An International Journal of Advanced Computer Technology

Designing of Semantic Nearest Neighbor Search

Pawar Anita R.[1], Mulani Tabssum H.[2], Pansare Rajashree B.[3], Bandgar Shrimant B.[4]

[1] Student BE Computer, S.B. Patil College of Engineering, Pune , Maharashtra India

[2] Student BE Computer,S.B. Patil College of Engineering,Pune , Maharashtra, India

[3] Student BE Computer,S.B. Patil College of Engineering,Pune , Maharashtra India

[4] Asst. Professor,S.B. Patil College of Engineering,Pune , Maharashtra India

Abstract: Finding out nearest neighbor is successfully implemented with distance as a base for comparison of who is nearest. But in real life we always not satisfied for getting only nearest by distance. Along with shortest distance we may have some goal in our mind for search such as find nearest Medical store in which diabetes medicines are available. In a geography information system, range search can be deployed to find all medical stores in a certain area, while nearest neighbor retrieval can discover the Medical store closest to a given address. Nearest neighbor search useful for any application but in this paper we will implement only medical application, useful for finding nearest medical along with tablet information .the medical is not globally near but medical is nearest along with all requirement of the user that means tablet. User have the authenticated user that means user already registered its account on server is necessary. Medical administrator is 2nd party to modify the backend database user friendly manner. it has the right of modification to all information stored in database server.

Keywords: spatial Index, K-mean, keyword-based Apriori item-set, neighbour search

I. INTRODUCTION

This project aims at development of a Semantic Nearest Neighbor Search Utility which facilitates the customers to manage their tablet information, and the Medical administrators to modify databases User-Friendly manner.

There are effort free way to support question that combine spatial & text features for example we could first fetch all medical whose contain the set of keyword (tablet name)& then from previous medical find nearest one. Similarly one could also do it reversely by targeting first the spatial condition –brows all medical in ascending order (by distance) to the query point until finding one whose menu has all the keywords.

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 Felipe et al. They nicely combine two known concepts: *R-tree*, 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 remains as it is 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.

An SI index remains as it is the spatial locality of data point and comes with R tree built on every inverted list at little space, by ids we merging traditional inverted list same as we can sequentially merge multiple lists. we can also control the R tree to brows the point of all important list in ascending order by distance to query point or address.

The spatial index it is basic concept to provide the data in a sequential manner .The spatial index provides only validated data as per user requirement. The spatial index which makes combinations of tablet regarding information. The spatial index having provides searching facilities. For example the user wants this query regarding information that not be sufficient knowledge in every user, so provides the spatial index that index whole combination just like alphabetical manner what exactly want that user as per choice, that is related and after that tablet oriented data should be provided. The spatial index to provide the tablet and medical related information.

IR2 tree stands for inverted index retrieval tree. The IR2 tree maintained relation in hierarchical manner. That keeps the wanted data. They have provided as possible as immediate location wanted data. The example show the IR2 tree concept suppose we wants the upper as well as middle level data, but user wants don't know actual way of searching, hence only name insert targeted location and search then that search not go to directed on exact location.

Keyword-based nearest neighbour search is concept based on keyword. What exactly user search data, whole keywords search based on the query. The key nearest neighbour search its depends upon the user query. What user search data in medical database. The keyword search it's provided types of algorithms that help to finding the nearest search element.

The keyword search finding whole possible combination at what exactly search, but at that time result providing time not all the result sent only from that location nearest location and location regarding data should be sent. This concept should be more useful to this project.

II. PROPOSED SYSTEM

To introduce an effective third party medical administrator or knowledge of the user's data. Thus our system is for finding or searching. Our treatment of nearest neighbor search falls in the general topic of spatial keyword search, which has also given rise to several alternative problems. A complete survey of all those problems goes beyond the scope of this project. 1. Strictly speaking, this is not precisely true because merging may need to jump across different lists; however, random I/Os will account for only a small fraction of the total overhead as long as a proper prefetching strategy is employed.

System Architecture

Our proposed system consists of 7 phases : the defining of architecture to provided data to helping of valid user. It provide the medical store database, way of providing the medical data as per requirements. During First phase user fire the query on server side then server check user already registered or not if registered then check authentication. in second phase merging inverted list accept the query, if user is authorize it go to third phase medical store database. after that server check the information or data which is request by the user. Then fourth phase Data point inverted index server check all tablet related information like price, expiry date etc. and go to fifth phase spatial inverted index in that collect all information that means create number of objects.

