

# A SURVEY ON LOCATION BASED SEARCH USING SPATIAL INVERTED INDEX METHOD

N.Minojini<sup>1</sup>, D.Daniel<sup>2</sup>, S.Jothi<sup>3</sup>

<sup>1</sup>PG scholar, Department of CSE, Dr. N.G.P Institute of Technology, Coimbatore, India

<sup>2</sup>Assistant Professor, Department of CSE, Dr.N.G.P Institute of Technology, Coimbatore, India

<sup>3</sup>PG scholar, Department of CSE, Dr. N.G.P Institute of Technology, Coimbatore, India

## Abstract

Conventional spatial queries, nearest neighbor retrieval and vary search consists solely conditions on objects geometric property. But today, several fashionable applications support new kind of queries that aim to seek out objects that satisfies each spatial knowledge and their associated text. As an example instead of considering all the hotels, a nearest neighbor queries would instead elicit the building that's nearest to among people who offer services like pool, internet at a similar time. For this a sort of questioning a variant of inverted index is employed that's effective for multidimensional points associate degreed come with an R-tree which is constructed on each inverted list, and uses the algorithm of minimum bounding methodology which will answer the closest neighbor queries with keywords in real time.

**Keywords:** Spatial database, nearest neighbor search, spatial index, keyword search.

\*\*\*

## 1. INTRODUCTION

A spatial information could be an information that stores multidimensional objects like points, rectangles, and etc. some spatial databases permit representing easy geometric objects like lines, points and polygons. Some spatial databases handle additional advanced structures like 3D objects, topological coverage's, linear networks. Based on totally different choice criteria spatial information provides quick access to 3D objects. In spatial information real entities are sculptured in geometric manner, as an example location of hotels, hospital, restaurants are described as points on maps, whereas larger space like landscapes, lakes, parks is described as a mix of rectangles. Spatial information system will utilize in geographic information systems, during this vary search may be utilized to seek out all restaurants in a very sure space, whereas nearest neighbor retrieval will notice the building nearer to a given address. Queries about abstraction information became more and more important to recent years with the increasing quality of some services like Google Earth and Yahoo Maps, as well as different geographic applications. Today, wide used of search engines has created it realistic to put in writing abstraction queries in an exceedingly new means. Historically, queries target objects solely geometric properties, as an example whether or not a point is in parallelogram or however two purposes are shut from every other. Some new application permits users to browse objects based on each of their geometric coordinate and their associated texts. Such sort of queries referred to as abstraction keyword question. For example, if a groundwork engine will be used to find nearest building that provides facilities like pool and internet at identical time. From this question, we have a tendency to might the initial obtain the complete building whose services contain the set of keywords, then notice the closest one from the retrieved

restaurant. The foremost disadvantage of this approach is that, on the troublesome input they are doing not offer real time answer. For example, from the question purpose the important neighbor lies quite far away, whereas all the nearer neighbors square measure missing a minimum of one of the question keywords. Abstraction keyword queries have not been wide explored. Within the past years, the cluster of people has shown interest in learning keyword searches in relational databases. Recently the eye has preoccupied to flat knowledge [5][6]. The simplest technique for nearest neighbor search with keywords is owing to Felipe et al. [5]. They mix the abstraction index R-tree [7] and signature file [8]. So that they developed a structure referred to as IR<sup>2</sup>-tree. This tree has the flexibility of each R-tree and signature files. Like R-tree it stores the abstraction proximity of object and like signature file it filters those objects that don't include all question keywords.

## 2. RELATED WORK

Cao et al. [1] projected collective spatial keyword queries, they conferred the new down side of retrieving a group of spatial objects, and every related to a collection of keywords. They develop approximation algorithms with provable approximations bounds and precise algorithms to solve the two issues.

Lu et al. [2], combined the notion of keyword searches with reverse nearest neighbor queries. They propose a hybrid index tree known as IUR-tree (Intersection-Union R-Tree) to answer the Reverse spatial textual k Nearest Neighbor (RSTkNN) question that effectively combines location proximity with matter similarity. They style a branch-and-bound search rule that relies on the IUR-tree. To more increases the question process, they proposed to associate degree improved variants of the IUR-tree known as cluster IUR-tree and two corresponding improvement rules.

Zhang and Chee [3] introduced hybrid categorization structure BR\*-tree, that mixes the R\*-tree and bitmap indexing to method the m-closest keyword question that returns the spatially nearest objects matching m keywords. They utilized a priority based mostly to search strategy that successfully cut backs the search area and additionally planned two monotone constraints, distance mutex and keyword mutex to help effective pruning.

Ian DE Flipe[4] given in economical technique to answer top-K spatial keyword question. They planned in index structure IR<sup>2</sup>-tree that mixes signature files and R-tree to allow keyword search on spatial knowledge objects that every have restricted range of keywords. Exploitation the IR<sup>2</sup>-tree inefficient progressive rule is given to answer the spatial keyword queries.

