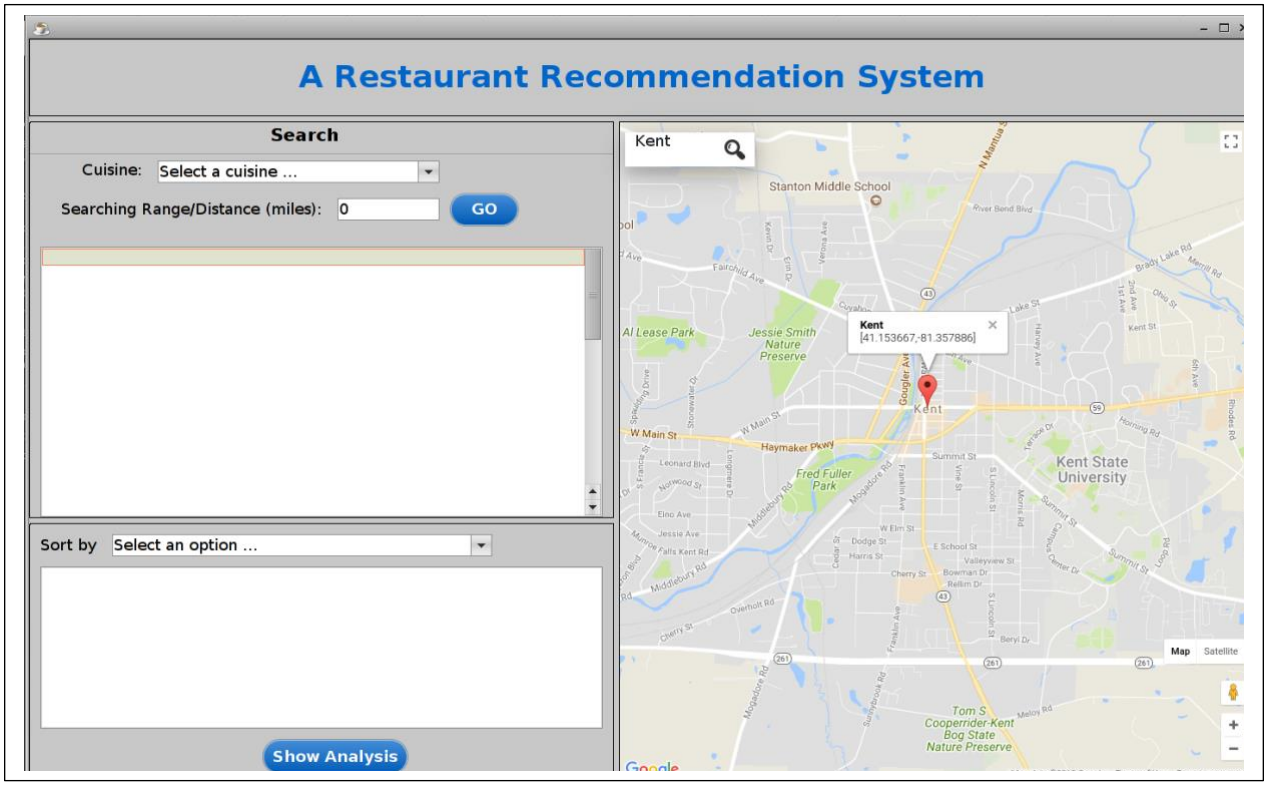

Fig 1: Shows the user interface and how users execute query and then results displayed

V. Baseline Tools

We have adopted a number of various tools to come up with a complete functional system meets user's demands. The following shows the major tools have used to construct this system and the advantage of using them.

1. JAVA 8: it is the main programming platform was used in this system due to its ability to integrate with the different operating platforms.
2. JSON: (JavaScript Object Notation) is a lightweight data-interchange format. It was used to parse and generate because of its easy formats when reading and writing.
3. Google Maps API: give ways of embedding Google Maps into web pages or retrieving data from Google Maps, and allow for either simple use or extensive customization. Google Maps is mainly used in this system to show how to visualize restaurant data on Google maps.
4. JAVA Libraries: a number of Java libraries were involved to connect between the functions and formats of our system platform and make everything works smoothly.