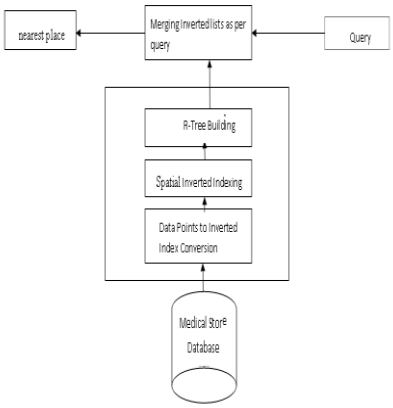

Fig.Flow Of Project

Fig.1 For Ex, reading 10 sequential pages at a time.

R-Tree building is sixth phase define the relation tree, with the help of that firstly server search number of places and then find the nearest place. This all information merging in merging inverted list and send the nearest path with information to user in nearest search which is last phase.

III. DESIGN ANALYSIS

UML diagrams

UML stands for Unified Modeling Language. Like other languages UML is also a langu**age which** provides a vocabu**lary** and the rules for combining words in the vocabulary for the purpose of communication . a modeling language is a language whose vocabulary and rules focus on the conceptual and physical representation of a system. A UML is thus a standard language for software blueprint. With UML we can represent any system with words and pictures. UML is used for visualization or imagination of the final product, specification of the requirements, constructing the artifacts, and documenting.

Building blocks of UML:

The vocabulary of the UML encompasses 3 kinds of building blocks

a. Things
b. Relationships
c. Diagrams

a. Things:

Things are the data abstractions that are first class citizens in a model. Things are of 4 types Structural

Things, Behavioral Things, Grouping Things, An notational Things

Relationships: Relationships tie the things together. Relationships in the UML are Dependency, Association, Generalization 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 (relationships).You draw diagrams to visualize a system from different perspectives, so a diagram is a projection into a system. Here we design UML diagrams for "Semantic Nearest Neighbor Search" which is for medical application
The UML diagrams as follows:

Class Diagram
The class diagram shows the classes that represent the concept. in this diagram class is shows as a solid outline rectangle with any by horizontal liner. A class diagram is a diagram that shows a set of classes ,interfaces and collaborations and their relationships. graphically ,a class diagram is a collection of vertices and arcs.

Component of class diagram
Classes, objects
Interfaces
Collaborations
Relationships like dependency, generalization and association

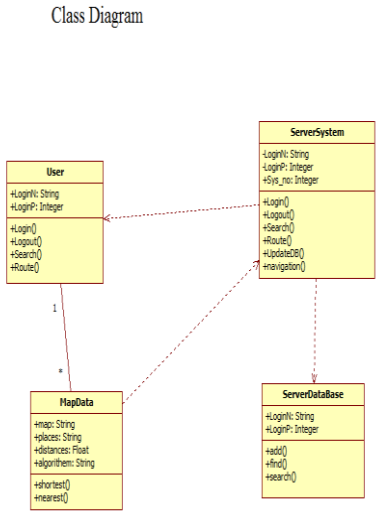

Fig.2: Class Diagram

Use case Diagram
A use case diagram establishment the capability system as a whole.

Component of use case diagram
1. Actor
2. Use cases
3. System boundary
4. Relationships

An actor is someone or something that must interact with the system under development. use case is pattern of behavior. Use case are useful in capturing and communicating functional requirements.

