IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 11, November 2017

Fast nearest Neighbor Search with Keywords

Miss Netra Katariya[1], Prof. Pawan Kumar[2]

G. H. Raisoni College of Engineering andManagement, Chas, Ahmednagar, India[1]

G. H. Raisoni College of Engineering andManagement, Wagholi, Pune, India[2]

Abstract: Fast nearest Neighbor Search with Keywords involve condition on object's geometric properties.In today's application many queries that aim to seek out out objects satisfying each a abstraction predicate, and a predicate on their associated texts. example of the above condition is consider all restaurants , a nearest neighbor question would instead extract the eating house that's the highest among those whose menus contain steak, spaghetti, brandy all at an equivalent time. in this paper effective resolution to such queries is predicated on the IR2-tree, features a few deficiencies that seriously impact its potency we have to develop a replacement access methodology known as the abstraction inverted index.the purpose of inverted index is to address flat knowledge and come with algorims that may answer nearest neighbor queries with keywords in real time.

Keywords: R-tree, UML, diagrams, brandy, index, R-tree.

I. INTRODUCTION

A spatial database has multidimensional objects (such as points, rectangles, etc.) that will provide faster access to those objects based on different selection criteria.spatial databases has importance that is reflected by the convenience of modeling entities of reality in a geometric manner. The example of above condition is mentioned over here points in a map are represented by locations of restau-rants, hotels, hospitals and combination of rect-angles as parks, lakes, and landscapes . spa-tial database has many many functionalities are useful in various ways in specific con-texts. In a geography information system ,range search is useful in find all restaurants in a certain area.The nearest neighbor retrieval can discover the restaurant closest to a given ad-dress.search engines are used for spatial quries purpose.these quries are focus on objects geo-metric properties only, such as whether a point is in a rectangle, or how close two points are from each other. modern application has some Organized by PCCOE, In Association with BOS(Comp Engg), SPPU,Pune call for the ability to select objects based on both of their geometric coordinates and their associated texts. It would be very useful if any search engine can be used to find the nearest restaurant that offers "brandy,spagheti,steak" all at same time, note that it is not "glob-ally"nearest restaurant (It would be returned by nearest neighbor query),the nearest restaurant only those who are providing foods and drinks spatial inverted index is called SI-index we desing a variant of inverted index that used to optimized for multidimensional points.This SI-index method is used for co-ordinates the points into a conventional inverted index with small extra space,for the purpose of delicate compact storage scheme. SI-index is also useful for preserving the spatial locality of data point.R-tree built on every inverted list as little space overhead.due to this or result of this condition is offers two competing ways for query pro-cessing.In the inverted lists we are merging the mutiple lists by ids.R-trees are used to browse the points of relevant lists in ascending order of their distances to the query point.SI-index are used for outperformed the IR2tree in query effienciency by factor of magnitude.

II. EXISTING SYSTEM

Exploration of the spatial queries have not been explored.In relational database studying keyword search is their.mutidimentional data has divert the attention.finding top-K Nearest Neighbors, where each node has to match the whole querying keywordsdata objects in the spatial space does not consider the density.in the incremental query these method is very low. We have finished explaining how to build the leaf nodes of an R-tree on an inverted list. As each leaf is a block, all the leaves can be stored in a blocked SI-index as described in Section Building the non leaf levels is trivial, because they are invisible to the merging-based query algorithms, and hence, do not need to preserve any common ordering. We are free to apply any of the existing R-tree construction algorithms. It is note worthy that the non leaf levels add only a small amount to the overall space overhead because, in an R-tree, the number of non leaf nodes is by far lower than that of leaf nodes.

III. PROPOSED SYSTEM

A spatial info manages dimensional objects (such as points, rectangles, etc.), and provides quick access to those objects supported totally different choice criteria. The importance of spa-tial databases is mirrored by the convenience of modeling entities of reality in an exceed-ingly geometric manner. for instance, locations of restaurants, hotels, hospitals so on square measure typically described as points in an exceedingly map, whereas larger extents like parks, lakes, and landscapes typically as a mix of rectangles. several functionalities of a spatial info square measure helpful in varied

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 11, November 2017

ways in which in specific contexts. as an example, in an exceedingly geographics system, vary search will be deployed to search out all restaurants in an exceedingly sure space, whereas near-est neighbor retrieval will discover the eating place nearest to a given address. Furthermore, because the SI-index relies on the traditional technology of inverted index, it's without delay incorporable in an exceedingly business com-puter programme that applies large similarity, implying its immediate industrial deserves

