ISSN 2320 - 2602
Volume 3, No.5, May 2014
International Journal of Advances in Computer Science and Technology
Available Online at http://warse.org/pdfs/2014/ijacst01352014.pdf

Ranking Spatial Data by Quality Preferences Using Spatial Techniques

Pooja Das, Akshay Baravkar, Saarah Sayyed, Prof. P.T Borse
India, poojadas1610@gmail.com
India, akshaybaravkar991@gmail.com
India, saarahsayyed.mail@gmail.com

ABSTRACT

The quality of features of an object is ranked based on the spatial preference query. For example, using a Landed property agency database of flats for lease, a customer may want to rank the flats with respect to the correctness of their location, defined after aggregating the qualities of other features (e.g., restaurants, cafes, hospital, market, etc.) within their spatial neighborhood. Such a neighborhood concept can be specified by the user via various functions. It can be an overt circular region within a given distance from the flat. In this paper, we formally define spatial preference queries and propose suitable indexing techniques and search algorithms for them. Extensively evaluations of our methods on both real and synthetic data reveal that an optimized branch-and-bound solution is resourceful and strong with esteem to different parameters.
Key words: H.2.4.h Query processing, H.2.4.k Spatial databases, Ranking – Top-*k* queries, Optimized Operators.

1. INTRODUCTION

A query to a web search engine usually consists of a list of keywords, to which the search engine responds with the best or "top" k pages for the query [3]. This top-k query model is prevalent over multimedia collections in general, but also over "structured" data for applications where users do not expect exact answers to their queries, but instead a rank of the objects that best match the queries. Top-k spatial preference queries return a ranked set of the k finest data objects depending on the scores of feature objects in their spatial neighborhood. In spite of the widespread range of location-based applications that depend on spatial preference queries, prevailing algorithms sustain non-negligible processing cost resulting in high response time. The reason was that computing the score of a data object requires examining its spatial neighborhood to find the feature object with utmost score. In [1] there is a proposed novel technique to speed up the performance of top-k spatial preference queries, mapping of pairs of data and feature objects to a distance-score space is done , which in turn allows us to identify and materialize the minimal subset of combination that was sufficient to riposte any spatial preference query. In order to handle spatial data efficiently, as required in computer aided design aided -data applications, a database system needs an index mechanism that will help to retrieve data items quickly according to their spatial locations However, traditional indexing techniques were not well suited to data objects of non-zero size located m multidimensional spaces therefore it is described in [2] a dynamic index structure called an R-tree which meets this need, and proposed algorithms for searching and updating it.

1.1 Spatial Techniques

For example, Figure 1 presents a spatial area containing data objects p (hotels) together with feature objects t (restaurants) and v (cafes) with their respective scores (e.g. rating). Consider a tourist interested in hotels with good restaurants and cafes in their spatial neighbourhood. The tourist specifies a spatial constraint (in the figure depicted as a range around each hotel) to restrict the distance was of the eligible feature objects for each hotel. Thus, if the tourist wants to rank the hotels based on the score of restaurants, the top-1 hotel was p3 (0.8) whose score 0.8 was determined by t4. However, if the tourist wants to rank the hotels based on cafes, the top-1 hotel was p1 (0.9) determined by v2. Finally, if the tourist was interested in restaurants and cafes (e.g. summing the scores), the top-1 hotel was p2 (1.2).

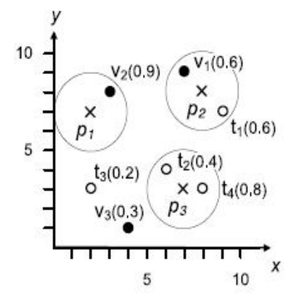

Figure 1: Spatial area containing data and feature objects.

Object ranking is a popular retrieval task in various applications. In relational databases, we rank tuples using an aggregate score function on their attribute values [3].Spatial database systems manage large collections of geographic entities, which apart from spatial attributes contain non-spatial information (e.g., name, size, type, price, etc.). In our paper, we study an interesting type of preference queries, which select the best spatial location with respect to the quality of facilities in its spatial neighborhood [4].

