International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

Find The Nearest Hotels on Google Map Using IR2-tree & Spatial inverted index

Hemalatha Kanumala[1]& Dr. Shaik Abdul Muzeer[2]
[1]M-Tech Dept. of CSE Megha Institute of Engineering & Technology for Women
[2]Professor & principal Dept. of CSE Megha Institute of Engineering &Technology for Women

Abstract–
Many search engines are used to search anything from anywhere; this system is used to fast nearest neighbor search using keyword. Existing works mainly focus on finding top-k Nearest Neighbors, 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. But in intended system, for example when there is search for nearest restaurant, instead of considering all the restaurants, a nearest neighbor query would ask for the restaurant that is, closest among those whose menus contain spicy, brandy all at the same time, 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. The spatial inverted index is the technique which will be the solution for this problem. Conventional spatial queries, nearest neighbor retrieval and range search consists only conditions on objects geometric property only. But today, many modern applications support new form of queries that aim to find objects that satisfies both spatial data and their associated text. For example instead of considering all the hotels, a nearest neighbor query would instead ask for the hotel that is closest to among those who provide services such as pool, internet at the same time. For this type of query a variant of inverted index is used that is effective for multidimensional points and comes with an R-tree which is built on every inverted list, and uses the algorithm of minimum bounding method that can answer the nearest neighbor queries with.

Keywords: -Nearest Neighbor Search; IR2-tree; nearest; Range search; Spatial inverted index

1. INTRODUCTION

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. There are easy ways to support queries that combine spatial and text features. For example, for the above query, we could first fetch all the restaurants whose menus contain the set of keywords {steak, spaghetti, brandy}, and then from the retrieved restaurants, find the nearest one. Similarly, one could also do it reversely by targeting first the spatial conditions – browse all the restaurants in ascending order of their distances to the query point until encountering one whose menu has all the keywords. The major drawback of these straightforward approaches is that they will fail to provide real time answers on difficult inputs. A typical example is that the real nearest neighbor lies quite faraway This paper gives importance spatial queries with keywords [5] [6] [9] [10]. Spatial queries with keywords take arguments like location and specified keywords and provide web objects that are arranged depending upon spatial proximity and text relevancy. Some other approaches take keywords as Boolean predicates [1] [2], searching out web objects that contain keywords and rearranging objects based on their spatial proximity. Some approaches use a linear ranking function [7] [8] to combine spatial proximity and textual relevance. Earlier study of keyword search in relational databases is gaining importance. Recently this attention is diverted to multidimensional data [3] [4]. N. Rishe, V. Hristidis and D. Felipe has proposed best method to develop neighbor search with keywords. For keyword-based retrieval, they have integrated R-tree with spatial index and signature file. By combining R-tree and signature they have developed a structure called the IR2-tree IR2-tree has merits of both R-trees and signature files. The IR2-tree preserves object's spatial proximity which important for solving spatial queries.

2. RELATED WORK

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. The best method to date for nearest neighbour search with keywords is due to Felipe et al. They nicely integrate two well-known concepts: R-tree, a popular spatial index, and signature file, an effective method for keyword-based document retrieval. By doing so they develop a structure called the IR2 -tree, which has the strengths of both R-trees and signature files. Like R-trees, the IR2 - tree preserves objects' spatial proximity, which is the key to solving spatial queries efficiently. On the other hand, like signature files, the IR2 -tree is able to filter a considerable portion of the objects that do not contain all the query keywords, thus significantly reducing the number of objects to be examined.

Disadvantages of existing system:

Fail to provide real time answers on difficult inputs.

The real nearest neighbor lies quite far away from the query point, while all the closer neighbors are missing at least one of the query keywords.

Proposed system:

In this paper, we design a variant of inverted index that is optimized for multidimensional points, and is thus named the spatial inverted index (SI-index). This access method successfully incorporates point coordinates into a conventional inverted index with small extra space, owing to a delicate compact storage scheme.

Meanwhile, an SI-index preserves the spatial locality of data points, and comes with an R-tree built on every inverted list at little space overhead. As a result, it offers two competing ways for query processing.

We can (sequentially) merge multiple lists very much like merging traditional inverted lists by ids. Alternatively, we can also leverage the R-trees to browse the points of all relevant lists in ascending order of their distances to the query point. As demonstrated by experiments, the SI-index significantly outperforms the IR2 -tree in query efficiency, often by a factor of orders of magnitude.

Advantages of proposed system:

Distance browsing is easy with R-trees. In fact, the best-first algorithm is exactly designed to output data points in ascending order of their distances

It is straight forward to extend our compression scheme to any dimensional space

3. IMPLEMENTATION

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 [10] 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)[10].

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

Hybrid Index Structures for Location-based Web Search.

