IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 5, May 2016

Fast Nearest Neighbor Search with Keywords

Sarita Ahire[1], Anup Vishwakarma[1], Bhagwat Chavan[1], Vinay Verma [1]

Student, Department of Computer, P.E.S.MCOE, Savitribai Phule Pune University, Pune, India[1]

Abstract: Conventional spatial queries, such as nearest neighbor retrieval, involve only conditions on objects geometric properties. Today, many modern applications call for novel forms of queries that aim to find objects satisfying both a spatial predicate, and a predicate on their associated texts. For example, instead of considering all the restaurants, a nearest neighbor query would instead ask for the restaurant that is the closest among those whose menus contain steak, brandy all at the same time. Currently the best solution to such queries is based on the IR2-tree, has a few deficiencies that seriously impact its efficiency. Motivated by this, we develop a new access method called the spatial inverted index that extends the conventional inverted index to cope with multidimensional data, and comes with algorithms that can answer nearest neighbor queries with keywords in real time. As verified by experiments, the proposed techniques outperform the IR2-tree in query response time significantly, often by a factor of orders of magnitude.

Keywords: Nearest Neighbor Search, Keyword Search, Spatial Index.

I. INTRODUCTION

Today, many modern applications call for novel forms of queries that aim to find objects satisfying both a spatial predicate, and a predicate on their associated texts. For example, instead of considering all the restaurants, a nearest neighbor query would instead ask for the restaurant that is the closest among those whose menus contain steak, brandy all at the same time

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.his document is a template.

II. RELATED WORK

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 neighbor search with keywords is due to Felipeetal. 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. The IR2 -tree, however, also inherits a drawback of signature files: false hits.

We are going 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.

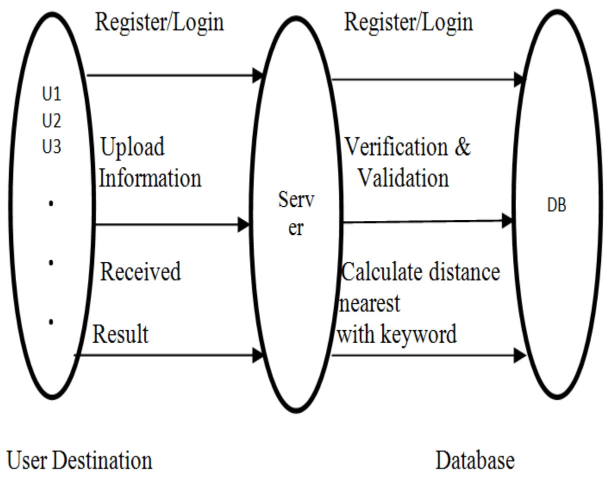

Fig. Venn diagram

III. LITERATURE REVIEW

TITLE	DETAIL			CONTENT
	Author name	**Journal name**	**Year**	
Spatial keyword querying	X. Cao, L. Chen	IEEE	2012	The techniques outperform the IR2-tree in query response time significantly.
Reverse spatial and textual k nearest neighbor search	G.Cong Et. al	ACM management of Data (SIGMOD	2011	RSTkNN query, which is the extension of RkNN query with the fusion of spatial information and textual description

IV. PROPOSED ALGORITHM

- **Description of the Proposed Algorithm:**

Aim of the proposed algorithm is to minimizing the time by calculating distance from source to destination.
The proposed algorithm is consists of following steps.

Step 1: Let Hl - list of hotels along with their latitude and longitude
Hm - be the menu available in respective hotel.
Step 2: Inverted index is generated first and preserved (HI).
Step 3: Keyword search is done by the user.
Step 4: HLc - list of hotel having all menu which satisfies query
Keyword search is done in menu of every hotel of that geometric region.
Step 5: R-tree is generated by merging all the points in HLc.
Step 6: Distance calculation is done using R-tree.

a = Math.sin(dLat/2) * Math.sin(dLat/2)+ Math.cos(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat2)) * Math.sin(dLng/2) * Math.sin(dLng/2);
c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
dist = (float) (earthRadius * c)* 0.001f;

Step 7: Hotels are displayed to the user in ascending order of their distance.

V. RESULTS

The simulation results involve the admin home page, admin login, hotel registered, hotels with menu, keyword search, login by admin or user, result for hotels with distance, user home page, user registration.

- **Admin home page**

The user details such as userid, user name, gender, DOB, email id, mobile number are store or display on admin home page. By this details only registered user can access this system.

