ISSN 2347 - 3983

International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 106- 108 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp21.pdf

CONVESIONAL WAY TO SEARCH COMBINATION OF ITEM SETS FROM MUITDIMESIONAL DATASOURCES

[1] N.Mahendra Babu [2] A.Nageswara rao

[1] M.Tech C.S Student, S.V. College Of Engineering, Tirupati, AP, India, mahendrababu2910@gmail.com
[2] Professor, HOD of CSE, S.V. College Of Engineering, Tirupati, AP, India, hod_cse_svce@svcolleges@edu.in

ABSTRACT:

Clients may search for different type of things to update their knowledge. But Search results depend on the user posed query has to satisfy their searched properties that is stored in the spatial database. For this fast generation, we need a conventional way of searching is required to develop modern application. Here our aim is search has to meet two things that spatial predicate and associated texts. Suppose a client searching for items that is available in a restaurant which is close to his circle is the condition. In spite of comparing all searches related items (restaurants) for accurate one. Just think about the restaurant which is closed to the searcher that contains menu items like "steaks, spaghetti, brandy" available at the same time. At present the great solution to solve this problem of searching is possible by IR2 tree but, it is insufficient to get an accurate result exactly. For that we newly introduced a spatial inverted index algorithm that extends the gap of IR2 tree which is lagging to compute the multidimensional data. We are evaluating this experimental by posing different queries.

Keywords: spatial database & inverted index, keywords, neighbor search, IR-Tree, Approximation algorithm and Exact algorithm, BR*-tree.

I. INTRODUCTION

A spatial database is a database that store multidimensional objects such as points, rectangles, and etc. some spatial databases allow representing simple geometric objects such as lines, points and polygons. Some spatial databases handle more complex structures such as 3D objects, topological coverage's, linear networks. Based on different selection criteria spatial database provides fast access to multidimensional objects. In spatial database real entities are modeled in geometric manner, for example location of hotels, hospital, restaurants are represented as points on maps, while larger area such as landscapes, lakes, parks are represented as a combination of rectangles. Spatial database system can use in geographic information system, in this range search can be utilized to find all restaurants in a certain area, while nearest neighbor retrieval can find the restaurant closest to a given address.

Queries in spatial database have become increasingly important in recent years with the increasing popularity of some services such as Google Earth and Yahoo Maps, as well as other geographic applications. Today, widely used by search engines has made it realistic to write spatial queries in a new way. Traditionally, queries focus on objects only geometric properties, for example, whether a point is in rectangle or how two points are close from each other. Some new application allows users to browse objects based on both of their geometric coordinates and their associated texts. Such type of queries called as a spatial keyword query. For example, if a search engine can be used to find the nearest hotel that offer facilities such as pool and internet at the same time. From this query, we could first obtain the entire hotel whose services contains the set of keywords, and then find the nearest one from the retrieved restaurant. The major drawback of this approach is that, on the difficult input they do not provide real time answer. For example, from the query point the real neighbor lies quite far away, while all the closer neighbors are missing at least one of the query keywords. Spatial keyword queries have not been widely explored. In the past years, the group of people has showed interest in studying keyword search in relational databases. Recently the attention has preoccupied to multidimensional data [5][6]. The best method for nearest neighbor search with keywords is because of Felipe et al. [5]. They combine the spatial index R-tree [7] and signature file [8]. So they developed a structure called IR tree. This tree has the ability of both R-tree and signature files. Like an R-tree it stores the spatial proximity of the object and like signature file it filters those objects that do not include all query keywords.

II RELATED WORK

For search related query's keywords play a crucial role in effective searching. Researchers study it for many years to enhance the commercial search engine performance. We have different types of search techniques are being proposed in earlier, but they don't perform their task effectively. To do this task effectively, we have broken it into two sections. First, we have given rank for search related items based on priority.

