International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 27799-27811

http://www.ripublication.com

Searching the Nearest Neighbour with a keyword based on Max-Priority

ANUSHA VAKA
P.G Scholar,
Dept of CSE,
SVEC,Tirupati,
Mail id:anushavaka92@gmail.com

Mr. SHAIK SALAM,M.E(Ph.D)
Associate professor
Dept of CSE,
SVEC,Tirupati,
Mail id: shaiksalam17@gmail.com

Abstract –
The Nearest Neighbor (NN) are searched using a keyword but initially they used R-trees, IR^2 trees and conventional inverted index for finding neighbors which do not show exact real neighbors , so with the help of key word the nearest neighbors can be found real time in difficult inputs .For example instead of considering all the Restaurants a query can be raised that which Restaurants are closest among those whose menu contain (soup pizza & coke) then the Nearest Neighbors are found based on the Conventional Spatial inverted index. Extending it with a new access method Ball tree Nearest Neighbor Algorithm which can answer NN queries with max-priority in real time .
Keywords- Conventional, spatial ,inverted index ,Nearest Neighbor, ball tree ,max-priority

INTRODUCTION

Back Ground

Data is a situated of estimations of subjective or quantitative variables. Information is regularly the aftereffect of estimation and can be pictured utilizing charts or pictures.

Database is a sorted out gathering of information the formally "database" alludes to the information themselves and supporting information structures.

Data Mining is separating legitimate data from enormous data another effective it is an innovation with awesome potential to help organizations concentrate on the most critical data in their information stockrooms. It has been characterized as the quick examination of expansive or complex information sets keeping in mind the end goal to find noteworthy examples or patterns that would somehow or another go unrecognized

SPATIAL DATABASE

The spatial database stores huge measure of space related information spatial data types :they can capture the fundamental abstraction for line, point and region like locations and represents the data in the form of points or in multidimensional way and the access is very speed Spatial Database is otherwise called geospatial or geographic data it is the information or data that distinguishes the topographical area of highlights on earth spatial information

For the Nearest Neighbor search a spatial query rises in which we need to find out the real nearest neighbors Conventional spatial queries, range search and nearest neighbor retrieval involve only conditions on objects The importance of spatial database is reflected by the convenience of modeling entities of reality in a geometric manner For example location of restaurants , hotels, hospitals and So on are represented as points in a map The Spatial queries such as range search and nearest neighbor involve only conditions on objects geometric

properties The Nearest Neighbor (NN) are searched using a keyword For example instead of considering all the Restaurants a query can be raised that which Restaurants are closest among those whose menu contain the keywords(soup pizza &coke) by using Conventional Spatial Inverted Index Further extending it with a new access method the Ball tree Nearest Neighbor Algorithm that can answer NN queries with max-priority in real time

LITERATURE SURVEY

"A."Discovering queries that are closest to a given area that contains a group of keywords".

For the keyword search there are number of application present in the internet that provide user some wide list of keywords which contain spatial objects in their attributes For example, land business websites enable the users to seek the property with some unique keywords and search the distance and giving then specific ranking Tend to decision such queries spatial keyword queries a spatial query contain space query and a group of keywords the solution can be finding out the distance and then the distance first spatial keyword query the object square measure hierarchical for distance and keyword square measure for the conjunctive filter that eliminates the objects that are not present .current systems use ad-hoc combos of nearest neighbor (NN) and keyword search techniques to tackle the matter. For an example, associate points R-Tree is employed to seek out the closest neighbors associate points for every neighbor an inverted index is employed to envision if the query keywords area unit contained.

B." Searching the Nearest Neighbor utilizing NN seek"

The closest neighbor could be a prophetical strategy suitable for order models. Not care for option prophetical calculations, the training data is not filtered or transformed to shape the model. Rather, the guiding data is that the model. When a fresh out of the box new case or occurrence is presented to the model, the algorithmic tenet appearance in the scarcest focuses the data to pursuit out an arrangement of cases that are most just about like it and uses them to anticipate the outcome. There are two main drivers inside the calculation: the amount of closest cases to be utilized (k) and a metric to experience what's implied by closest. Each utilization of the algorithm needs that we have a tendency to determine a positive number worth for Nearest Neighbors.