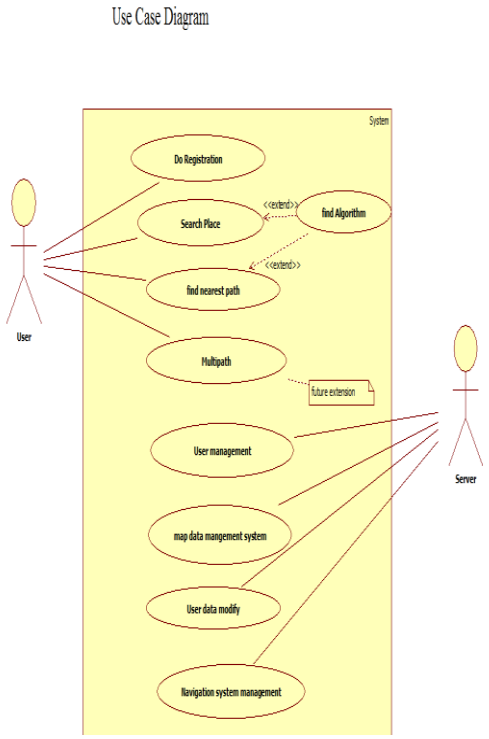

Fig.3: Use case Diagram

3)

4) Activity Diagram
Activity diagram in an UML diagram is used to describe dynamic aspect of the system. As well as to show message flow of control from one activity to another activity. Activity diagrams should be used in conjunction with other modeling techniques such as interaction diagrams and state diagrams. The main reason to use activity diagrams is to model the workflow behind the system being designed. Activity Diagrams are also useful for: analyzing a use case by describing what actions need to take place and when they should occur; describing a complicated sequential algorithm; and modeling applications with parallel processes.

Component of Activity diagram
1. Activity state and action states
2. Transition
3. Objects

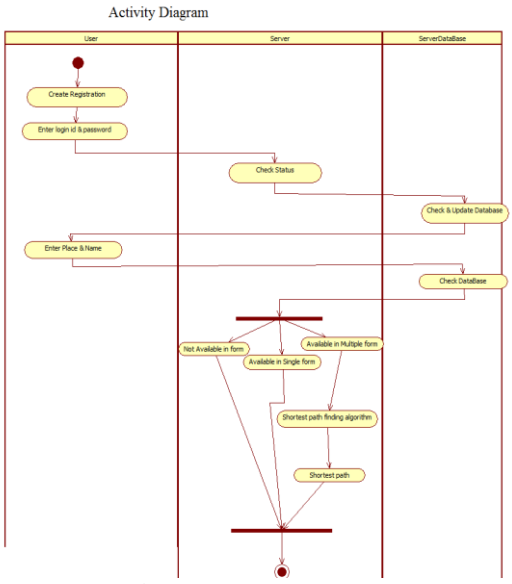

Fig.4: Activity Diagram

Sequence Diagram

The sequence diagram is one of the behavior diagrams that model the dynamic aspect of a system. The series of sequential steps over the time to represent the behavior is done by a sequence diagram. A sequence diagram is used to show the work flow and the messages exchanged among objects or among the elements.

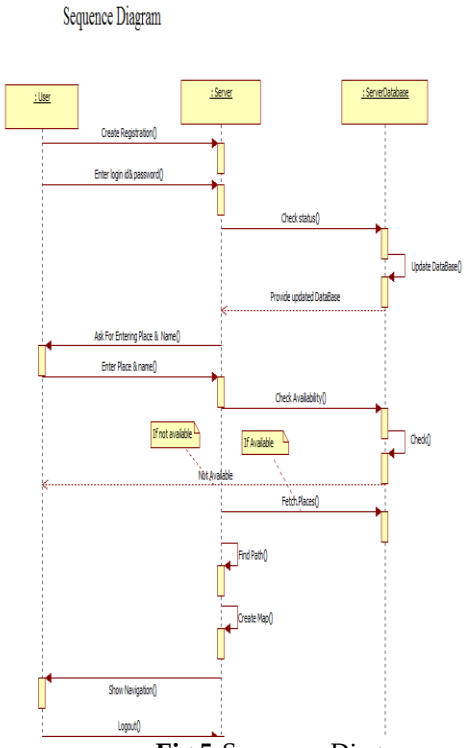

Fig.5: Sequence Diagram

5) Collaboration Diagram

A collaboration diagram is a UML interaction diagram which is describe a particular goal by representing set of chronological interaction between objects.

