IJCSE International Journal of Computer Science and Engineering Open Access
Research Paper Volume-2, Issue-3 E-ISSN: 2347-2693

Top-K Spatial Preference Query with Range Based Skyline Query in Mobile Environment

R. Prema Steffi [1*] and S. Sundaramoorthy [2]

[1*]Dept. of CSE, Angel College of Engineering and Technology, Tirupur, India. steffirajan@gmail.com
[2]Dept. of CSE, Angel College of Engineering and Technology, Tirupur, India. sundaramoorthy87@gmail.com

Received: 6 March 2014 Revised: 18 March 2014 Accepted: 27 March 2014 Published: 30 March 2014

Abstract— As the location based service makes the users to process their queries from anywhere and at anytime, the spatial query processing that with this provides the processing on the basics of the spatial attributes in the skyline. With thus provides the users to identify the nearest neighbour among their query given with its spatial attributes. With the existing, they have done the skyline query processing based on the range in the spatial attribute for analyzing the nearest among the other skyline data sets. They have done these using two novel algorithms as index and non index algorithm. We are going to concentrate by making with that range query to process on the basis of user's feedback as their rating for the skyline query processing. Here the processing is done by ranking on the spatial attribute in the skyline data sets. When the user makes the feedback as their rating for the result of the skyline, this will be also analyzed while considering the next query processing for the nearest neighbour in the skyline query processing. Here, we are going to consider the algorithms as nearest neighbour algorithm and branch and bound algorithm in which this makes the analysis on the basics of both nearest and the minimum bound within the skyline.

Index Term— Spatial data mining, Skyline Query Processing, Interesting Points

I. INTRODUCTION

We all know about the location based services in the mobile environment, which provides the users to make their query from anywhere and at anytime for their processing and thus supports the users very well. In this fast decaying world, this location based services helps the users highly with their instant services for the users. Conventional LBSs focus on processing proximity-based queries, which makes use of the range query and nearest neighbor (NN) query [8]. However, these queries are not sufficient for the applications that need to consider both spatial and nonspatial attributes of the objects being queried. For that, they makes typical scenario to find a nearby car park with cheap parking fee, with distance as a spatial attribute and parking fee as a nonspatial attribute. With this, the spatial attributes are also considered for the analysis of the space related data in the skyline which is given as the query from the users to make processing in the skyline. In these criteria, the processing is done with the location based along with the spatial attributes. On these bases, in the existing, they consider the range query for their processing and they consider the scenario as the car parking in the skyline which thus gives out the result on the basis of the nearest distance to park and with the minimum price as the parking fee. They have done this on the basis of the range query with the spatial attribute. With that, they analyzed the cheap and the best one within the skyline data sets in the skyline query processing.

We have extended our work with this range based skyline query for the query processing by giving the users feedback with the section for the users. In this the users will look at the query result and thus provides the query rating for those spatial attribute processing. With this we have done our work based on those preferences rating from the user along the range query will be analyzed for the users providing the interesting point in the skyline data sets. We chosen our data sets as the hospital and atm data sets for making analysis with the spatial data mining and thus to provide within the range based skyline points by their interesting points in the skyline data sets.

With thus gives the skyline results for the users in the mobile environment with it's LBS in the spatial data on analyzing the spatial attributes given by the users.

Park	Park Fee	Service Quality
A	15	4
B	14	5
C	13	5
D	18	3

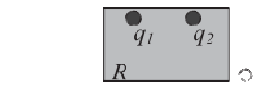

Fig 1.1 Car parking scenario for range based skyline queries.

In the above figure 1.1 it shows the scenario of car parking with thus gives the analysis on the basis of the cheap and best for parking the car in the given range. For this analysis we have taken as the hospital data sets as well as the atm data sets in the skyline. These data sets in the skyline makes the view regarding it's spatial attribute given by the user with it's range for making the skyline query processing.

