ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

New Access Method Where Queries Give Results That Satisfy Spatial and Text Predicates

K.Anil Kumar
M.Tech Student,
CMR Institute of Technology.

D.Shanthi
Assistant professor,
CMR Institute of Technology.

Dr.S.Arvind
HOD,
CMR Institute of Technology.

Abstract:

A spatial query takes a location and given keywords as arguments and returns objects that are ranked according to both spatial proximity and text relevance relative to the query. Spatial queries like nearest neighbor retrieval and range search, occupy only conditions on geometric properties of object. For finding objects which are satisfying a spatial predicate and predicate on their associated texts, novel form of queries are called by many applications. Presently the most effective resolution to such queries is predicated on the IR2-tree, which, as shown during this paper, features a few deficiencies that seriously impact its potency. Impelled by this, we tend to develop a replacement access methodology known as the abstraction inverted index that extends the standard inverted index to address flat knowledge, and comes with algorithms that may answer nearest neighbor queries with keywords in real time. As verified by experiments, the projected techniques outgo the IR2-tree in question latent period considerably, typically by an element of orders of magnitude.

Keywords:

IR Tree, Search, Spatial Database, Test, maps.

Introduction:

Nearest neighbor search (NNS), also known as closest point search, similarity search. It is an optimization problem for finding closest (or most similar) points. Nearest neighbor search which returns the nearest neighbor of a query point in a set of points, is an important and widely studied problem in many fields, and it has wide range of applications. We can search closest point by giving keywords as input; it can be spatial or textual.

A spatial database use to manage multidimensional objects i.e. points, rectangles, etc. Some spatial databases handle more complex structures such as 3D objects, topological coverage's, linear networks. While typical databases are designed to manage various NUMERIC'S and character types of data, additional functionality needs to be added for databases to process spatial data type's efficiently and it provides fast access to those objects based on different selection criteria.

Keyword search is the most popular information discovery method because the user does not need to know either a query language or the underlying structure of the data. The search engines available today provide keyword search on top of sets of documents. When a set of query keywords is provided by the user, the search engine returns all documents that are associated with these query keywords.

Solution to such queries is based on the IR2-tree, but IR2- tree having some drawbacks. Efficiency of IR2-tree badly is impacted because of some drawbacks in it. The solution for overcoming this problem should be searched. Spatial inverted index is the technique which will be the solution for this problem. Spatial database manages multidimensional data that is points, rectangles.

Related Work:

IR2 – Tree

The IR2 – Tree combines the R-Tree and signature file. First we will review Signature files. Then IR2-trees are discussed. Consider the knowledge of R-trees and the best- first algorithm for Near Neighbor Search. Signature file is known as a hashing-based framework and hashing -based framework is which is known as superimposed coding (SC).

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Drawbacks of the IR2-Tree:

IR2-Tree is first access method to answer nearest neighbour queries. IR2-tree is popular technique for indexing data but it having some drawbacks, which impacted on its efficiency. The disadvantage called as false hit affecting it seriously. The number of 940 false positive ratio is large when the aim of the final result is far away from the query point and also when the result is simply empty.In these cases, the query algorithm will load the documents of many objects; as each loading necessitates a random access, it acquires costly overhead.

Keyword search on spatial databases:

This work, mainly focus on finding top-k Nearest Neighbors, in this method each node has to match the whole querying keywords. As this method match the whole query to each node, it does not consider the density of data objects in the spatial space. When number of queries increases then it leads to lower the efficiency and speed. They present an efficient method to answer top-k spatial keyword queries. This work has the following contributions: 1) the problem of top-k spatial keyword search is defined. 2) The IR2-Tree is proposed as an efficient indexing structure to store spatial and textual information for a set of objects. There are efficient algorithms are used to maintain the IR2-tree, that is, insert and delete objects. 3) An efficient incremental algorithm is presented to answer top-k spatial keyword queries using the IR2-Tree. Its performance is estimated and compared to the current approaches. Real datasets are used in our experiments that show the significant improvement in execution times.

Disadvantages:

1. Each node has to match with querying keyword. So it affects on performance also it becomes time consuming and maximizing searching space.

Processing Spatial-Keyword (SK) Queries in Geographic Information Retrieval (GIR) Systems:

Location based information stored in GIS database. These information entities of such databases have both spatial and textual descriptions. This paper proposes a framework for GIR system and focus on indexing strategies that can process spatial keyword query.