There is more and more research interest in location-based web search, i.e. searching web content whose topic is related to a particular place or region. This type of search contains location information; it should be indexed as well as text information. text search engine is set-oriented whereas location information is two-dimensional and in Euclidean space. In previous paper we see same two indexes for spatial as well as text information. This creates new problem, i.e. how to combine two types of indexes. This paper uses hybrid index structure, to handle textual and location based queries, with help of inverted files and R*-trees. It considered three strategies to combine these indexes namely: 1) inverted file and R*-tree double index.2) first inverted file then R*-tree.3) first R*-tree then inverted file. It implements search engine to check performance of hybrid structure, that contains four parts:(1) an extractor which detects geographical scopes of web pages and represents geographical scopes as multiple MBRs based on geographical coordinates. (2) The work of indexer is use to build hybrid index structures integrate text and location information. (3) The work of ranker is to ranks

Spatial inverted index

The spatial inverted list (SI-index) is essentially a compressed version of an I-index with embedded coordinates. Query processing with an SI-index can be done either by merging, or together with R-trees in a distance browsing manner. Furthermore, the compression eliminates the defect of a conventional that an SI-index consumes much less space

Registration: In this module a User have to register first, and 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.

Hotel 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

4. EXPERIMENTAL RESULTS

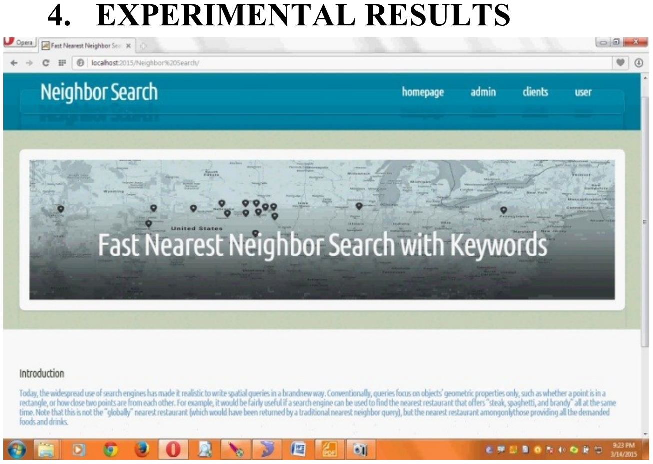

Fig:-1 Main Frame

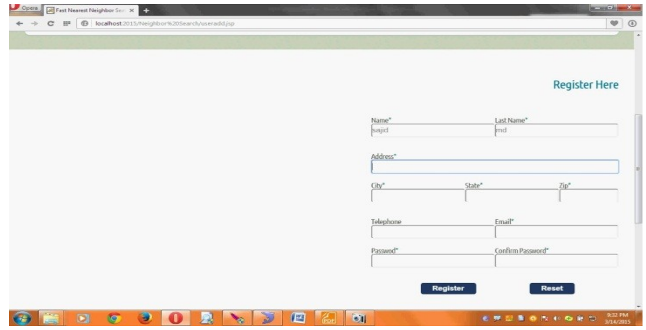

Fig:-2 User Registration

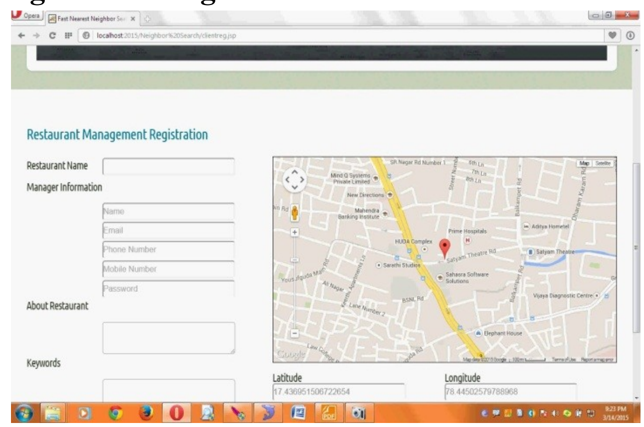

Fig:-3 Hotels Registration

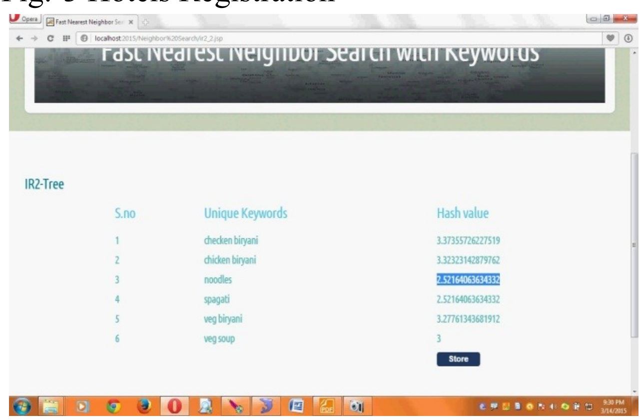

Fig:-4 working with IR2 Tree

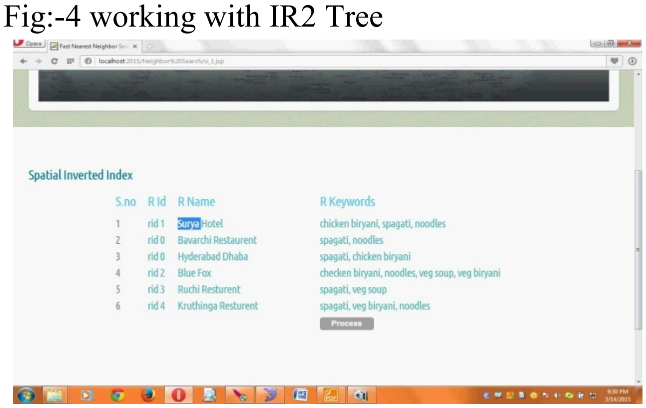

Fig:-5 Spatial Inverted Index

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

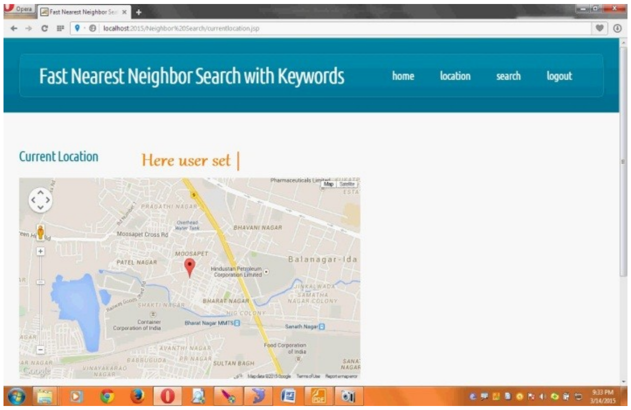

Fig:-6 Results On Map

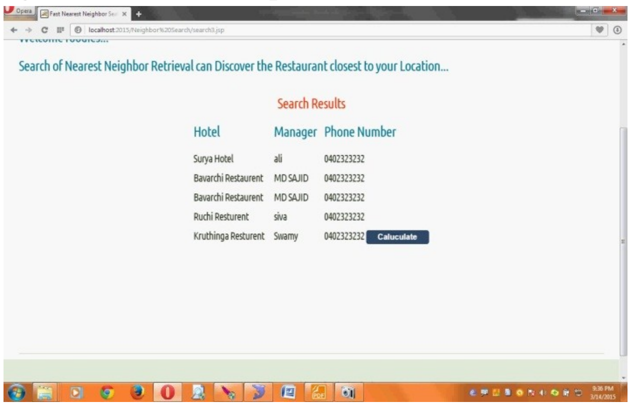

Fig:-7 Nearest Results Collection