Component of collaboration diagram

1.Objects
2. Links
3. Messages

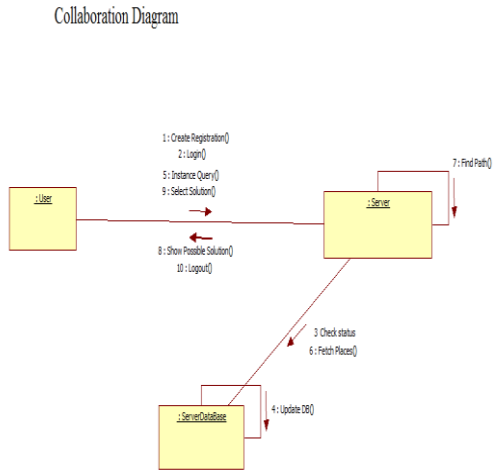

Fig.6: Collaboration Diagram

6) Deployment Diagram

A deployment diagram shows the physical relationships among software and hardware components in the derived system. each node of deployment diagram represent some kind of computational unit in most cases in piece of hardware. Container among nodes shows the communication path over which system will interact.
Components of deployment diagrams

1. Nodes
2. Dependency and association relationships

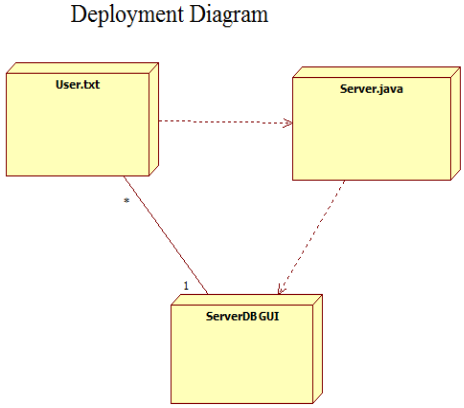

Fig.7: Deployment Diagram

IV. COMPARISON BETWEEN EXISTING SYSTEM AND PROPOSED SYSTEM

Existing System

Since this system provides the services where to put location and information about the product. Existing system is related to only one place for example Shopping center. In one shopping center number of floor then for the customer this system provides the shortest path from he or she stand. From this system user knows the shortest path only area of shopping center also can't find the information.

Demerits of Existing System

Although this application provides within minimum time as per user requirement but provide facility in limited area that means only one location, Also some restriction on their limited area. Medical database provides wanted information as well as regarding information with location where existing system provides application related information but not every information as well as location.

V. CONCLUSION

This paper describe, A user-set minimum support decides about which rules have high support .Once the rules are selected, they are all treated the same, irrespective of how high or how low their support. Their locations are uniformly distributed in Uniform, whereas in Skew, they follow the Zip f distribution. In this system we are proposing a way for providing all details of the tablet along with distance to the user. Comparison with existing system shows the efficiency of this system. We design this system which will be helpful for implementation of SNNS system. So, In this system of Semantic Nearest Neighbor Search Utility which facilitates the customers to manage their tablet information, and the Medical administrators to modify the backend databases in a User-Friendly manner.

VI. 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 DataEngin-eering (ICDE), pages 516, 2002.

[2] N.Beckmann, H.Kriegel,R.Schneider, and B.Seeger.The R*-tree:Ane_cient and robust access method for points and rectangles.InProcof ACM Management of Data(SIGMOD), pages 322331, 1990.

[3] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan. Keyword searching and browsing in databases using banksInProc.of International Conf-erence 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, (1):373384, 2010.

[6] Y .-Y .Chen,T.Suel, and A. Markowetz.E_cient query processing geographic web search engines. In Proc.of ACM Management of Data(SIGMOD),pages277288, 2006

[7] G .Cong, C. S .Jensen, and D. Wu. E_cient retrieval of the top-k most relevant spatial web objects. PVLDB, (1):337348, 2009.

[8] C .Faloutsos and S .Christodoulakis Signature _les:An access method for docu-6mentsand its analytical performance evaluation.ACM Trans-actions on Infor-mation Systems (TOIS),2(4):267288, 1984.

[9] I.D .Felipe, V.Hristidis, and N.Rishe. Keyword search on spatial databases.In Proc.of International Conference on Data Engineering(ICDE) pages 656665, 2008

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

[11] X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi .Collective spatial keyword quer-ying.In Proc. of ACM Management of Data (SIGMOD),pages 373384, 2011.

[12] I.Kamel and C.Faloutsos. Hilbert R-tree : An improved r-tree using fractals. In Proc. of Very Large Data Bases(VLDB), pages 500509, 1994.

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