2. SYSTEM ARCHITECTURE

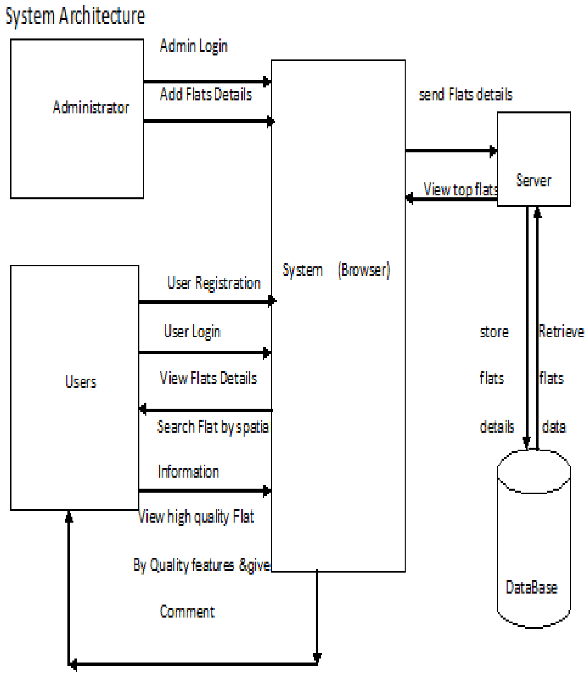

Figure 2: Architecture diagram of spatial database system.

2.1 Architectural Explanation

As we can see the system architecture depicts the heart of the system, the architecture contains administrator, user, system, server and database. Our system interacts with two subsystem or block i.e admin and user. Admin will initially handle the database he will be the first to interact with database. The functions which admin will perform are Admin Login i.e he will login with authorization, the next function performed by admin is to Add Flats Details. Now after Admin role comes the Users interaction with the system. The user will register himself. The user needs to properly login to view flats details. As the user would like to view the flats accordingly to his requirement the flats will be searched according to spatial and non-spatial information stored in database. As per our system he can request flat with various parameters that can be with respect to café, restaurant, market, hospital. These spatial information is ranked objects based on qualities of features in spatial neighborhood. These features were spatial, but in our system we have implemented interesting type of preference queries which apart from spatial attributes also contain non-spatial information (eg name, size, type, price etc).so the user will be able to view flats with high quality features also after viewing the flat he can vote for the flat on basis of quality. When user request to view the flat the system (browser) will retrieve information from database and give him and after viewing flat he again interacts with the system by voting and vote will make changes in database by server dynamically. Now this is all about user, admin and how system works .Now about database, database here in our system is the spatial database which manage large collection of geographical entities which apart from spatial attributes contain non-spatial information where spatial attributes are café, restaurant, hospital, market and non-spatial attributes are name, size, type, price and so on calculated the potential."

3. IMPLEMENTATION

3.1 Optimized Branch and Bound Algorithm

W_k:=new min-heap of size k (initially empty);
γ:=0; k-th score in W_k
algorithm(Node N)
1: V :={e|e∈ N};
2: if N is non-leaf then
3: for c:=1 to m do
4: compute Tc(e) for all e∈ V concurrently;
5: remove entries e in V such that T+(e) <= γ ;
6: sort entries e ∈ V in descending order of T (e);
7: for each entry e ∈ V such that T (e) > γdo
8: read the child node N' pointed by e;
9: BB(N');
10: else
11: for c:=1 to m do
12: computeTc(e)for all e ∈ V concurrently;
13: remove entries e in V such that r+(e) <=γ ;
14: update W_k (andγ) by entries in V ;
Algorithm Find_Result(Node N, Nodes L1,….,Lm)
1: for each entry e∈ N do
2: if N is non-leaf then
3: compute T (e) by entries in L1….,Lm;
4: if T (e) >γ then
5: read the child node N' pointed by e;
6: Find _Result(N', L1,….,.Lm);
7: else
8: compute r (e) by entries in L1,…..,Lm;
9: updateW_k (and γ) by e (when necessary);

3.2 Algorithm Explanation

Spatial preference query ranks objects based on quality of features in their spatial neighbourhood. So the ranking is done by assigning higher weights to features based on proximity of flat. So we have formally define spatial preference queries and proposed appropriate searching algorithms .Here we have evaluated of method on both real and synthetic data using branch and bound solution. In branch and bound what will happen, the objects will be examined but by significantly reducing number of observation whereas earlier in GP (group probing) the procedure was expensive as it examined all objects in D and computed their component scores. As per branch and bound the key idea is to compute for non-leaf entries e in object tree D, an upper bound T(e) of score r(p) for any point p in subtree of e, thus we can save numerous score computations .BB is called with N being the root of D. If N is non-leaf the scores will be computed .T(e) for non-leaf entries e concurrently. The equation will be evaluated for component scores Tc(e) known so far ,we can derive T+(e) an upper bound of T(e) and if T+(e)<=γ then subtree of e cannot contain better result than W_k and it is removed from V.In order to obtain points with high scores early, we sort entries in descending order of T(e) before invoking the above procedure recursively on child entries in V. The branch and bound algorithm was responsible for reducing number of computations .After that the feature join algorithm is used for evaluating top-k spatial preference query by a multiway spatial join on various features whose quality is to be judged.F1, F2….Fm are those features, also the feature join is used to obtain combination of feature points which can be in neighbourhood of some object from D..Hence the object with top-k highest score is retrieved by feature join.