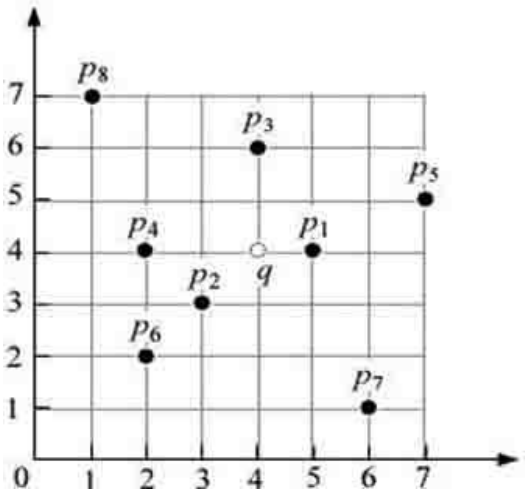


fig (a)points of the location

p	W_p
p_1	$\{a, b\}$
p_2	$\{b, d\}$
p_3	$\{d\}$
p_4	$\{a, e\}$
p_5	$\{c, e\}$
p_6	$\{c, d, e\}$
p_7	$\{b, e\}$
p_8	$\{c, d\}$

fig (b)texts associated

C. "A System for Keyword-Based Search over spatial indices"

Spatial indices are utilized by spatial (databases which store data identified with questions in space) to advance spatial inquiries. Conventional index don't proficiently handle spatial indices, for example, how far two points contrast, or whether points fall within a spatial area

- It is utilized for making files on the grounds that there is tremendous measure of information need to be put away for looking that information put away as xml documents.
- If the data put away made as indices then space needed is less likewise time required for keyword search is comparatively less

D."The R- tree An Efficient and Robust Access Method for Points and Rectangles"

R-Tree which are which represent the closest node that is the root and the tree which consists of the children where every each new node that has entered it shows the highest preference that of all makes utilization of singularly Associate in Nursing R-Tree association .

MOTIVATIONS

For searching nearest neighbors the main motivation is from the spatial inverted index where it can find only the nearest but using Ball tree nearest neighbor algorithm we can find the nearest neighbors who are having highest priority and this can give the exact priority based nearest neighbors with the key word

PROBLEM DOMAIN

The issue is that while looking on an inquiry for the Nearest Neighbors utilizing the spatial database which need to speak to the information as focuses which really don't demonstrate the genuine neighbors as focuses in the guide it just can demonstrate the redesigned areas just like the from the point give is that if the closest neighbor lies a long way from the question point and missing no less than one inquiry watchword For discovering protests in space it neglects to give ongoing replies on troublesome inputs It don't show max-need focuses which are the real continuous Nearest Neighbors

PROBLEM DEFINITION

For discovering protests in space it neglects to give ongoing replies on troublesome inputs It don't show max-need focuses which are the genuine continuous Nearest Neighbors Preposterous space ingesting or unfit to give ongoing answers direct methodologies is that they will neglect to give continuous replies on troublesome inputs. is that the genuine closest neighbor lies far-far from the question point, while all the closer neighbors are absent no less than one of the inquiry pivotal words so continuous nearest neighbors are not found.

PROBLEM STATEMENT

For finding objects in space it fails to provide real time answers on difficult inputs It do not show max-priority points which are the actual real time closest Nearest Neighbors

PROBLEM FORMULATION

The queries for finding objects in space are left with space ingestion or incapable to provide real time answers is that if the Nearest Neighbour lies quite far from the query point and missing at least one query keyword This problem can be resolved by using Ball tree Nearest Neighbor algorithm which is a space partitioning data structure for organizing points in a multi-dimensional space which provides the real closest neighbors with max-priority

ASSUMPTIONS

- The max-priority points which are the actual real time closest nearest neighbors
- Ball –tree Nearest Neighbor Algorithm should show the real closest neighbor with the Max-priority of real closest neighbours in real time

MATHEMATICAL MODEL

$$Qp = \{ p \in P \mid Qw \subseteq Pw \}$$

P:point s
Q:query
QP: query point
Qw: set of keywords
Pw: set of points
p € P: associated with a set of words
p € P Pw and termed the document of p

VERIFICATION AND VALIDATION

STEP 1: First start the process for point search p

STEP 2: $_{Qw}$ set of keywords P{A,B,C,D}

IF searching for point {A,B} input Attributes {A,B,}

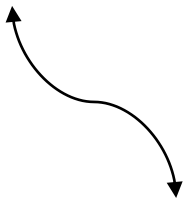
STEP 3:

p € P associated with a set of words

p1={A,A}

p2={B,D}

p3={A,B}



p4={A,B} set of keywords Qw

STEP 4:

