

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

Enhancing the Performance of Geometric Algorithms in Geospatial Data

R.Kavitha, N.Sivakumar

P.G. Student, Department of CSE, University College of Engineering, Anna University (BIT Campus), Trichirappalli, Tamilnadu, India

ABSTRACT: Geometric Algorithms is a computer science issue in Geographic Information System. However some algorithms are better than the others due to the response time and accuracy. This paper describes the performance of proposed Geo-Algorithm that deals with the GeoSpatial data. GeoSpatial data identifies the geographic location and features of earth such as location and related latitude and longitude. A new algorithm is proposed, that helps to find the nearest neighbor. Experimental studies shows that proposed Geo-Algorithm is outperforming when compared to the existing one. This algorithm helps to find accurate Nearest Neighbor. For this we have used weather data and location related latitude and longitude dataset.

KEYWORDS: Geometric Algorithm, Geographic Information System, GeoSpatial data, Nearest Neighbor.

I. INTRODUCTION

Computational Geometry is one that describes the studies of data structure and algorithm for Geometric problems. Geometric Algorithm used to solve Geometric Problems such as Range Searching, Point Location, Nearest Neighbor and Ray Tracing. Spatial data mining denotes the spatial relationships, extraction of knowledge, or other interesting patterns that are not clearly stored in spatial data sets. Spatial data are mainly required for GIS whose information is related to geographic locations. The amount of spatial data that can be collected in such systems is also increasing exponentially. The complexity of the data found in these databases means that it is not possible for humans completely to analyze the data being collected. Spatial database is one of large database that manage multidimensional objects such as lines, points etc..

Spatial partitioning is a step to define, generate and represent partitioned data. Spatial Data is partitioned based on spatial attribute. Spatial data is partitioned as tree and recursive subdivision of space. KD tree algorithm is used for spatial partitioning. The purpose is to hierarchically decompose space into small number of cells, such that no two cells contain too many objects. Fast way to access any input object is by its position. Algorithm construct kd tree by partitioning point set recursively along different dimensions. Space partitioning is the process of dividing a space into two or more disjoint subsets. Space-partitioning means that divide large regions into several small regions. Then the partitioning technique applied to small regions recursively. These regions are organized into a tree called as spatial-partitioning tree.

Cluster analysis or clustering is the task of grouping a set of objects, that objects will be similar means they are in the same group or otherwise in different group. To build a hierarchy of clusters called hierarchical cluster analysis. Agglomerative hierarchical clustering is one of the bottom-up clustering where large clusters have some small sub-clusters, which in turn have several small sub-clusters, etc. Agglomerative hierarchical clustering starts with every single object in a single small cluster. In each successive iteration, it agglomerates or merge with the closest pair of sub-clusters, until all of the data is in one large cluster.

Nearest Neighbor technique is simple and efficient that based on the structure of data. Various techniques available for solving the nearest neighbor problems. K-Nearest Neighbor is simple and fast technique deals with large data samples

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

that use the nearest neighbor rule. Ball Tree K-Nearest Neighbor deal with high dimensional data to improve speed to find K-Nearest Neighbor to perform Geometric learning tasks. Quad tree is partition the space into two dimensions that means root node partitioned exactly into four children. Octree is data structure that partition space into three dimension space such as recursively subdivides space into eight octants. R Tree is one of spatial indexing technique that indexing multi dimensional information such as geographical coordinates that store geometric objects and also used for spatial access methods. R+ Tree differ from R Tree that avoids overlapping of coordinates. R* Tree is robust and efficient technique used to structure geometric objects.

The aim of this paper is to propose a new algorithm to find the nearest neighbor and describe the performance of algorithm in Geo Spatial data. This paper is organized as follows. Section 2 gives the review of previous work on finding nearest neighbor. Section 3 elaborates proposed methods and analysis and the discussion about our work is noted in Section 4. Section 5 presents our conclusions and gives a brief overview of future work.

II.RELATED WORKS

A few works related to nearest neighbor techniques include Mohammed Otair (2013)[1],who proposed a technique that uses spatial partitioning and clustering using kd tree. In[2],Daniel Mullner have discussed about the hierarchical and agglomerative clustering. In [3], Nitin Bhatia and Vandana (2010) gives the survey of various nearest neighbor techniques and their observations. In[4], Sudhir Singh and Nasib Singh Gill has proposed a nearest neighbor clustering algorithm to define the complexity among the use of the algorithm. In[5],Jerome h.friedman et al(1977) define the nearest neighbor problem and find best matches of an algorithm. This work focuses on finding the nearest neighbor that take input as an location along with related latitude and longitude.