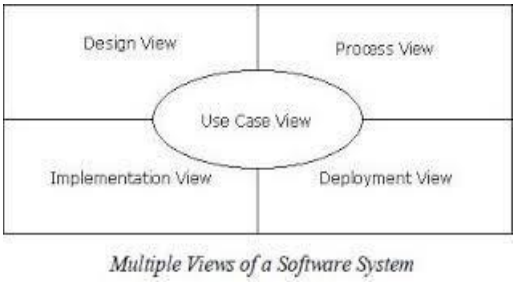

Fig. 1. Modeling a System Architecture using views of UML

IV. DESIGN ANALYSIS

UML Diagrams: UML is a method for describ-ing the system architecture in detail using the blueprint. UML represents a collection of best engineering practices that have proven suc-cessful in the modeling of large and complex systems. UML is a very important part of de-veloping objects oriented software and the soft-ware development process. UML uses mostly graphical notations to express the design of software projects. Using the UML helps project teams communicate, explore potential designs, and validate the architectural design of the software. Definition: UML is a general-purpose visual modeling language that is used to spec-ify, visualize, construct, and document the ar-tifacts of the software system. UML is a lan-guage: It will provide vocabulary and rules for communications and function on conceptual and physical representation. So it is modeling language. UML Specifying: Specifying means building models that are precise, unambiguous and complete. In particular, the UML address the specification of all the important analysis, design and implementation decisions that must be made in developing and displaying a soft-ware intensive system. UML Visualization: The UML includes both graphical and textual repre-sentation. It makes easy to visualize the system and for better understanding. UML Construct-ing: UML models can be directly connected to a variety of programming languages and it is sufficiently expressive and free from any ambi-guity to permit the direct execution of models. UML Documenting: UML provides variety of documents in addition raw executable codes.

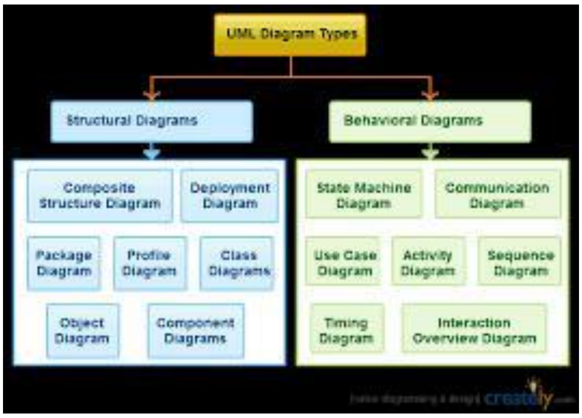

Fig. 2. UML Diagram Types

The use case view of a system encompasses the use cases that describe the behavior of the system as seen by its end users, analysts, and testers. The design view of a system encom-passes the classes, interfaces, and collaborations that form the vocabulary of the problem and its solution. The process view of a system en-compasses the threads and processes that form the system's concurrency and synchronization mechanisms. The implementation view of a system encompasses the components and files that are used to assemble and release the phys-ical system. The deployment view of a system encompasses the nodes that form the system's hardware topology on which the system exe-cutes.

Uses of UML: The UML is intended pri-marily for software intensive systems. It has been used effectively for such domain as En-terprise Information System, Banking and Fi-nancial Services, Telecommunications, Trans-portation, Defense/Aerospace, Retails, Medical Electronics, Scientific Fields, Distributed Web. Building blocks of UML: The

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 11, November 2017

vocabulary of the UML encompasses 3 kinds of building blocks Things Relationships Diagrams Things: Things are the data abstractions that are first class citi-zens in a model. Things are of 4 types Structural Things, Behavioral Things, Grouping Things, An notational Things Relationships: Relation-ships tie the things together. Relationships in the UML are Dependency, Association, Gener-alization and Specialization UML Diagrams: A diagram is the graphical presentation of a set of elements, most often rendered as a connected graph of vertices (things) and arcs (relation-ships). There are two types of diagrams, they are: Structural and Behavioral Diagrams

4.1 Structural Diagrams

Structural Diagrams: The structural diagrams represent the static aspect of the system. These static aspects representthose parts of a dia-gram which forms the main structure and therefore stable.These static parts are represents by classes, interfaces, objects, components and nodes. The fourstructural diagrams are: Class diagram Object diagram Component diagram Deployment diagram