Pw and termed the document of P

Qp = matched pair is {p3,p4}

The n P={A,B}Keyword match found it is the point p for Qp

STEP 5:

$$Qp = \{ p \in P \mid Qw \subseteq Pw \}$$

hence the Nearest Neighbor are found with max priority

Frame Work

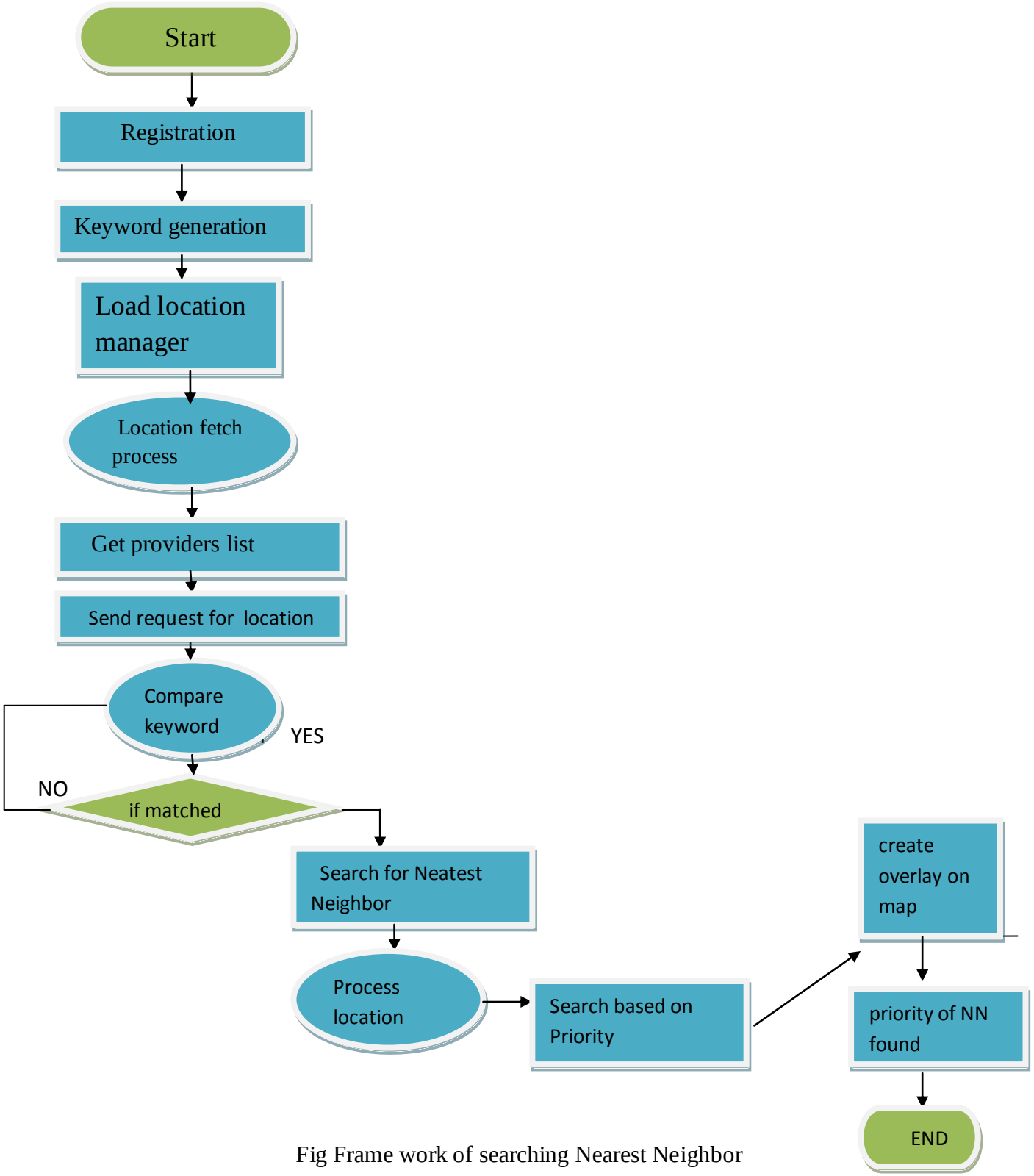