ISSN 2347 - 3983

International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 106- 108 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp21.pdf

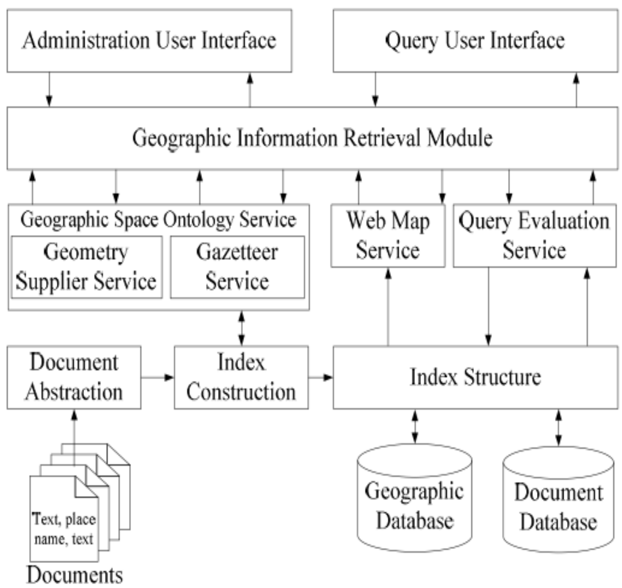

Fig 1: System Design

Query progressing have two sections, converting it into a conventional way to search the items quickly and select the appropriate search technique which is related to the situation. These keywords are stored in the form of the matrix, then we are converting it into the required form, another section we are iteratively checking for constructing the inverted list which is easy for searching. With this we can easily compare the search related things, find out the distance and time by following modules.

III IMPLEMENTATION

A. IR-Tree, Approximation algorithm and Exact algorithm:

This method is used to retrieve a group of spatial web objects such that the query's keywords are covered by group's keywords and objects are near to the query location and have the lowest inter object distances. This method addresses the two instantiation of the group keyword query. First is to find the group of objects that cover the keywords such that the sum of their distances to the query is minimized. Second is to find a group of objects that covers the keywords such that between of the maximum distance between objects in a group of objects and query and maximum distance among two objects in a group of objects is minimized. Both of these sub problems are NP-complete. Greedy algorithm is used to provide an approximation solution to the problem that utilizes the spatial keyword index IR-tree to reduce the search space. But in some application query does not contain a large number of keywords, for this exact algorithm is used that uses the dynamic programming.

B. IUR-tree (Intersection union R-tree)

Geographic objects associated with descriptive texts are becoming common. This gives importance to spatial keyword queries that take both the location and a text description of the content. This technique is used to analyze the problem of reverse spatial and textual k nearest neighbor search i.e finding objects that take the query object as one of their spatial textual similar objects. For this type of search hybrid index structure is used that successfully merge the location proximity with textual similarity. For searching, branch and bound algorithm are used. In addition, to increase the speed of query processing a variant of IRtree and two optimization algorithm is used. To enhance the IU R-tree, text clustering is used, in this object of all the databases is grouped into clusters according to their text similarity. Each node of the tree is extended by the cluster information to create a hybrid tree which is called as cluster IUR-tree. To enhance the search performance of this tree two optimization methods are used, the first is based on outlier detection and extraction and second method is based on text entropy.

B. BR-tree:*

This hybrid index structure is used to search m-closest keywords. This technique finds the closest tuples that match the keywords provided by the user. This structure combines the R*-tree and bitmap indexing to process the closest keyword query that returns the spatially closest objects matching keywords To reduce the search space a priori based search strategy is used. Two monotone constraints are used as a priori properties to facilitate efficient pruning which is called as distance matrix and keyword mutex. But this approach is not suitable for handling ranking queries and in this number of false hits is large.

C. IR2-trees:

The growing number of applications requires the efficient execution of nearest neighbor queries which is constrained by the properties of spatial objects. Keyword search is very popular on the internet so these applications allow users to give list of keywords that spatial objects should contain. Such queries called as a spatial keyword query. This consists of query area and a set of keywords. The IR2-tree is developed by the combination of R-tree and signature files, where each node of the tree has spatial and keyword information.

ISSN 2347 - 3983

International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 106- 108 (2015)
Special Issue of NCTET 2K15 - Held on June 13, 2015 in SV College of Engineering, Tirupati
http://warse.org/IJETER/static/pdf/Issue/NCTET2015sp21.pdf

```
IR2NearestNeighbor(p,W,U)
  while not U.IsEmpty()
    E ← U.Dequeue()
    if E is a non-Leaf Node
      for each (NodePtr,MBR,S) in E
        if S matches W
          U.Enqueue(LoadNode(NodePtr),Dist(p,MBR))
    else if E is a Leaf Node
      for each (ObjPtr,MBR,S) in E
        if S matches W
          U.Enqueue(ObjPtr,Dist(p,MBR))
    else /* E is an object pointer */
      return E as next nearest object pointer to p
IR2TopK(R,Q)
  initialize a list L
  Initialize a priority queue U
  U.Enqueue(R.RootNode,0)
  W ← Signature(Q.t)
  c ← 0
  while c < Q.k
    ObjtPtr ← IR2NearestNeighbor(Q.p,W,U)
    T ← LoadObject(ObjPtr)
    if T.t contains all keywords in Q.t
      c ← c + 1
      L.add(T)
  return L
```

