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

An innovative system for Nearest Neighbor Searching Method Using Keywords

N. Rani Reddy
MTech Student
Sridevi Women's Engineering College
V.N.Pally, Near Gandipet, Hyderabad, Telangana

Mrs. N Sujatha
Assistant Professor
Sridevi Women's Engineering College
V.N.Pally, Near Gandipet, Hyderabad, Telangana

Abstract: ***Conservative spatial queries, such as range search and nearest neighbor reclamation, involves only conditions on objects' numerical 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, spaghetti, brandy" all at the same time. Currently the best solution to such queries is based on the Information Retrieval 2-tree, which has a few deficiencies that seriously impact its efficiency. Motivated by this, there is a development of 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 Information Retrieval 2-tree in query response time significantly, often by a factor of orders of magnitude.***

Keywords: ***Information retrieval, spatial index, keyword search.***

1. INTRODUCTION:

Keyword search in document performed with various approaches ranked retrieval results, clustering search results & identifying the nearest neighbor Keyword search on xml document categorized as two different approaches one is Keyword search on xml document which can be performed by ranking the searched results based on match or the answer to keyword & finding the nearest neighbor of keyword by using GST or by Xpath Query. In paper [2] the problem of returning clustered results for keyword search on documents the core of the semantics is the conceptually related relationship between keyword matches, which is based on the conceptual relationship between nodes in trees. Then, we propose a new clustering methodology for search results, which clusters results according to the way they match the given query.

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.

Today, the widespread use of search engines has made it realistic to write spatial queries in a brand new way. Conventionally, queries focus on objects' geometric properties only, such as whether a point is in a rectangle, or how close two points are from each other. We have seen some modern applications that call for the ability to select objects based on both of their geometric coordinates and their associated texts. For

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

example, it would be fairly useful if a search engine can be used to find the nearest restaurant that offers “Chicken 65, Biryani, and brandy” all at the same time. Note that this is not the “globally” nearest restaurant (which would have been returned by a traditional nearest neighbor query), but the nearest restaurant among only those providing all the demanded foods and drinks. 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 far away from the query point, while all the closer neighbors are missing at least one of the query keywords. 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 boom of Internet has given rise to an ever increasing amount of text data associated with multiple dimensions (attributes), for example, customer reviews in shopping websites (e.g., Amazon) are always associated with attributes like price, model, and rate. A traditional OLAP data cube can be naturally extended to summarize and navigate structured data together with unstructured text data. Such a cube model is called text cube [1]. A cell in the text cube aggregates a set of documents/ items with matching attribute values in a subset of dimensions.

Keyword query, one of the most popular and easy-to use ways to retrieve useful information from a collection of plain documents, is being extended to RDBMSs to retrieve information from text-rich attributes [2], [3]. Given a set of keywords, existing methods aim to find relevant items or joins of items (e.g., linked by foreign keys) that contain all or some of the keywords.

Traditional IR techniques can be used to rank documents according to the relevance. In a large text database, however, the number of relevant documents to a query could be large, and a user has to spend much time reading them. If a document is associated with attribute information, in a data cube model (a multidimensional space induced by the attributes), e.g., the text cube, a cell aggregates the documents with matching values in a subset of attributes. Such a collection of documents is associated with each cell, corresponding to an object that can be directly recommended to the user for the given query.

This paper studies the problem of keyword-based top-k search in text cube, i.e., given a keyword query, find the top-k most relevant cells in a text cube. When users want to retrieve information from a text cube using keyword question, believe that relevant cells, rather than relevant documents, are preferred as the answers, because: 1) relevant cells are easy for users to browse; and 2) relevant cells provide users insights about the relationship between the values of relational attributes and the text data.

2. 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 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.

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

Disadvantages of existing System:

- Existing System involves only conditions on objects geometric properties Mainly focus on top-k nearest neighbors, does not find a nearest restaurant for the user given address.
- Low efficient for the incremental query.

3. 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

Advantages of Proposed System:

• An effective access method called Spatial Inverted Index (SI-Index)is proposed to discover the restaurant closest to the user given address
• Spatial database manages multidimensional objects.
• Fast access to the objects in order to retrieve the information to the user

4. PROPOSED SYSTEM ARCHITECTURE

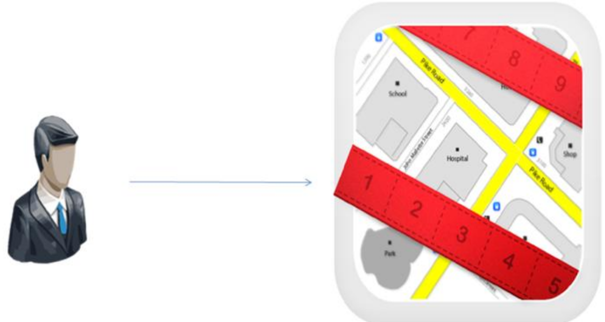

Fig.1: Proposed System Architecture

MODULES:
• Registration
• Login
• Hotel_Registration
• Search Techniques
• Map_view
• Distance_Search

Registration:

In this module a 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.

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:

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

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.

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 (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 milli-seconds. 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 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] Paidekondala Mahesh Goud & B.Rama Ganesh, An Efficient Nearest Neighbor Searching Method Using Keywords, IJMETMR http://www.ijmetmr.com/oljune2015/PaidekondalaMaheshGoud-BRamaGanesh-129.pdf , Volume No: 2 (2015), Issue No: 6 (June)

[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 30–39, 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 277–288, 2006.

[9] E. Chu, A. Baid, X. Chai, A. Doan, and J. Naughton. Combining keyword search and forms for ad hoc querying of databases. In Proc. of ACM Management of Data (SIGMOD), 2009.

[10] G. Cong, C. S. Jensen, and D. Wu. Efficient retrieval of the top-k most relevant spatial web objects. PVLDB, 2(1):337–348, 2009.

[11] C. Faloutsos and S. Christodoulakis. Signature files: An access method for documents and its analytical performance evaluation. ACM Transactions on Information Systems (TOIS), 2(4):267–288, 1984.

[12] I. D. Felipe, V. Hristidis, and N. Rishe. Keyword search on spatial databases. In Proc. of International Conference on Data Engineering (ICDE), pages 656–665, 2008.

[13] R. Hariharan, B. Hore, C. Li, and S. Mehrotra. Processing spatial keyword (SK) queries in geographic information retrieval (GIR) systems. In Proc. of Scientific and Statistical Database Management (SSDBM), 2007.

[14] G. R. Hjaltason and H. Samet. Distance browsing in spatial databases. ACM Transactions on Database Systems (TODS), 24(2):265–318, 1999.

[15] V. Hristidis and Y. Papakonstantinou. Discover: Keyword search in relational databases. In Proc. of Very Large Data Bases (VLDB), pages 670–681, 2002.