Corresponding Author: *R.Prema Steffi,steffirajan@gmail.com*

II. RELATED WORKS

The Skyline Query Processing is made based on the spatial attribute that is given by the users. This makes the evaluation by making use of the spatial data within it. With the skyline, the users can identify the interesting points in the skyline and thus helps to make further process with those results. With this skyline in the spatial data we have made the literature survey.

Skyline query is one of the most widely used preference queries. It is based on the dominance relationship between tuples (Leewuen, 2010) [9]. Skyline query filters out set of all interesting point in the dataset. Here the point is said to be interesting if it is not dominated by any other point. Here the point may dominates another point if it is good or better in all dimensions and better in at least one dimension.The Skyline Query is based on the dominance relationship between tuples.

Assuming that smaller values are preferable in all dimensions, in which the dominance relationship and the skyline is defined as follows (Leewuen, 2010) [9].Example the following dataset taken from [9] clearly explains the concept skyline Queries. In which they considers the dataset which has the information about hotels, their distance to the beach and the price. A two dimensional plot of the dataset is shown in Figure 2.1; the distance and price are assigned to the X, Y axis of the plot respectively, [1].The goal is to find a hotel whose distance to beach and the price are both minimum. [1], dataset may not have one single point that satisfies both desirable properties. In which, here the user is presented with a set of interesting points that partly satisfies the imposed constraints. Here the interesting points in the dataset will be:

'a' indicates shortest distance from the beach

'k' indicates least price

'i' has neither shortest distance nor the least price

All the other points are dominated by set of points {a, i, k}, i.e., both distance and price values are greater than one or more skyline points.

In [2], Skyline query processing was first introduced into the database community by Borzonyi,[2] a number of algorithms have been proposed from then on. These algorithms can be divided into two categories[2], the first category is nonindex algorithms and the representatives are Black Nested Loop (BNL) and Divide-and-Conquer (D&C). In the main advantage of [2], makes the operation filters out a set of interesting points from a potentially large set of data points.

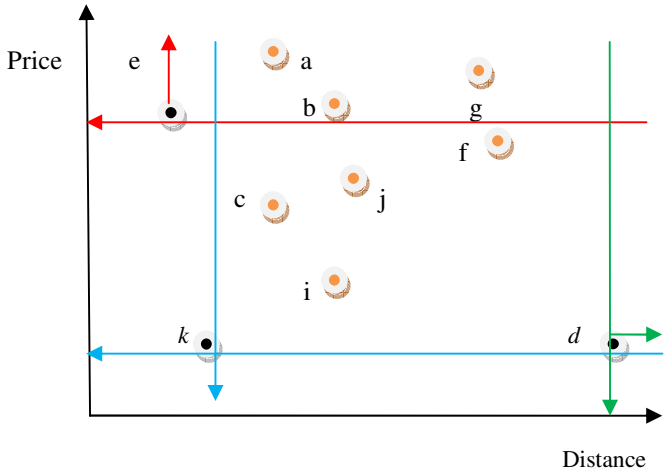

Figure 2.1 Hotel data set

Here they consider the point is interesting if it is not dominated by any other point. With this they consider an example, a hotel might be interesting for somebody traveling to Nassau [2] if no other hotel is both cheaper and closer to the beach. This [2] also have some issues that it doesn't consider the location based services with mobile environment and becomes difficult to identify the exact region in the mobile environment with skyline queries, because consider the result only database based services. In mobile environments, [3] users do not have the accurate knowledge about their locations to specify the query points because all location identification methods have errors. Even if they have such knowledge, [3] they may not want to expose these locations to the service providers for privacy reasons. [3], RNN queries address these issues by allowing users to specify ranges rather than points for NN queries.