III.PROPOSED WORK

A new algorithm is proposed to solve Geometric problems of Nearest Neighbor. It gives better performance when compared to existing one. It helps to find nearest neighbor in quick and efficient manner.

A. Spatial Partitioning and tree construction

Kd Tree is used to partition the data space hierarchically into number of subspaces and are used for storing the spatial coordinates. This data structure is like a binary search tree useful for the range and nearest neighbor searches. It is one type of partitioning based on the level in a tree along with their dimensions. Kd tree divide space into small number of cell where no one contain more than two inputs. Tree has root node and internal nodes. Trees have some values that are stored in leaves. Splitting criteria differ on the levels in a tree.

Constructing Kd tree involves following steps

Step 1:Start with root node, and moves down the tree recursively, To insert a point based on the splitting criteria first note down the level that is even and then determine whether it goes to left or right depend on the point having value less than or greater than the root along x-dimensions.

Step 2:Update their root assign current node as root node, to insert point ,Check whether the level is odd and then determine whether goes to left or right depend on the point have value less than or greater than the root along y-dimensions.

Step 3: The algorithm ends when all the points are inserted based on spatial partitioning.

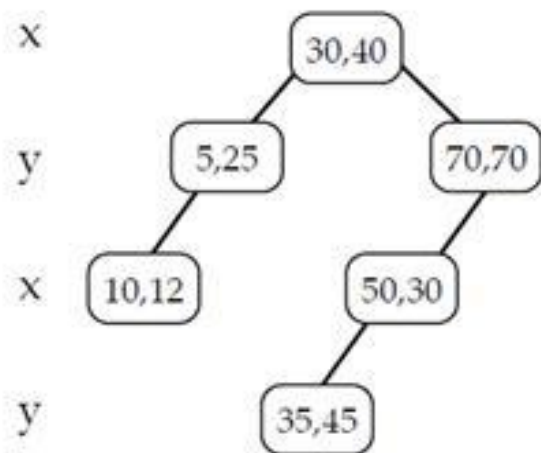
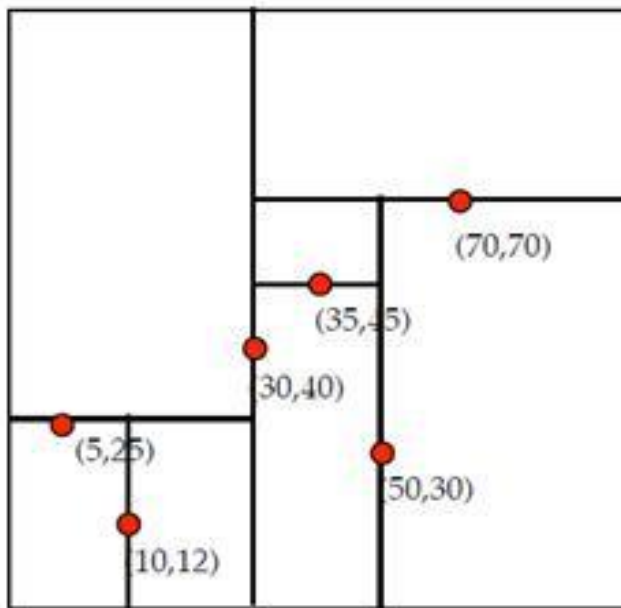
International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

Spatial Partitioning:

Insert:(30,40),(5,25),(10,12),(70,70),(50,30),(35,45)



Algorithm:

```

Construct kd-tree(intlevel,set coordinates)
Assign a Root Node(x,y)as initial node
Construct n child nodes;
Assign levels for constructing a tree
If level=odd
For 1 do upto n
If Root x value>child x value as the x-dimension
Insert into left subtree else insert into right subtree
Elseif
Consider y dimension
Repeat until.
  