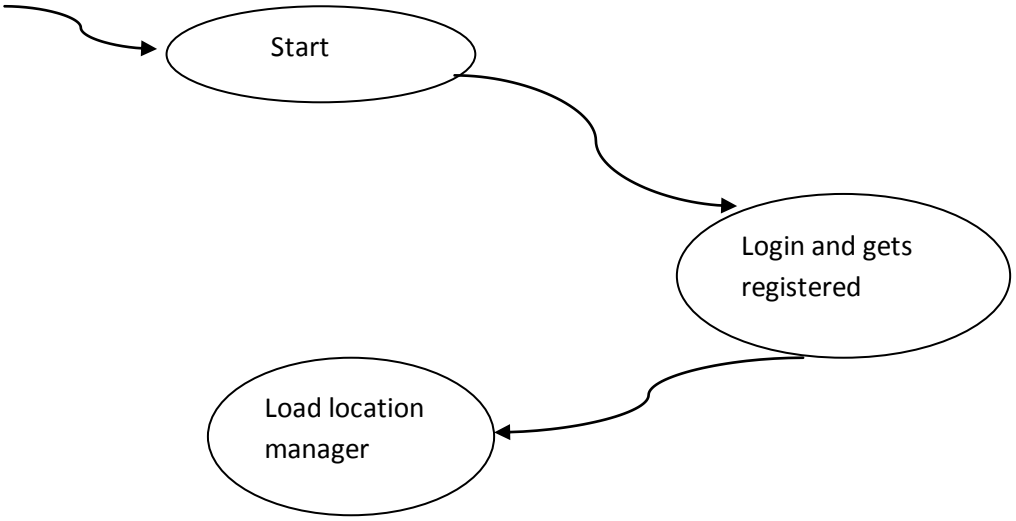
Fig Frame work of searching Nearest Neighbor

EXPERIMENTAL DESIGN

- **Phase 1:**

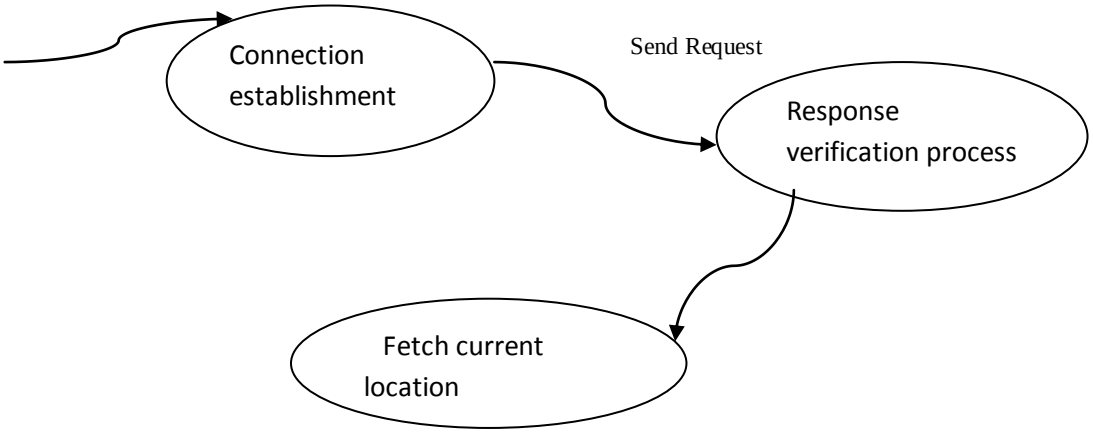
  starts the process login and gets registered and then it loads the location manager