The main advantage of [3], makes Ranges as (hyper-) rectangles and propose e□cient in-memory processing and secondary memory pruning techniques for RNN queries in both 2D and high-dimensional spaces. These techniques that are used as generalized for kRNN queries, which thus return the k nearest neighbors for every point in the range. It also has issues in [3] that are inefficient to submit many PNN queries individually to the server. None of the method combines both range based skyline queries. [4] Introduced the skyline query problem in the context of LBSs and proposed a continuous skyline query processing algorithm called CSQ for moving clients. [3], assuming a linear movement model, CSQ process the skyline query at the starting point of the query segment and tries to predict the possible changes to the answer set when the client moves. Skyline algorithms have so far dealt mainly with queries of static query points over static datasets. In [4], it has advantages as Skylines at different times are indicated by different line chains. In which, here the situation becomes more complex when all data points can move, which is frequent in real-time applications like e-games and digital war systems. It thus have issues in [4, 5], nevertheless, in which the previous studies have limitations as they simply

assume that the query location is an exact location point or a line segment. Skyline queries assume a static setting, in which the distances from the query point to the data points do not change. [6] Trusted server is needed to cloak at least K users' locations for protecting location privacy. In order to implement K-anonymity, here one of its trusted server is set up to collect user location information and perform cloaking procedures in which the exact location of the query requester is blurred as a cloaked spatial area whose boundary is defined by the locations of K−1 other users. Then, the trusted server will send the location-dependent query along with the cloaked spatial area [6-7] to location-based service providers to retrieve location-dependent data. Query location is an area instead of a single query point, location- dependent service providers should fetch those query results based on the cloaked spatial region [6]. Also it has advantages by providing the network distance spatial query solutions which can preserve user privacy by utilizing K-anonymity mechanisms. It also has feasibility and efficiency of their approach through extensive simulations. However there is no solution for solving queries on spatial networks with privacy protection.

In [8], the skyline of a set of d-dimensional points contains the points that are not dominated by any other point on all dimensions. Skyline computation has recently received considerable attention in the database community, especially for progressive (or online) algorithms that can quickly return the first skyline points without having to read the entire data file. Here currently, the most efficient algorithm is NN (nearest neighbors), which applies the divide- and-conquer framework on datasets indexed by R-trees. [8], Although NN has some desirable features, it also presents several inherent disadvantages. In this research develop BBS, [8] a progressive algorithm also based on nearest neighbor search, [8] which is IO optimal, i.e., it performs a single access only to those R-tree nodes that may contain skyline points. [8] Furthermore, it does not retrieve duplicates and its space overhead is significantly smaller than that of NN. Finally, BBS is simple to implement and can be efficiently applied to a variety of alternative skyline queries. With this [8], BBS is simple to implement and can be efficiently applied to a variety of alternative skyline queries. Here, an analytical and experimental comparison shows that BBS outperforms NN (usually by orders of magnitude) under all problem instances. Unlike dynamic nature of the spatial attribute makes skyline queries in LBSs unique and challenging.

III. EXISTING SYSTEM

It makes relax the assumption and propose a more general skyline query range-based skyline query (RSQ), which takes a spatial range as the input of user location, and when opposed to a point or a line in existing LBS skyline studies. Given a range-based skyline query [1], the service provider should return a collective set of skyline results for every possible query point of the user in R. [1], Unfortunately; this is not easy as the number of possible points in R is infinite.

Existing solutions focus on solving point- or line-based skyline queries, i.e when the query location is an exact location point or a line segment. [1] However, due to privacy concerns and limited precision of localization devices, and thus the input of a user location is often a spatial range. In this paper it makes the studies as a new problem of how to process such range-based skyline queries. To address this issue, [1] proposed a novel index-based algorithm, called I-SKY. [1], the idea is to precompute the skyline scopes for all objects by their dominance relations. And thus by indexing the skyline scopes, in this, a range-based skyline query can be efficiently processed through a root-to-leaf traversal of the index tree. In that, they have taken an example of range based skyline query with car parking scenario. To handle frequent movements of the objects being queried, and thus have proposed incremental versions of I-SKY and N-SKY, [1] which avoid recomputing the query index and results from scratch. Additionally, they have developed efficient solutions for probabilistic and continuous range-based skyline queries.