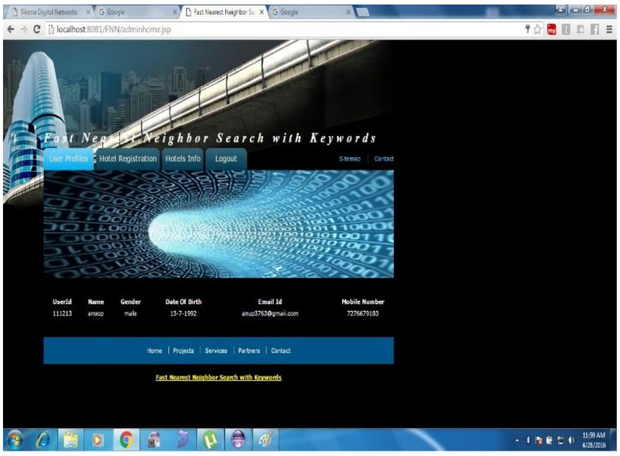

Fig1. Admin home page

- **Admin login**

Admin can registered here by id and password only admin can change, modify and delete the contents on system.

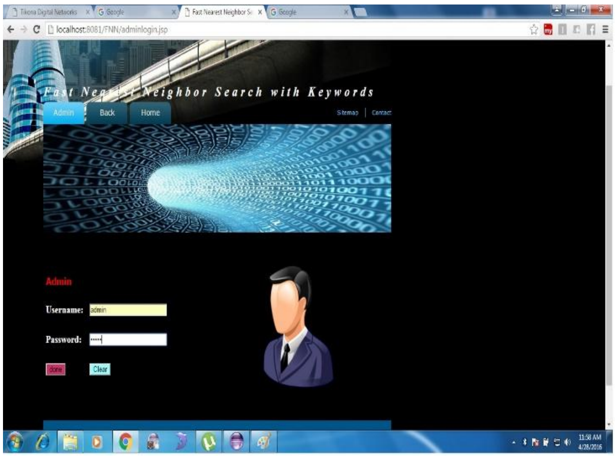

Fig2.Admin login

- **Registered hotels**

Here hotels are registered with hotel id, latitude and longitude with some actions.

Fig3. Hotels registration

- **Hotel with menu**

Here shows the hotels name with its menu to the user.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 5, May 2016

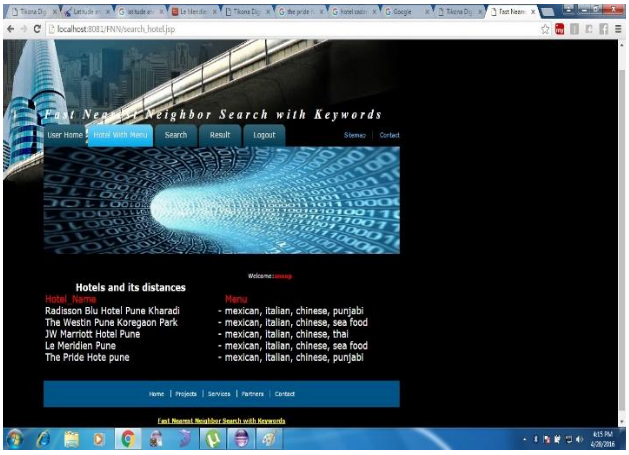

Fig4. Hotels with menu

- **Keyword Search:**

Here keyword is search with the user position by get location and enter the keyword.

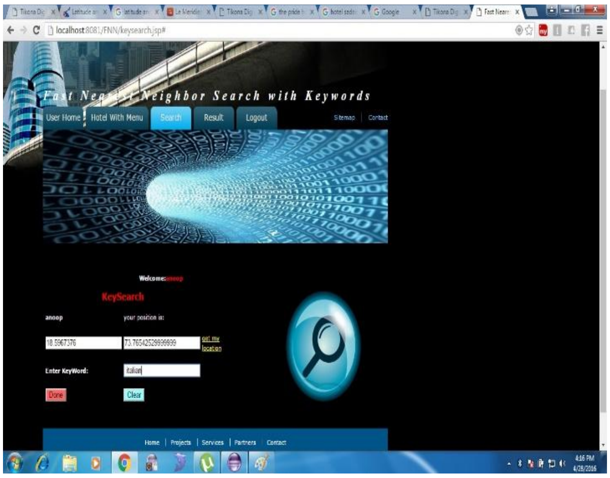

Fig5. Keyword Search

- **Login Page**

Here admin and user can login.

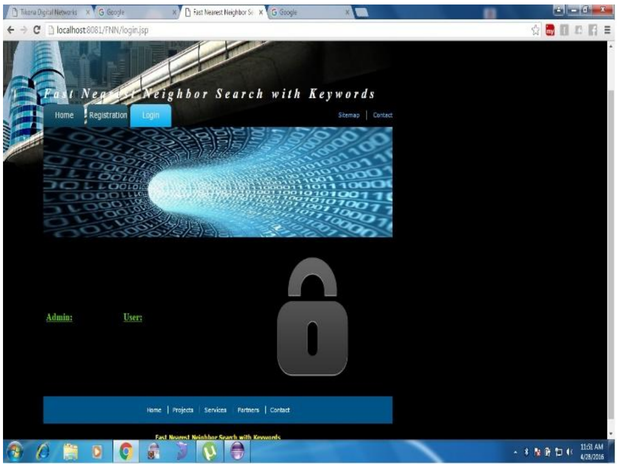

Fig6. Login page

- **Result:**

Here actual result is shown by the hotels with hotel name and its distance.