- Search for nearest point based on keyword



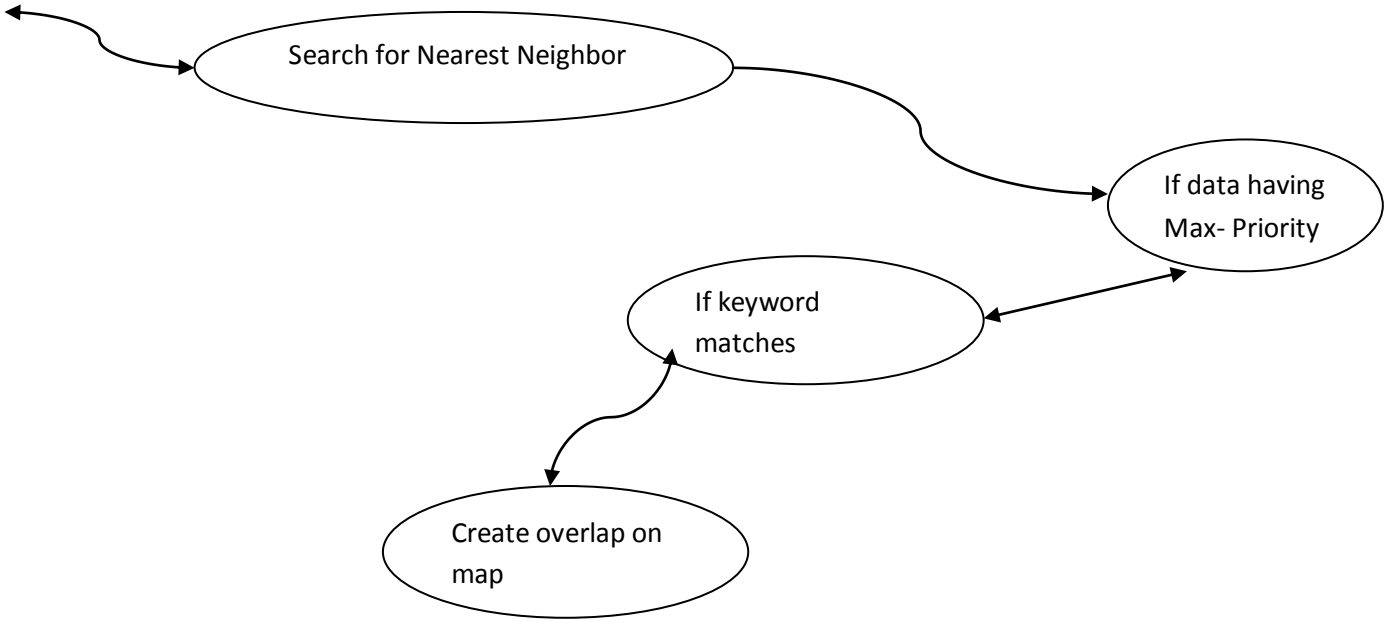
- **Phase 2:**

  User will serve as input to the next phase user sends the request and then receive the response.



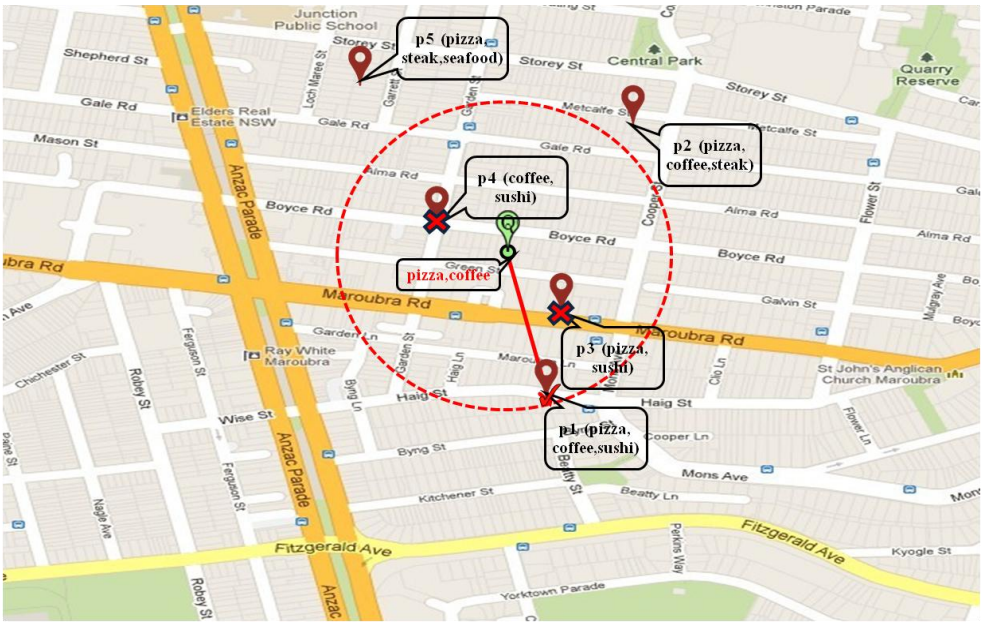
- **Phase 3:**

This is the most important phase of the system. Initially it waits for the request from and search nearest neighbor then converts the points into one dimension.