VI. Results

When the query results are turned to the users. The result restaurants are displayed in both the map and the result list as shown in Figure 2 below:

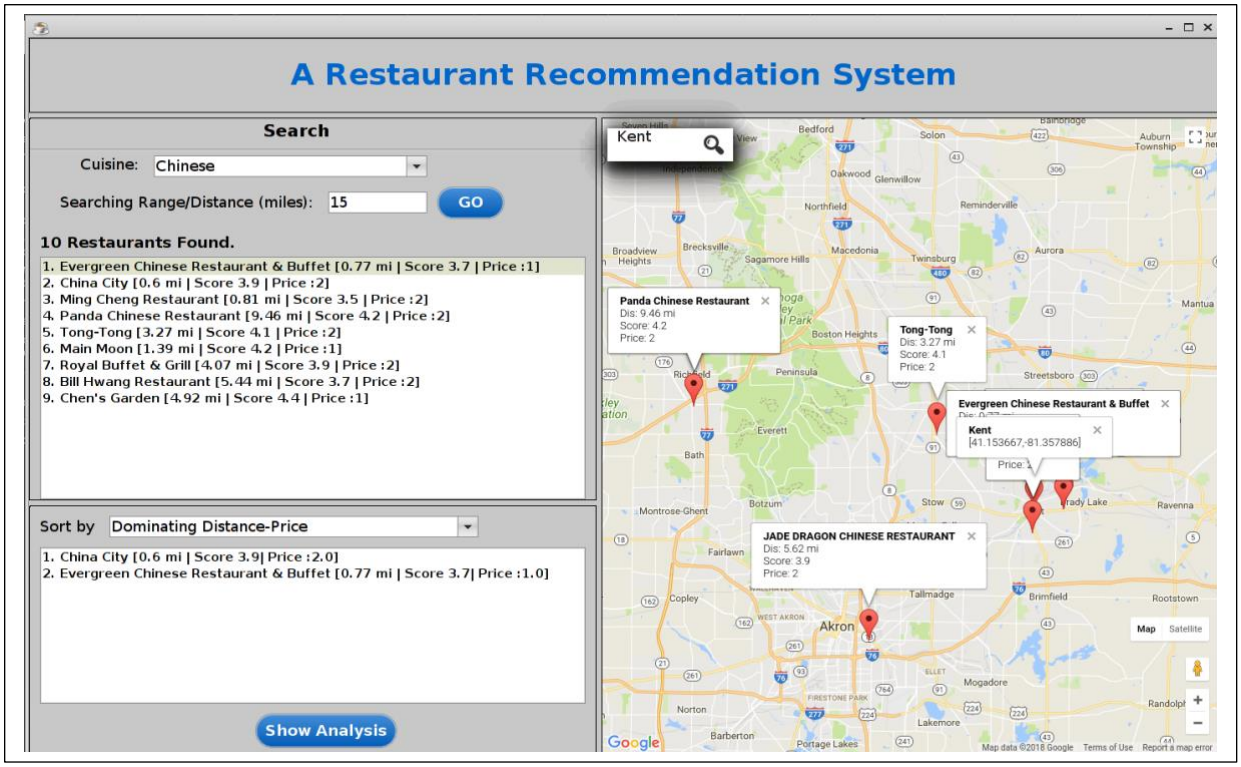

Fig 2: Illustration of submission range and Skyline query

The users are able to change the sort of the retrieved restaurants to see different recommended restaurants. Specifically, the users can select the ranking(sorting) option via a drop-down list, as shown in the bottom left hand side of Figure 2, where four options are provided including sorting by dominating distance-price-score, by distance-price, by distance-score, and by price-score.

VII. Results Visualization

When user's query is processed, a visualizer displays the distribution of the retrieved restaurants on a multidimensional graph for each ranking option selected by the user. This visualizing feature enables the user to see the dominance of results retrieved after ranking over the whole results. Our system supports visualization of 2 and 3 dimensional spaces. The advantage of visualizing the obtained results in our system is just to let users have a clear picture of what kind of restaurants they are looking and to find suitable locations. For example, the users may realize that by changing the value of the distance attribute to expand their search to restaurants further away from their locations, they may be able to find a location that is highly dominated amongst the other people in terms of the five different attributes adopted in this system, namely cuisine, distance, price, score, and time. Below, Figure 3 and 4 show 2D and 3D visualization samples of obtained results after selecting raking option by a user.