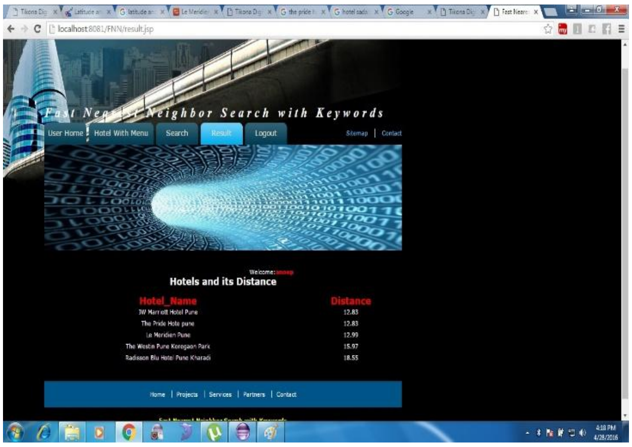

Fig7. Result

- **User Home Page:**

On user home page show the result by Map view of Nearest restaurant with time calculation.

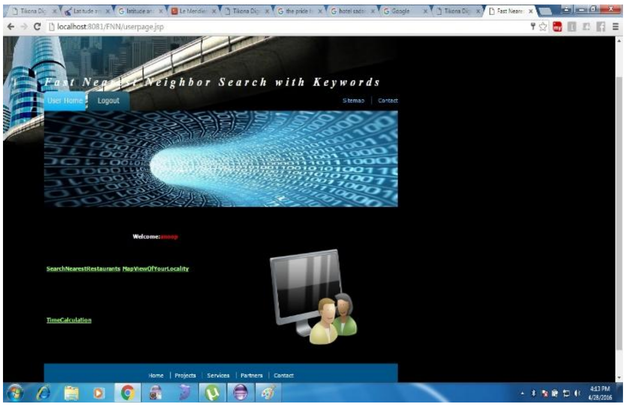

Fig8. User home page

- **User Registration:**

Here user can registered by user name, DOB, mob number, emailed etc. only registered user can access the system for their purpose.

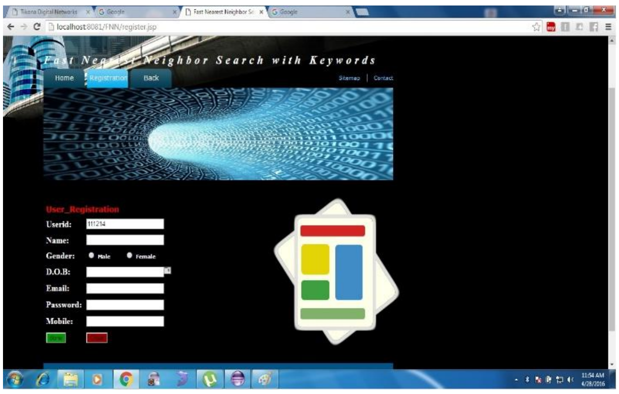

Fig9. User Registration

VI. 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

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 3, Issue 5, May 2016

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 (SI-index). Not only that the SI-index is fairly space economical, but also it has the ability to perform keyword-augmented nearest neighbor search in time that is at the order of dozens of milliseconds. Furthermore, as the SI-index 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.

REFERENCES

[1] S. Agrawal, S. Chaudhuri, and G. Das. Dbxplorer: A system for keyword-based search over relational databases. In Proc. Of International Conference on Data Engineering (ICDE), pages 516, 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 322331, 1990.

[3] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudar-shan. Keyword searching and browsing in databases using banks. In Proc. of International Conference on Data Engineering (ICDE), pages 431440, 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 1629, 2012

[5] X. Cao, G. Cong, and C. S. Jensen. Retrieving top-k prestige-based relevant spatial web objects. PVLDB, 3(1):373384, 2010.

[6] X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi. Collective spatial keyword querying. In Proc. of ACM Management of Data (SIG-MOD), pages 373384, 2011.

[7] B. Chazelle, J. Kilian, R. Rubinfeld, and A. Tal. The bloomier filter: an efficient data structure for static support lookup tables. In Proc. of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 3039, 2004.

[8] Y.Y. Chen, T. Suel, and A. Markowetz. Efficient query processing in geographic web search engines. In Proc. of ACM Management of Data (SIGMOD), pages 277288, 2006.

[9] D. Zhang, Y. M. Chee, A. Mondal, A. K. H. Tung, and M. Kitsure-gawa. Keyword search in spatial databases: Towards searching by document. In Proc. of International Conference on Data Engineering (ICDE), pages 688699, 2009

BIOGRAPHY

Sarita Ahire, Anup Vishwakarma, Bhagwat Chavan, Vinay Verma are a Student of the Computer Department, modern college of engineering, Savitribai Phule Pune University. We are appearing in last year of engineering in 2015-2016.