Fig 2: IR2 tree algorithm

This method is efficient answering the top-k spatial keyword queries. In this signature is added to the every node of the tree. An able algorithm is used to answer the queries using the tree. An Incremental nearest algorithm is used for the tree traversal and if the root node signature does not match the query signature then it prunes the whole subtrees. But IR2-tree has some drawbacks such as false hits where the object of final result is far away from the query or this is not suitable for handling ranking queries.

E. Spatial inverted index and Minimum bounding method:

So, new access method, spatial inverted access method is used to remove the drawbacks of previous methods such as false hits.

```
NearestNeighbor(p,U)
    /* priority queue U initially contains root node of R with distance 0 */
  while not U.IsEmpty()
    E ← U.Dequeue()
    if E is a non-Leaf Node
      for each (NodePtr,MBR) in E
        U.Enqueue(LoadNode(NodePtr),Dist(p,MBR))
    else if E is a Leaf Node
      for each (ObjPtr,MBR) in E
        U.Enqueue(ObjPtr,Dist(p,MBR))
    else /* E is an object pointer */
      return E as next nearest object pointer to p
```

Fig 3: incremental nearest neighbour algorithm

This method is the variant of inverted index using for multidimensional points. This index stores the spatial region of data points and on every inverted list Rtree is built. The Minimum bounding method is used for traversing the tree to prune the search space.

IV. CONCLUSION

In our proposal different techniques are used to search for nearest neighbour things from a spatial database by overcoming various drawbacks like space consumption, efficiency to calculate the large amount of data, mainly it gets related objects which is taken from scattered data from different sources but it was not giving accurate results because it is not comparing the search related things along with their properties. For this we use a spatial index algorithm for point out where exactly the data situated at the source and for comparing the search related thing IR2-tree is used to give exact result.

REFERENCES

[1] X. Cao, G. Cong, C.S. Jensen, and B.C. Ooi, “Collective Spatial Keyword Querying,” Proc. ACM SIGMOD Int’l Conf. Management of Data, pp. 373-384, 2011.
[2] J. Lu, Y. Lu, and G. Cong, “Reverse Spatial and Textual k Nearest Neighbor Search,” Proc. ACM SIGMOD Int’l Conf. Management of Data, pp. 349-360, 2011.
[3] D. Zhang, Y.M. Chee, A. Mondal, A.K.H. Tung, and M. Kitsuregawa, “Keyword Search in Spatial Databases: Towards Searching by Document,” Proc. Int’l Conf. Data Eng. (ICDE), pp. 688-699, 2009.
[4] G. Cong, C.S. Jensen, and D. Wu, “Efficient Retrieval of the Top-k Most Relevant Spatial Web Objects,” PVLDB, vol. 2, no. 1, pp. 337- 348, 2009.
[5] I.D. Felipe, V. Hristidis, and N. Rishe, “Keyword Search on Spatial Databases,” Proc. Int’l Conf. Data Eng. (ICDE), pp. 656-665, 2008.
[6] Yufei Tao and Cheng Sheng, “Fast Nearest Neighbor Search with Keywords”, IEEE transactions on knowledge and data engineering, VOL. 26, NO. 4, APRIL 2014.
[7] N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger, “The R - tree: An Efficient and Robust Access Method for Points and Rectangles,” Proc. ACM SIGMOD Int’l Conf. Management of Data, pp. 322-331, 1990.
[8] C. Faloutsos and S. Christodoulakis, “Signature Files: An Access Method for Documents and Its Analytical Performance Evaluation,” ACM Trans. Information Systems, vol. 2, no. 4, pp. 267-288, 1984.
[9] G. R. Hjaltason and H. Samet. Distance browsing in spatial databases. ACM Transactions on Database Systems (TODS), 24(2):265–318, 1999s.