5. CONCLUSION

We have seen plenty of applications calling for a search engine that is able to efficiently support novel forms of spatial queries that are integrated with keyword search. The existing solutions to such queries either incur prohibitive space consumption or are unable to give real time answers. In this paper, we have remedied the situation by developing an access method called the spatial inverted index (SIindex). Not only that the SI-index is fairly space economical, but also it has the ability to perform keyword augmented nearest neighbour search in time that is at the order of dozens of milli-seconds. Furthermore, as the SIindex is based on the conventional technology of inverted index, it is readily incorporable in a commercial search engine that applies massive parallelism, implying its immediate industrial merits.

6. REFERENCES

[1]. S. Agrawal, S. Chaudhuri, and G. Das.Dbxplorer: A system for keyword-based search over relational databases. In Proc. OfInternational Conference on Data Engineering (ICDE), pages 5–16, 2002.

[2]. N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger. The R*-tree: An efficient and robust access method for points and rectangles. In Proc. of ACM Management of Data (SIGMOD), pages 322–331, 1990.

[3]. G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan.Keyword searching and browsing in databases using banks. In Proc. of International Conference on Data Engineering (ICDE), pages 431– 440, 2002.

[4]. X. Cao, L. Chen, G. Cong, C. S. Jensen, Q. Qu, A. Skovsgaard, D. Wu, and M. L. Yiu. Spatial keyword querying. In ER, pages 16–29, 2012.

[5]. X. Cao, G. Cong, and C. S. Jensen.Retrieving top-k prestige-based relevant spatial web objects.PVLDB, 3(1):373–384, 2010.

[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, 1999.

[10]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.

Author Profile

Dr.Shaik Abdul Muzeer
Professor & Principal
Megha institute of Engineering & Technology for Women

Dr.S.A.Muzeer, at present working as a principal of Megha institute of engineering & Technology has completed his PG and P.HD in Electronics & Communication Engineering and published around 25 Papers in National & International Journals. His area of research is Digital signal processing and Bio-medical engineering