The range-based skyline problem becomes even more challenging for dynamic data sets [1] where the objects being queried can move and update their spatial attributes frequently by making non index-based algorithm. The main advantage of this existing system, is to handle the movement of the objects being queried and thus the probabilistic range-based skyline problem to reduce both the result set size and computation cost. But the issues in [1] are, the range based skyline queries are less result to find the points and it does not generate skylines queries with combinational way. To address these issues, we have made these combinatorial skyline data sets for the skyline query processing as well as the spatial preference query in the spatial attribute with thus done with the user's preference in the skyline query results.

IV. PROPOSED SYSTEM

First we generalize the concept of skyline query and introduce a novel type of query, in which the skyline query is considered as combinatorial, i.e to find the outstanding combinations from all combinations of the given tuples. In this project, we first introduce the concept of the combinatorial skyline query (CSQ) and explain the difficulty in resolving this type of query. Then we are going to focus on the user ranking i.e the feedback given by the user with the result of the skyline query processing in the spatial data and thus we proceed the further skyline query processing on these ranking given by the user in order to make the their preferences in the query processing to give the result as the interesting points in the skyline data sets.

A. Architecture of the Proposed System

In the figure 3.1, it shows the architecture diagram of the proposed system. In which, the user makes the query in the spatial data to make the process with the spatial attributes. Without using the skyline query processing it gives out lot of results for those query given. The query based result in them

may have unrelated results also within them. When it's processed with the skyline in those spatial attributes given by the users, it gives the interesting points in the skyline data sets. With that we have done the part of existing work along with our extended work by making the processing based on both index based RSQ and non index based RSQ, in which they makes use of static data sets as well as dynamic data sets for it's processing. Then we make our extended work by giving the users to make the ranking based on the results shown by the skyline query processing. When the user makes the ranking it will be made to store it in the server automatically. Then when the skyline query is given with it's spatial attributes, it will be processed by the preference given by the user and thus makes the processing with the skyline data sets.

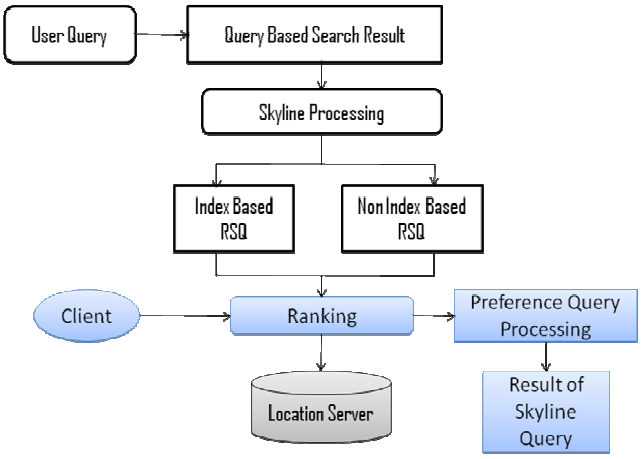

Fig 3.1 Architecture diagram of proposed system

With those preferences in the skyline data sets, the analysis is done and thus gives out the interesting points in the skyline based on its spatial attributes. Again whenever the user makes the ranking with the result and thus makes the server to be updated automatically to make the further process in the skyline query processing when the query is given.

B. Algorithm used in the proposed system

In the proposed system, we have done our work based on the two algorithms mainly. One is based on the nearest neighbour algorithm and the other is based on the branch and bound algorithm. In the nearest neighbour algorithm, we make the analysis for considering the skyline query processing with the skyline data sets as the nearest one in the skyline. This is done for viewing the interesting point in the skyline with its spatial attributes given. Likewise, by using the branch and bound algorithm, we make the analysis for processing the skyline query to have the criteria either as in the minimum bounded rectangle or as the maximum bounded rectangle in the skyline. i.e. In query processing it may be choose as maximum in any of its spatial attributes and minimum in other spatial attributes. On analyzing, this criterion in the skyline data sets, it will give out the interesting points in the skyline.