The following contributions in this paper: 1) It gives framework for query processing in Geo- graphic Information Retrieval (GIR) Systems. 2) Develop a novel indexing structure called KR*-tree that captures the joint distribution of keywords in space and significantly improves performance over existing index structures. 3) This method have conducted experiments on real GIS datasets showing the effectiveness of our techniques compared to the existing solutions. It introduces two index structures to store spatial and textual information.

A) Separate index for spatial and text attributes: Advantages:

1. Easy of maintaining two separate indices.

2. Performance bottleneck lies in the number of candidate object generated during the filtering stage.

Disadvantages:

1. If spatial filtering is done first, many objects may lie within a query is spatial extent, but very few of them are relevant to query keywords. This increases the disk access cost by generating a large number of candidate objects. The subsequent stage of keyword filtering becomes expensive.
B) Hybrid index

Advantages and limitations:

1.When query contains keywords that closely correlated in space, this approach suffer from paying extra disk cost accessing R*-tree and high overhead in subsequent merging process.

EXISTING SYSTEM:

Spatial queries with keywords have not been extensively explored. In the past years, the community has sparked enthusiasm in studying keyword search in relational databases. It is until recently that attention was diverted to multidimensional data. Existing works mainly focus on finding top-k Nearest Neighbours, where each node has to match the whole querying keywords .It does not consider the density of data objects in the spatial space. Also these methods are low efficient for incremental query.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

DISADVANTAGE:

Failure of finding objects in space.

PROPOSED SYSTEM:

A spatial database manages multidimensional objects (such as points, rectangles, etc.), and provides fast access to those objects based on different selection criteria. The importance of spatial databases is reflected by the convenience of modeling entities of reality in a geometric manner. For example, locations of restaurants, hotels, hospitals and so on are often represented as points in a map, while larger extents such as parks, lakes, and landscapes often as a combination of rectangles.

Many functionalities of a spatial database are useful in various ways in specific contexts. For instance, in a geography information system, range search can be deployed to find all restaurants in a certain area, while nearest neighbor retrieval can discover the restaurant closest to a given address.

ADVANTAGE:

Finding objects in space by using Google-Maps.

PROBLEM STATEMENT:

There is a disadvantage of finding objects in space from the existing system. In this proposal, we are going to allocate the objects in space by using spatial database.

SCOPE:

We are going to search the nearest neighbor.

Architecture:

Modules :

1.Registration
2.Login
3.Registration
4.Search Techniques
5.Map_view
6.Distance_Search

Modules Description

Registration:

In this module an User have to register first,then only he/she has to access the data base.

Login:

In this module,any of the above mentioned person have to login,they should login by giving their email id and password .

Registration:

In this module Admin registers the hotel along with its famous dish. Also he measures the distance of the corresponding hotel from the corresponding source place by using spatial distance of Google map.

Search Techniques:

Here we are using two techniques for searching the document 1)Restaurant Search,2)Key Search.

Key Search:

It means that the user can give the key in which dish that the restaurant is famous for .This results in the list of menu items displayed.

Restaurant Search:

It means that the user can have the list of restaurants which are located very near. List came from the database.

Map_View:

The User can see the view of their locality by Google Map(such as map view, satellite view) .

Distance_Search:

The User can measure the distance and calculate time that takes them to reach the destination by giving speed. Chart will be prepared by using these values. These are done by the use of Google Maps.

Conclusion:

In this report, we have surveyed a Searching Nearest Neighbor based on Keywords using Spatial Inverted Index and evaluate the needs and challenges present in Nearest Neighbor Search. This report covers existing techniques for that and also covers upon new improvements in current technique. In this paper, we have surveyed topics like IR2 – Tree, Drawbacks of the IR2-Tree, Spatial keyword search, Solutions based on Inverted Indexes.

References:

[1] Yufei Tao & Cheng Sheng, Fast Nearest Neighbor Search with Keywords, IEEE Transactions on Knowledge and Data Engineering (TKDE), (Volume:26 , Issue: 4)

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] I.D. Felipe, V. Hristidis, and N. Rishe, "Keyword Search on Spatial Databases," Proc. Int'l Conf. Data Eng. (ICDE), pp. 656-665, 2008.

[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, 1999.