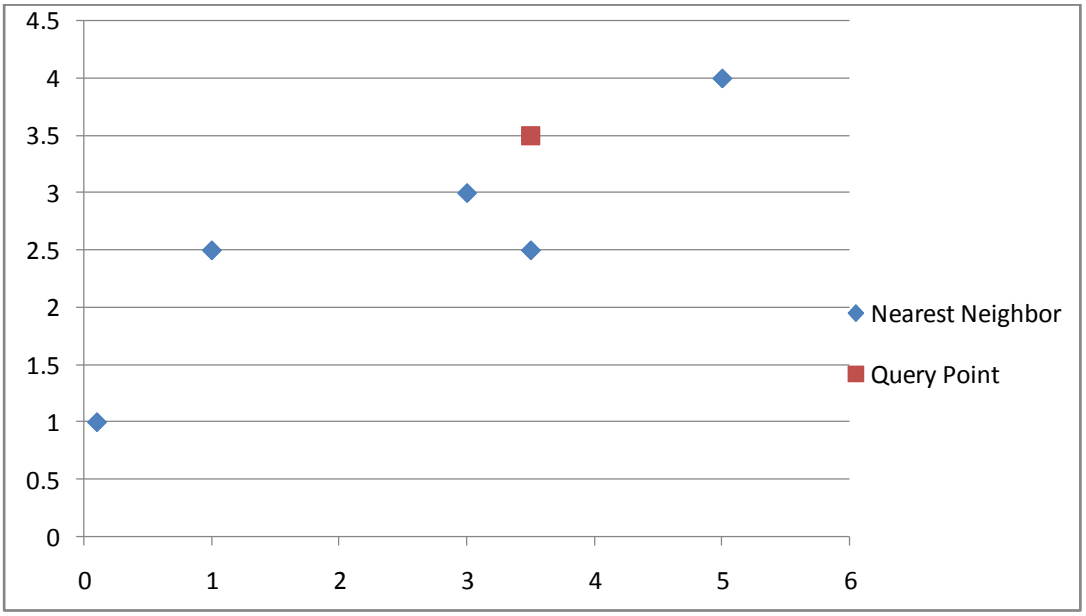
- **Phase 4**

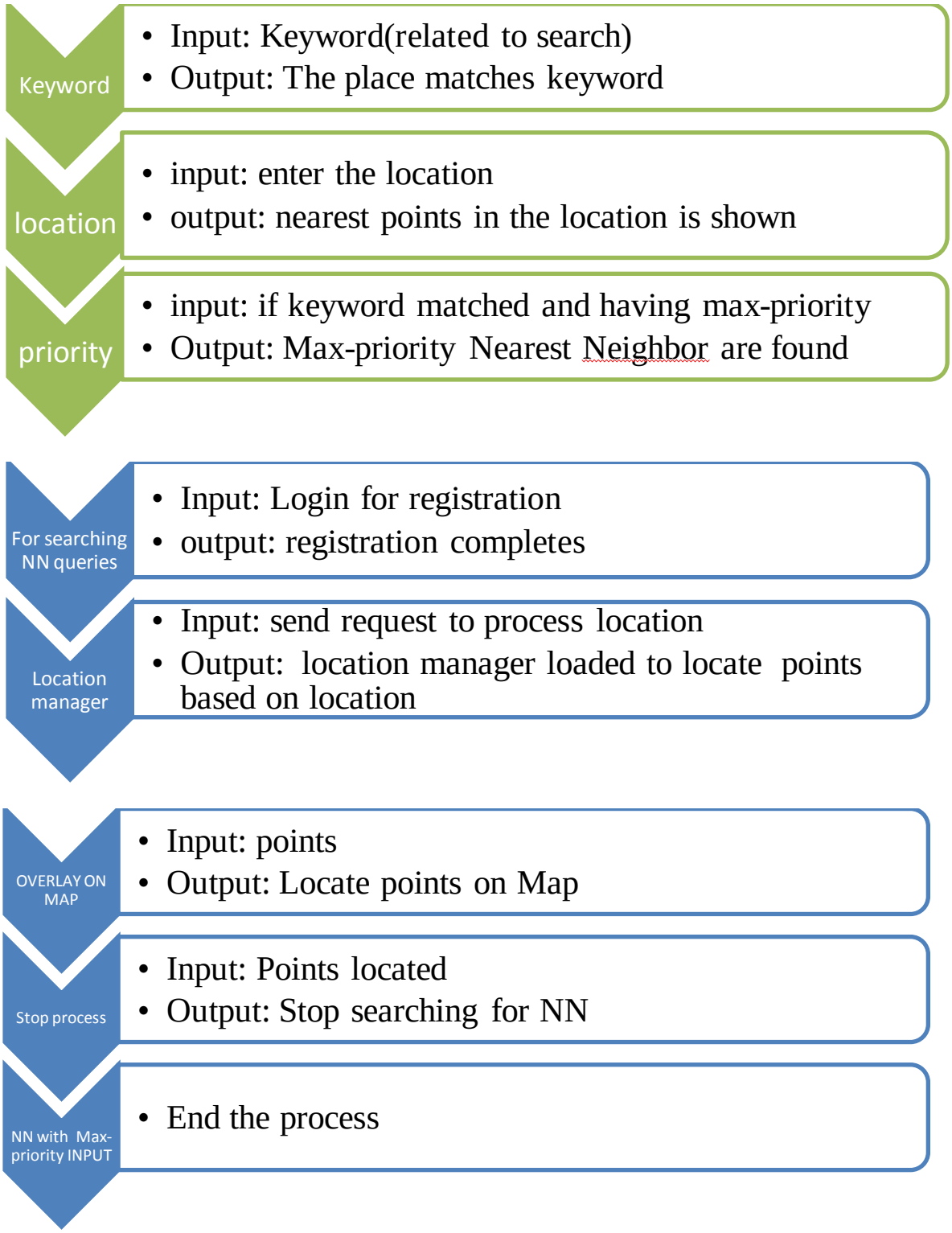
This is the phase of the system. It waits for the data to be displayed then initiates if the data is found and then key word matches it creates overlay on the map and Ball tree nearest neighbor algorithm for max- priority of closest nearest neighbor if the NN priority point is found then the search is stopped and points are plotted on map



An example for searching the Nearest Neighbor with Max-priority

RESULTS AND SENSITIVITY ANALYSIS




Keyword
• Input: Keyword(related to search)
• Output: The place matches keyword
location
• input: enter the location
• output: nearest points in the location is shown
priority
• input: if keyword matched and having max-priority
• Output: Max-priority Nearest Neighbor are found
For searching NN queries
• Input: Login for registration
• output: registration completes
Location manager
• Input: send request to process location
• Output: location manager loaded to locate points based on location
OVERLAY ON MAP
• Input: points
• Output: Locate points on Map
Stop process
• Input: Points located
• Output: Stop searching for NN
NN with Max-priority INPUT
• End the process

Comparison of Result

x-axis	Y axis	Average time for searching NN with Ball tree	Average time for searching NN with IR^2 tree
0	Ball tree	5ms	22ms
10	Ball tree	10ms	30ms
20	$1R^2$ tree	12ms	35ms
30	IR^2 tree	15ms	49ms