4.2 Behavioral Diagrams

Behavioral Diagrams: Any system can have two aspects, static and dynamic. So a model is considered as complete whenboth the aspects are covered fully. Behavioral diagrams basically capture the dynamic aspect of a system. Dy-namic aspect can befurther described as the changing/moving parts of a system. UML has the following five types of behavioral diagrams: Use case diagram Sequence diagram Collab-oration diagram Statechart diagram Activity diagram

4.3 Use-Case diagram

Use-Case diagram: Use case diagrams are a set of use cases, actors and their relationships. They represent the usecase view of a system.A use case represents a particular func-tionality of a system.So use case diagram is used to describe the relationships among the functionalities and theirinternal/external con-trollers. These controllers are known as actors.

4.4 ClassDiagram

ClassDiagram: Class diagrams are the most common diagrams used in UML. Class diagram consists of classes, interfaces, associations and collaboration. Class diagrams basically repre-sent the object oriented view of a system which is static in nature. Active class is used in a class diagram to represent the concurrency of the system. Class diagram represents the object ori-entation of a system. So it is generally used for development purpose. This is the most widely used diagram at the time of system construction.

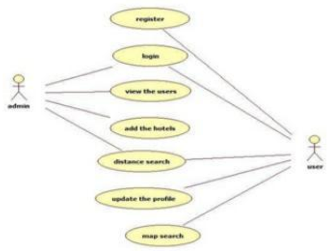

Fig. 3. .UML Class Diagram with Relationships

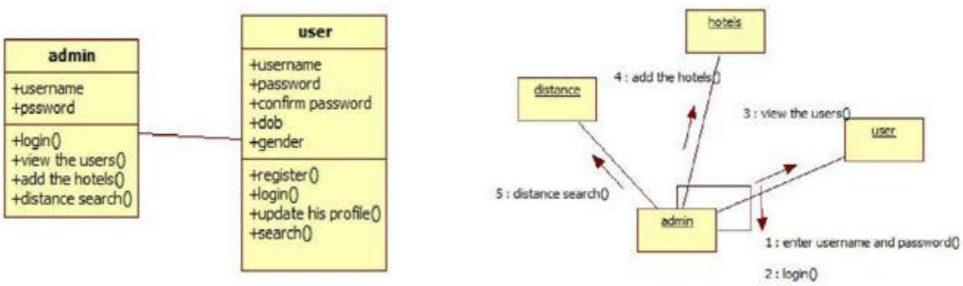

Fig. 4. .UML Class Diagram with Relationships

4.5 Sequence diagram

Sequence diagram: The processes are repre-sented vertically and interactions are show as arrows. This article explains the purpose and the basics of Sequence diagrams.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 11, November 2017

4.6 Collaboration diagram

Collaboration diagram: Communication dia-gram was called collaboration diagram in UML

1. It is similar to sequence diagrams but the focus is on messages passed between objects. The same information can be represented using a sequence diagram and different objects. State machine diagrams: State machine diagrams aresimilar to activity diagrams although notations and usage changes a bit. They are sometime known as state diagrams or start chart dia-grams as well. These are very useful to describe the behavior of objects that act different accord-ing to the state they are at the moment. Below State machine diagram show the basic states and actions.

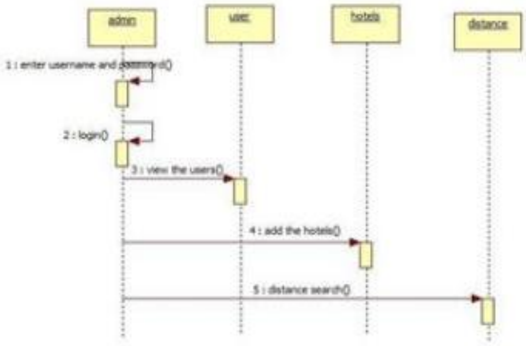

Fig. 5. .. Basic Sequence Diagram

4.7 Activity diagram

Activity diagram: Activity diagrams describe the workflow behavior of a system. Activity diagrams are similar to state diagrams because activities are the state of doing something. The diagrams describe the state of activities by showing the sequence of activities performed. Activity diagrams can show activities that are conditional or parallel. Activity diagrams show the flow of activities through the system. Dia-grams are read from top to bottom and have branches and forks to describe conditions and parallel activities. A fork is used when multiple activities are occurring at the same time. The diagram below shows a fork after activity1.