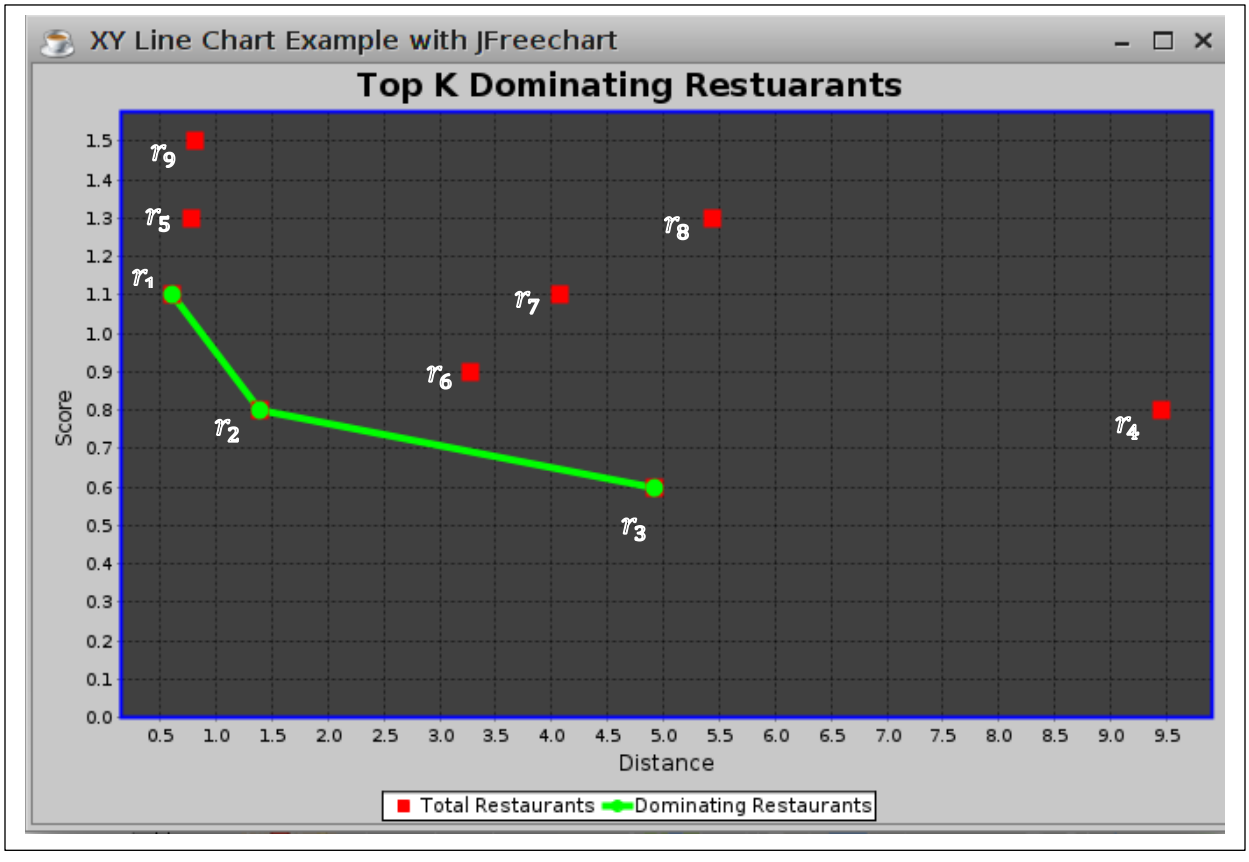

Fig 3 (a): Displays a visualization sample of 2D for obtained restaurants.

In Figure 3 (a), the objects linked to each other represent the dominated restaurants among the others. Specifically, restaurant r_2 dominate the restaurants r_4, r_6, r_7, r_8, restaurant r_1 dominate the restaurants r_9, r_5 and restaurant r_3 dominates nothing.