G. Cong, C.S. Jensen, and D. Wu dialect [5] planned to associate approaches that computes the relevancy between the documents of associate objects and a question. This relevancy is then incorporated with the geometrician distance between an object and question to calculate associate overall similarity of object to query.

### 3. NEAREST NEIGHBOR SEARCH TECHNIQUE

#### 3.1 IR-Tree, Approximation Algorithmic Rule and Precise Algorithm

This technique is employed to retrieve a bunch of spatial internet objects specified the query's keywords measure cowl by group's keywords and objects area unit around the question location and have very cheap bury object distances. This technique addresses two internal representation of the cluster keyword question. First is searching out the cluster of objects that cowl the keywords such that the add of their distances to the question is minimized. Second is searching out a bunch of objects that cowl the keywords specified add of the most distance among associate object in cluster of objects and question and maximum distance among two objects in cluster of objects is reduced. Each of those sub issues area unit NP-complete. Greedy rule is employed to produce associate approximation solutions to the matter that utilizes the spatial keyword index IR-tree to scale back the search house. However in some application question doesn't contain an oversized range of keywords, for this actual rule is employed that uses the dynamic programming [1]

#### 3.2 IUR-Tree (Intersection union R-Tree):

Geographic objects related to descriptive texts area units becoming common. This provides importance to special keyword queries that take each the situation and text description of content. This system is employed to research the problem of reverse spatial and matter k- nearest neighbor search i.e finding objects that takes the question object in concert of their spatial matter similar objects. For this kind of search hybrid index structure are employed that with success merge the situation proximity with matter similarity.

For searching, the branch and sure algorithmic program is employed. Additionally to increase the speed of question process a variant of IUR- tree and two improvement algorithmic programs is employed. To reinforce the IUR-tree text cluster is employed, during this object of all the information base is cluster into clusters in keeping with their text similarity. Every node of the tree is extended by the cluster information to make a hybrid tree that is termed as cluster IUR-tree. To boost the search performance of this tree two improvements way is employed, an initial relies on outlier detection and extraction and second technique is predicated on text entropy. [2]

#### 3.3 BR\*-Tree:

This hybrid index structure is employed to look m-closest keywords. This method finds the nearest tuples that match the keywords provided by the user. This structure combines the R\*-tree and classification assortment to method the m-closest keyword question that returns the spatially nearest objects matching m keywords to scale back the search house apriority based mostly search strategy is employed. Two monotone constraints are employed as a priority properties to facilitates efficient pruning that is named as distance mutex and keyword mutex. However this approach isn't appropriate for handling ranking queries and during this range of false hits is large.[3]

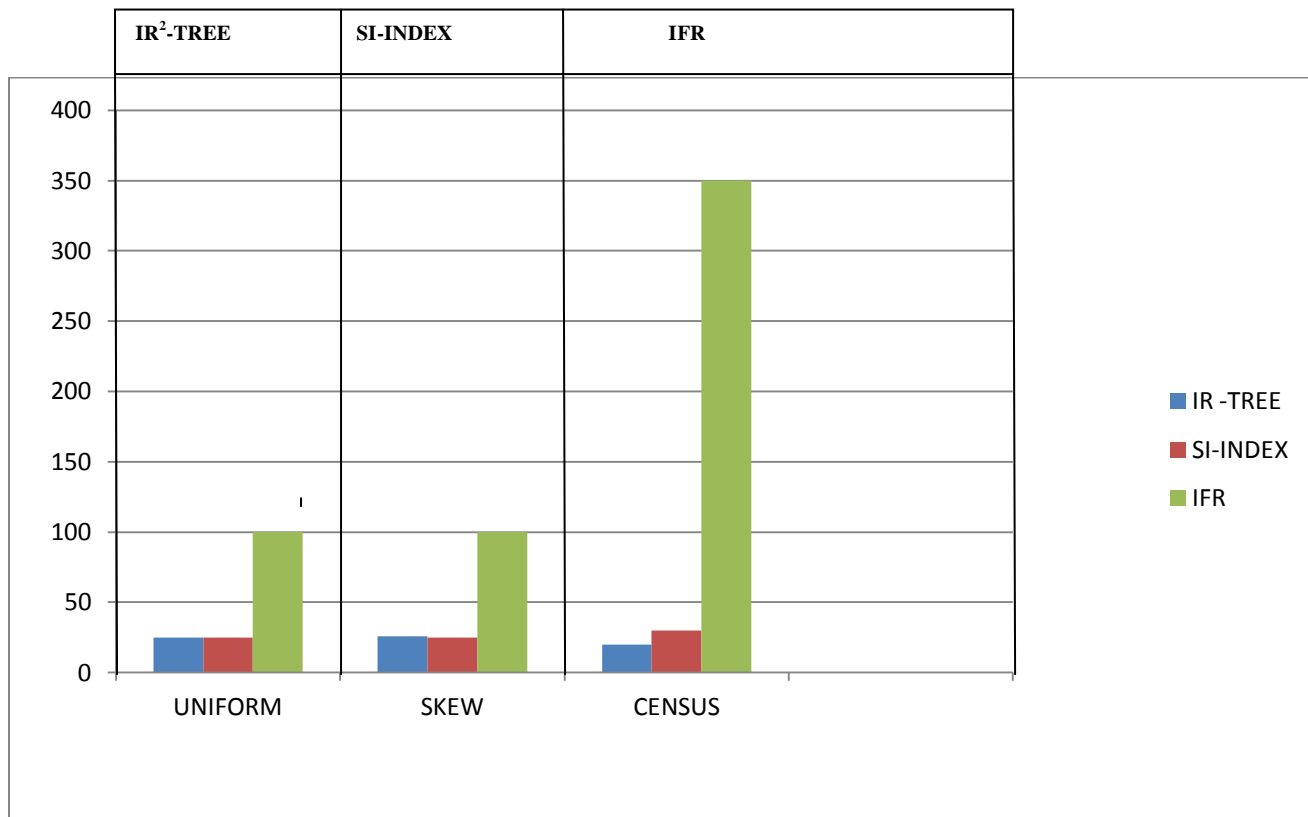
#### 3.4 IR<sup>2</sup>-Tree:

The growing range of applications needs the efficient execution of nearest neighbor queries that is constrained by the properties of spatial objects. Keyword search is extremely common on the net therefore these applications allow users to present list of keywords that spatial objects should contain. Such queries known as a spatial keyword query. This can be consisted of question space and set of keywords. The IR<sup>2</sup>-tree is developed by the mix of R-tree and signature files, wherever every node of tree has spatial and keyword data. This technique with its efficiency provides solution to top-k spatial keyword queries. Also it facilitates the signature is added to each nodes of the trees. Associate in a position rule is employed to answer the queries exploitation the tree. Progressive nearest algorithm is employed for the tree traversal and if root node signature doesn't match the question signature then it prunes the whole sub trees. However IR<sup>2</sup>-tree has some drawbacks such as false hits wherever the thing of ultimate result's isolated from the question or this can be not appropriate for handling ranking queries.[4]

#### 3.5 Spatial Inverted Index and Minimum Bounding Method

So, new access technique spatial inverted access technique is used to get rid of the drawbacks of previous strategies such as false hits. This technique is that the variant of inverted index using for two-dimensional points. This index stores the spatial region of information points and on each inverted list R-tree is built. Minimum bounding technique is employed for traversing the tree to prune the search area.

#### 4. PERFORMANCE ANALYSIS



**Fig 1:** Comparison of performance

The **SI-index** is proposed in query algorithms, and it is an excellent trade-off between space and query efficiency.

Compared to **IFR**, it consumes significantly less space and answers queries much faster. **IFR** occupies large space cost, because it needs to duplicate the coordinates of a data point  $p$ , as many times as the number of distinct words in the text description of  $p$ .

Compared to **IR<sup>2</sup> TREE**, its performance in query time is lower by a factor of orders of magnitude. **IR<sup>2</sup> tree** fails to give the real-time answers. The deficiency of **IR<sup>2</sup> tree** is mainly due to the vast number of false hits.

#### 5. CONCLUSION

This paper presents the survey of varied techniques for nearest neighbor search for spatial information. As in the previous ways, there have been several drawbacks. The present solutions incur too expensive area consumption or they are unable to present real-time answers. Therefore, to beat the drawbacks of previous ways, a new technique relies on a variant of inverted index and R-tree, and a formula of minimum bounding technique is employed to scale back the search space. This technique can increase the potency of nearest neighbor search too.

#### 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, 1999.

## BIOGRAPHIES



**N. MINOJINI** received the B.TECH degree in Computer science and engineering from Karunya University on 2012. She is currently with the Post graduate in Dr.N.G.P IT, now works on the project in Data Mining, and Continues research on location based search using spatial inverted index.



**D. Daniel** received the B.TECH degree in Information Technology from Karunya University in 2009. He received his M.TECH degree on 2011 from Karunya University, he Continues research on adaptive scheduling techniques and multimedia usage in cloud computing.



**S. JOTHI** received the B.E degree in Computer science and engineering from KTVR Knowledge Park of Engineering and Technology on 2013. She is currently with the Post graduate in Dr.N.G.P IT, now works on the project in Networking, and Continues research on overall network performance based on STR in zigbee.