V. SCREENSHOTS

In the screenshots they mentioned the mod-ules.The following figer mentioned that how the modules are proceed

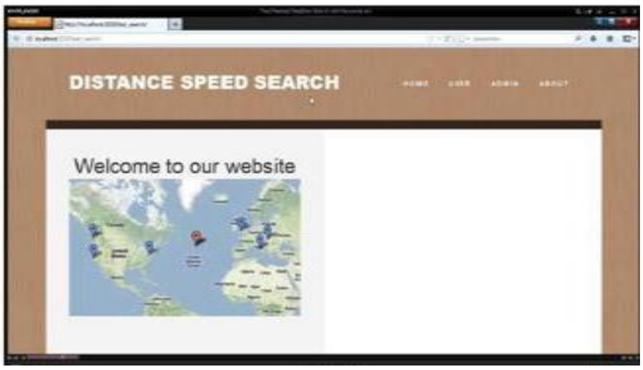

Fig. 6. ..

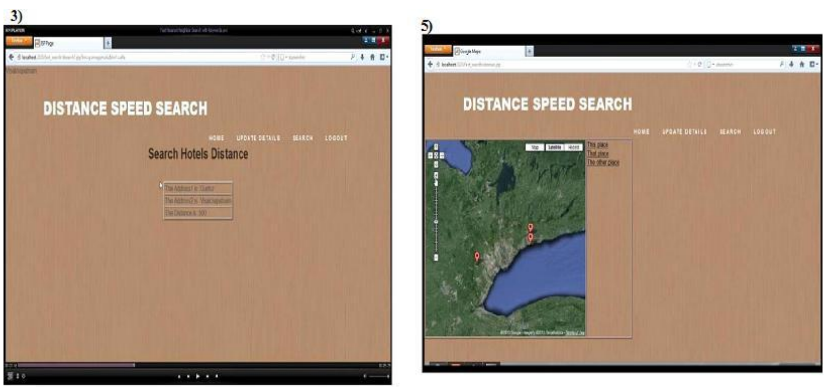

Fig. 7. ..

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 11, November 2017

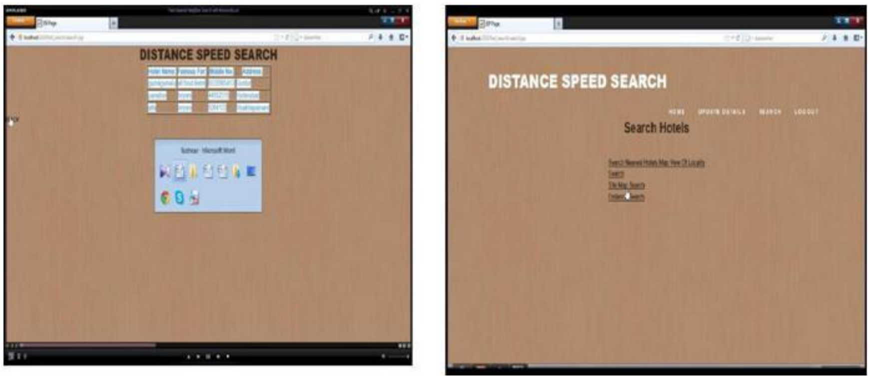

Fig. 8. ..

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 ex-isting 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 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.

ACKNOWLEDGMENT

I would like to thank my guide **Prof. Pawan kumar** for his personnel involvement, technical guidance, valuable suggestions and continuous consolation throughout the work. Also thank to all authors of papers referred in this paper for their work. It was knowledge gaining and interesting process with best of outcomes.

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 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 ngineering (ICDE), pages 431440, 2002.
[4] X. Cao, L. Chen, G. Cong, C. S. Jensen, Q. Qu, A. Skovs-gaard, 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):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.

BIOGRAPHIES

Miss Netra Dilip Katariya, PG Scholar Computer Engineering, G. H. Raisoni College of Engineering and Management, Savitribai Phule Pune University,Chas, Ahmednagar.414001, India.

Prof. Pawan Kumar, Assistant Professor, Computer Engineering, G.H. Raisoni College of Engineering and Management, Savitribai Phule Pune University, Wagholi, Pune 411015, India.

 DOI 10.17148/IJARCCE.2017.61120