V. EXPERIMENTAL RESULT

In this we have done our experimental setup using java along with MySQL. To make the analysis for the skyline queries and to find the interesting points in the skyline data sets, we have used the WAMP Server. Here we have taken the data sets as the hotel and restaurant, for the scenario to find the interesting points in the skyline. In this both hotel and restaurant we have taken 30 objects within the data sets. With this we have taken the operating system as windows 7 and with apache along with php, we have generated the skyline points along with their interesting points. In this data set we have presented the data for the making the analysis in the server. In this, we considered the skyline as minimum distance from hotel to the restaurant and the maximum quality rating in the restaurant. By these criteria, it makes the analysis on the skyline query processing and thus proceeds in the skyline data sets.

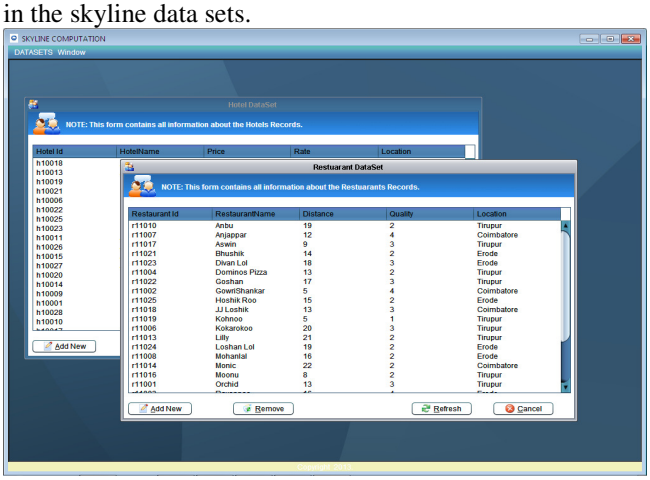

Fig 5.1 Shows the data sets of hotel and restaurants

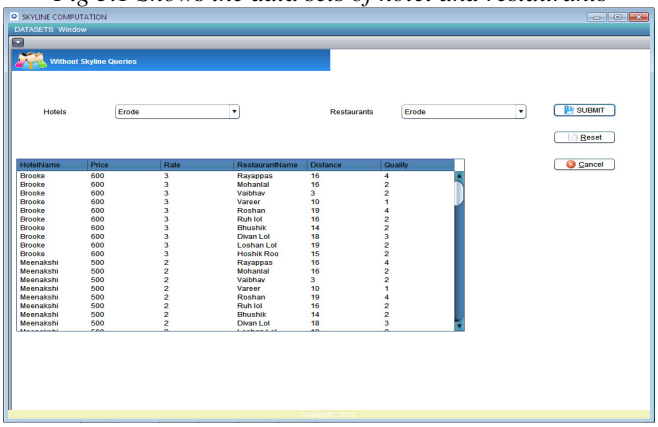

Fig 5.2 Shows the query processing without skyline

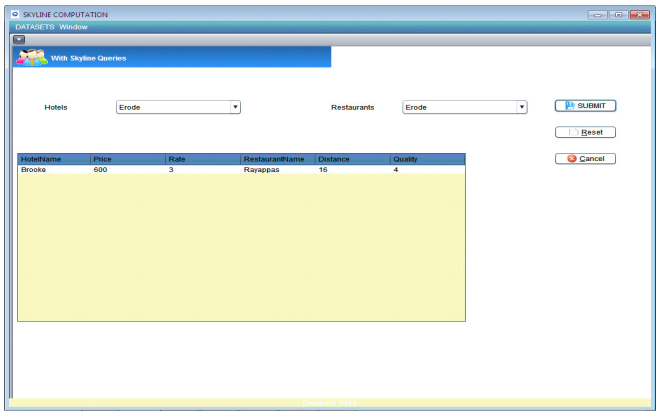

Fig 5.3 Shows the query processing with skyline

VI. CONCLUSION

Thus we have done our work in the skyline query processing by using the skyline data sets in the WAMP server. We made the analysis on the skyline to identify the interesting points in the skyline data sets and thus viewed the difference with the non skyline query processing and the skyline query processing. Up to this we have done the query processing analysis with the skyline and thus we have doing the further process as it's continue to these data sets for making the ranking analysis on the preference made by the user. We also extended our future work to make its analysis on the basis of automatic updates on the skyline result in the continuous moment in the location.