3.3 Abbreviations:

e :an entry in an R-tree
D;the object dataset
T(e):upper bound score of an R-tree entry e of D
γ:represents the top k(ie.lowest score in W_k
W_k:min-heap of size k

4. FLOW-GRAPH

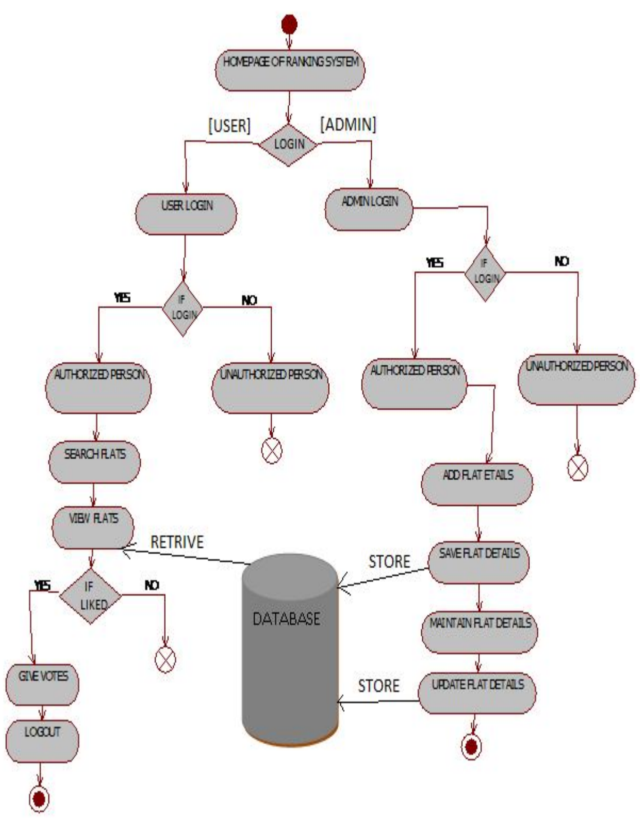

Figure 3: Ranking Spatial data Generation.

4.1 Flow-Graph Explanation

As per the flow of our system first comes the homepage of Ranking Spatial System. After visiting homepage, the person can choose to login, so the person who wants to login can be admin or user .Now we will see these two main components of system one by one. First the admin, if the person is admin he will enter into admin login page and he will enter id and password. If authorized person then he can add flat details and if unauthorized person the flow will end so the admin can add flat details, save the details ie store in database, manipulations in database ie maintain flat details. All this interaction with database is of admin and its work is done .this side can also be said as admin rating. Now on user side if the user is not authorized the flow will stop otherwise if the user is authorized he can search flats then for him the flats will be retrieved from database and viewed to him. After his use of flat details, he can also rate for it and logout. This is how the flow of our system goes.