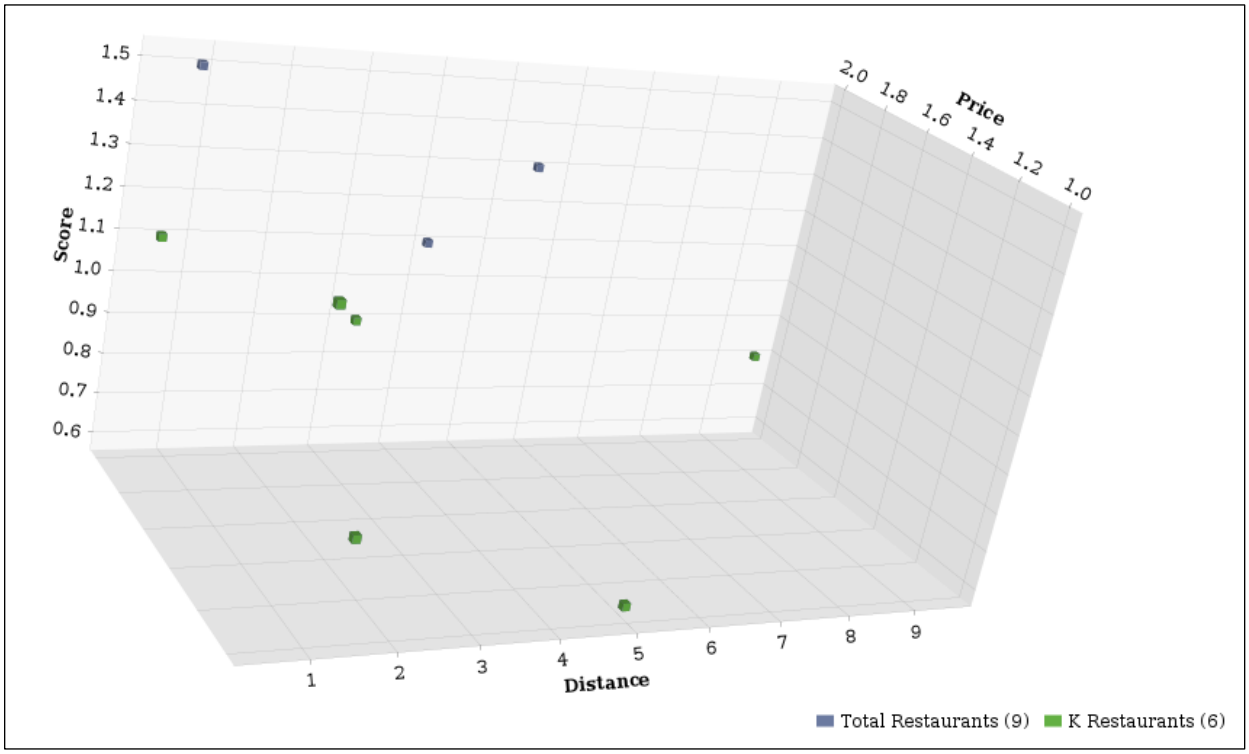

Fig 3 (b): Displays a visualization sample of 3D for obtained restaurants.

VIII. Performance

In our system, the work performed can be divided into three categories: searching, ranking (sorting by), and visualizing. The overall processing time varies from system to system, but in our system, we reduced the complexity time much, especially when using Skyline queries. In this system, complexity time of Skyline algorithm is $\mathcal{O}(K)$ where K is the number of the restaurants retrieved after the range query. Thus, the total complexity time of our system is $\mathcal{O}(K) + \mathcal{O}(nlogn)$ where $\mathcal{O}(nlogn)$ is the time spent for sorting results by distance.

IX. Conclusion

In this project, we developed a restaurants recommendation system based on preference queries. The system adopts a browser model and is based on real-time Google dataset, which supports skyline and range queries. Additionally, the system supports visualizing the recommended

/retrieved restaurants in 2 and 3 dimensional graphs to give users the opportunity to when making adjustments to their query.

X. Future Work

In addition to supporting convenient and flexible query submission and query result, the following may make our system more prominent and practical.

1. Allow the users specifying an interested region by drawing a rectangle in map to obtain restaurants in that area

2. Allow the users to write reviews on the restaurants and score them such that the information of restaurants is updated

3. The system can be implemented on a mobile platform which provides user current locations using mobile GPS

REFERENCES

[1] Xiaoye Miao, Yunjun Gao, Gang Chen, Huiyong Cui, Chong Guo, and Weida Pan. "SI^2P: A Restaurant Recommendation System Using Preference Queries over Incomplete Information".

[2] Alexei Yatskov, and Yasushi Kiyoki, "Restaurant Search with Predictive Multispace Queries".