REFERENCES

[1] S.Borzonyi, D.Kossmann, and K. Stocker, "Range based skyline queries in mobile environment" *Proc. Int'l Conf. Data Eng.,* pp. **421-430, 2013.**

[2] S.Borzonyi, D.Kossmann, and K. Stocker, "The Skyline Operator," *Proc. Int'l Conf. Data Eng.,* pp**. 421-430, 2001.**

[3] H. Hu and D.L. Lee, "Range Nearest Neighbor Query," IEEE Trans. Knowledge and Data Eng., vol. **18**, no. **1**, pp. **78-91**, Jan. **2006**.

[4] Z. Huang, H. Lu, B.C. Ooi, and K.H. Tong, "Continuous Skyline Queries for Moving Objects," IEEE Trans. Knowledge and Data Eng.,vol. **18**, no**. 12,** pp. **1645-1658**, Dec. **2006**.

[5] Ilaria Bartolini, Paolo Ciaccia, and Marco Patella, "The Skyline of a Probabilistic Relation" IEEE Trans. on Knowledge and Data Engineering, vol**. 25**, no.**7,** July **2013**.

[6] W.S. Ku, R. Zimmermann, W.-C. Peng, and S. Shroff, "Privacy Protected Query Processing on Spatial Networks," Proc. ICDE Workshop Privacy Data Management, pp. **215-220, 2007**.

[7] Parke Godfrey, Ryan Shipley, Jarek Gryz"Maximal Vector Computation in Large Data Sets"

[8] D.Papadias, Y.Tao, G.Fu, B. Seeger, "An Optimal and Progressive Algorithm for Skyline Queries" Proc. ACM SIGMOD Int'l conf. management of data,**2003**.

[9] D. Papadias, Y. Tao, G. Fu, and B. Seeger, "Progressive Skyline Computation in Database Systems," ACM Trans. Database Systems, vol. **30**, no. **1**, pp. **41-82, 2005**.

[10] M. Sharifzadehand C. Shahabi, "The Spatial Skyline Queries,"Proc. 32nd Int'l Conf. Very Large Data Bases (VLDB), **2006**.

[11] J. Xu, X. Tang, H. Hu, and J. Du, "Privacy-Conscious Location- Based Queries in Mobile Environments," *IEEE Trans. Parallel and Distributed Systems,* vol. **21**, no. **3**, pp. **313-326**, Mar. **2010**.

AUTHORS PROFILE

R.Prema Steffi pursuing M.E Computer Science and Engineering in Angel College of Engineering and Technology in Tirupur under Anna University Chennai. She did her B.E Computer Science and Engineering in Dr.Sivanthi Aditanar College of Engineering in Tiruchendur under Anna University Chennai. She is a member of IAENG, CSI. Her area of interest is Data Mining. Specialized area is Spatial Data Mining and Web Mining. Presented paper on "Arc Based Skyline Query" in a National Conference and has published an international journal in IJERT.

S.Sundaramoorthy working as Assistant Professor in Angel College of Engineering and Technology in Tirupur and had 3 year of experience. He completed his M.TECH Computer Science and Engineering in SNS College of Technology in Coimbatore under Anna University and completed his M.B.A in Bharathiyar University and B.E Computer Science in K.S.Rangasamy College of Technology under Anna University. He has published four papers in International journal and member in ISTE,IAENG. His Area of interest is Data Warehousing and Data Mining.