5. IMPLEMENTED SYSTEM

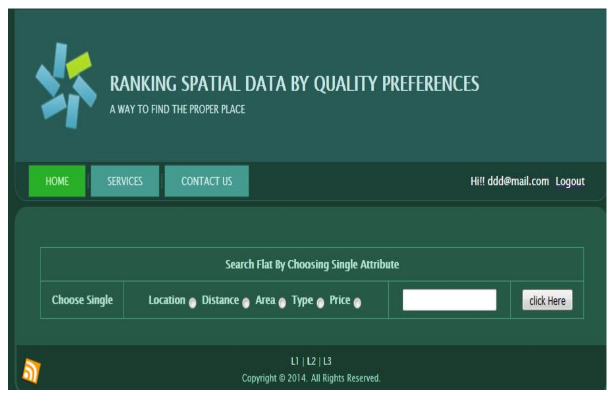

Figure 4: Single Attribute

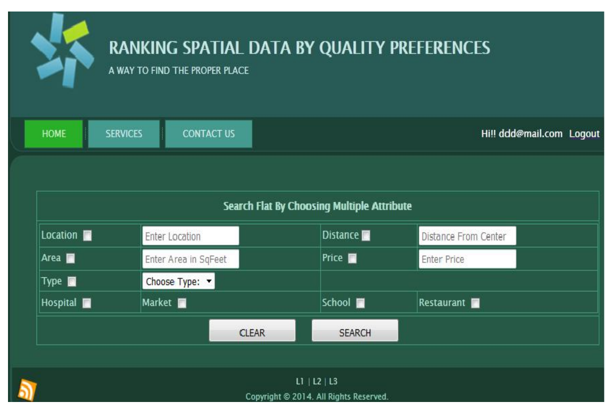

Figure 5: Multiple Attributes

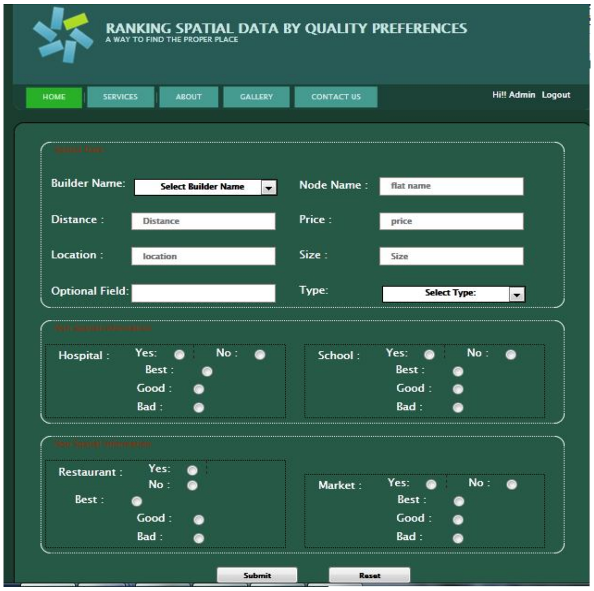

Figure 6: Admin_login

As per our Implementation, spatial preference query ranks the objects based on their spatial neighbourhood including the qualities of features. We used appropriate indexing techniques and search algorithms for calculating the accurate output. As per Figure4 we can see that proper calculating of attributes are fetched from our system, as per Figure4 user can access single attributes considering all Spatial data such as Location, area(sq.),distance and price And Figure5 describes with multiple attributes selection for the user benefit. Combination of both Spatial and non-spatial information result to accurate and perfect output and gives us the ranked data as per the requirement of the user. Figure6 shows the admin rating provided as per the survey of the spatial neighbourhood.

6. RESULT ANALYSIS

Performance is evaluated on basis of factor like ease of development, availability of hardware and re-usable code availability. The feasibility of running software is tested to be of minimum risk, these were selected as a platform for development. Performance is to estimate whether it is possible to develop the proposed system with the available hardware, software and network resources. Since all proposed hardware, software and network requirement are easily available; the development of application became feasible.

7. CONCLUSION

In this paper, we implemented a top-k spatial preference query, which provides a new type of ranking for spatial objects based on qualities of features in their neighborhood. The neighborhood of an object p is captured by the scoring function: (i) the range score restricts the neighborhood to a crisp region centered at p, whereas (ii) the influence score relaxes the neighborhood to the whole space and assigns higher weights to locations closer to p. The algorithm BB derives upper bound scores for non-leaf entries in the object tree, and prunes those that cannot lead to better results. The algorithm BB utilizes an optimized method for computing the scores of objects (and upper bound scores of non-leaf entries). The algorithm FJ performs a multi-way join on feature trees to obtain qualified combinations of feature points and then search for their relevant objects in the object tree. BB is scalable to large datasets and it is the strongest algorithm with respect to various parameters. However, FJ is the best algorithm in cases where the number m of feature datasets is low and each feature dataset is small. In the future, we will study the top-k spatial preference query on road network, in which the distance between two points is defined by their shortest path. The challenge is to develop alternative methods for computing the upper bound scores for a group of points on a road network.

ACKNOWLEDGEMENT

We sincerely thank our guide Prof. Mr. P.T.Borse for his support and valuable guidance. We even like to thank Prof. Mr. Shrikant Dhamdhere (Head of Computer Department at PGMCOE) for all the facilities that were provided.

REFERENCES

1. M. L. Yiu, X. Dai, N. Mamoulis, and M. Vaitis, **Top-k Spatial Preference Queries**, in ICDE, 2007.
2. A. Guttmann, **R-Trees: A Dynamic Index Structure for Spatial Searching**, in SIGMOD, 1984
3. N. Bruno, L. Gravano, and A. Marian, **Evaluating Top-k Queries over Web-accessible** Databases, in ICDE, 2002.
4. Man Lung Yiu, Hua Lu, Nikos Mamoulis, and Michail Vaitis **,Ranking Spatial Data By Quality Preferences,** VOL.23, NO. 3, March 2011