Table 1 Searching Time for NN

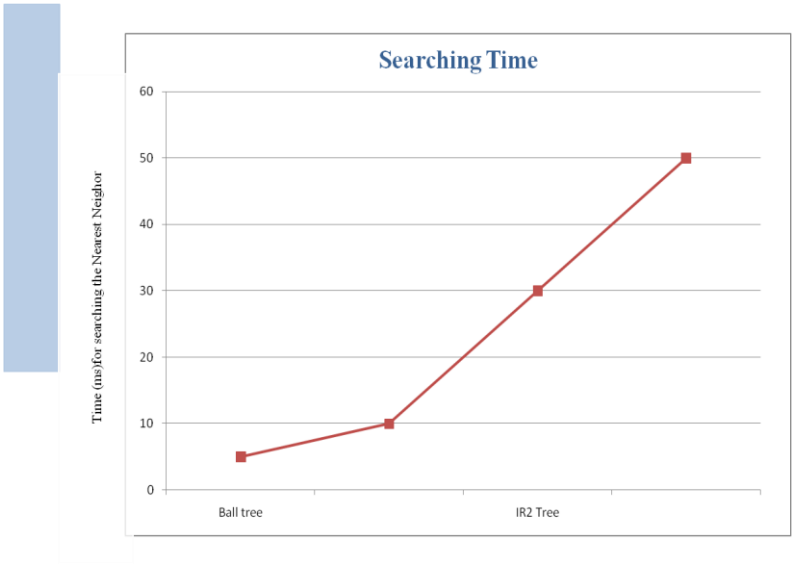


Fig Time in searching the NN

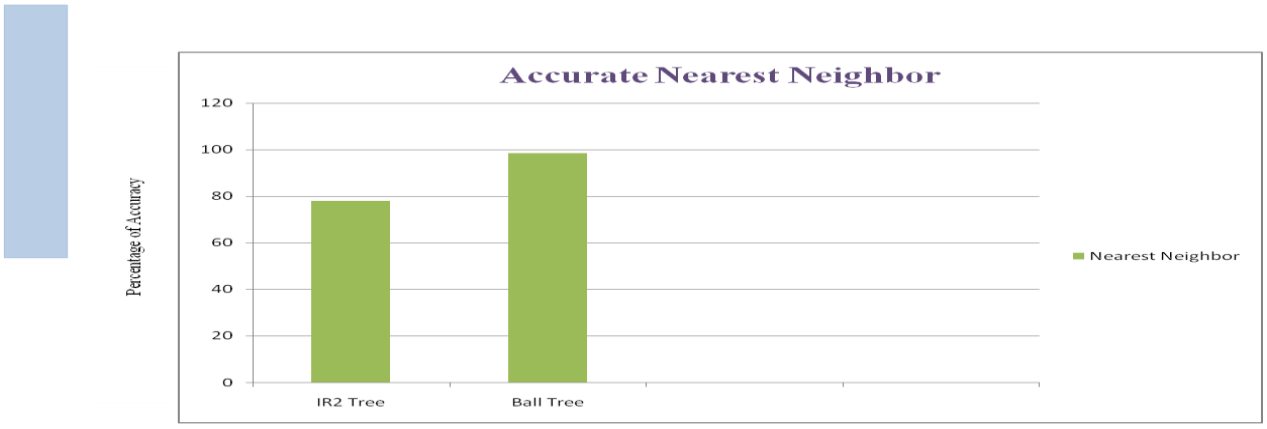


FIG Accuracy in percentage of NN

Table 2 Percentage of Accuracy in finding NN

X-axis	Y-axis	Accuracy of Nearest Neighbour	Percentage of accuracy in NN
0-60	IR^2 tree	78	78%
60-100	Ball tree	99	99.95%

Justification of Result

	INPUT	OUT PUT
For searching NN queries Location manager Keyword Location priority overlay on map stop process	Login for registration Send request to process location Keyword(related search) Enter the location If keyword matches and having max-priority points points are located	Registration completes Location manager loaded to locate points based on location The place matches keyword Nearest point in the location is shown Max-priority nearest Neighbor are found Locate points on map Stop searching for NN

Conclusion

The Important properties of indexing techniques support k spatial keyword search and the Nearest Neighbor (NN) are found based on the max-priority points which are the actual real time closest nearest neighbors by using Ball tree one could be able to find them for exact real time answers and the Ball-tree Nearest Neighbor algorithm efficiently support top k spatial keyword search with Max-priority

Future work

The further work can be by adding mobile tools to the prioritized Nearest Neighbor search one can find out different inputs by travelling and their neighbors gets changed by movement of the mobile signal changes added to the spatial data base may also give the real Nearest Neighbor.

REFERENCES

[1] IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 4, APRIL 2014.“ finding objects that are nearest to a given location that contains a group of keywords”.

[2] 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.

[3] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan, “Keyword Searching and Browsing in Databases Using Banks,” Proc. Int’l Conf. Data Eng. (ICDE), pp. 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,” Proc. 31st Int’l Conf. Conceptual Modeling (ER), pp. 16-29, 2012.

[5] X. Cao, G. Cong, and C.S. Jensen, “Retrieving Top-k PrestigeBased Relevant Spatial Web Objects,” Proc. VLDB Endowment, vol. 3, no. 1, pp. 373-384, 2010.

[6] 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.

[7] B. Chazelle, J. Kilian, R. Rubinfeld, and A. Tal, “The Bloomier Filter: An Efficient Data Structure for Static Support Lookup Tables,” Proc. Ann. ACM-SIAM Symp. Discrete Algorithms (SODA), pp. 30- 39, 2004.