```

B. Spatial Clustering

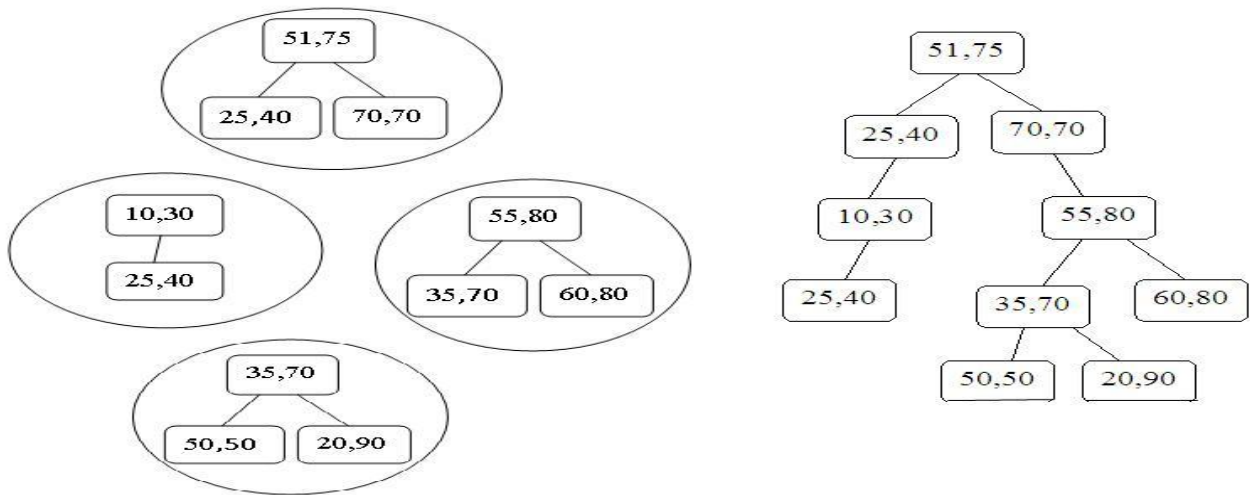
Clustering is major task to find pair of clusters to merge with the nearest neighbor cluster to form single cluster. Each node follows the nearest neighbor in a tree which has sequence of nodes. Sequence of iteration follows each tree is merged with the nearest neighbor tree .Finally a tree is obtained.

Nearest- neighbor chain algorithm is a method that can be used to perform several types of agglomerative hierarchical clustering, using an amount of memory that is linear in the number of points to be clustered.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015



Algorithm:

```

After the kd tree construction
Merge the nearest neighbor tree
For 1 do upto n
{
For 1 do upto n
{
For (i th node=find nearest(jth node) ;
cluster k;
k++;}}

```

C. Nearest Neighbor Search using Kd Tree

Nearest Neighbor search is to find which set of coordinates in a set is closest to another set of coordinates in a set. It is similar to the search of a nearest neighbor search in Binary Search Tree. Kd tree is useful for the searching of coordinate. Kd tree is quick one that help for nearest neighbor search and range search.

Algorithm:

```

Nearest Neighbor(q(x,y),Root)
Assign n nodes
If distance (q(x.value)<Root(x.value) then
Return Root(x value)
Else
Nearest Neighbor value of Root(x,y)

```

D. Proposed Nearest Neighbor Algorithm(Geo-Algorithm)

In this system to propose a new algorithm named as Geo-Algorithm, helps to find nearest neighbor. Location related latitude and longitude of ten cities are passed to the algorithm proceed as follows:

Step 1: Get the input latitude and longitude from user by selecting location via google map.

Step 2: Construct kd tree and start with user selecting city as root city containing points as an latitude and longitude.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

Step 3: Then move down the tree from new city point being inserted.

Step4: Starting with root node, the algorithm recursively moves down the tree levels and set of coordinates as an city (lat, long).

Step 5: If the level is even, compute the x-value of the root node of tree. Then insert set of coordinates in left that have the x-value less than the x-value of the root node of the tree.

Step 6: Otherwise, Insert set of coordinates in right that have x-value greater than the root node of the tree.

Step 7: if the level is odd, compute the y-value of the root node of the tree. Then insert set of coordinates in left that have y-value less than the y-value of the root node of tree.

Step 8: Otherwise, Insert set of coordinates in right that have y-value greater than the y-value of the root node of tree.

Step 9: After the construction of kd trees, Each tree connects with the another kd tree.

Step 10: Finally, Merge all the tree to form one tree then find nearest neighbor using kd tree.

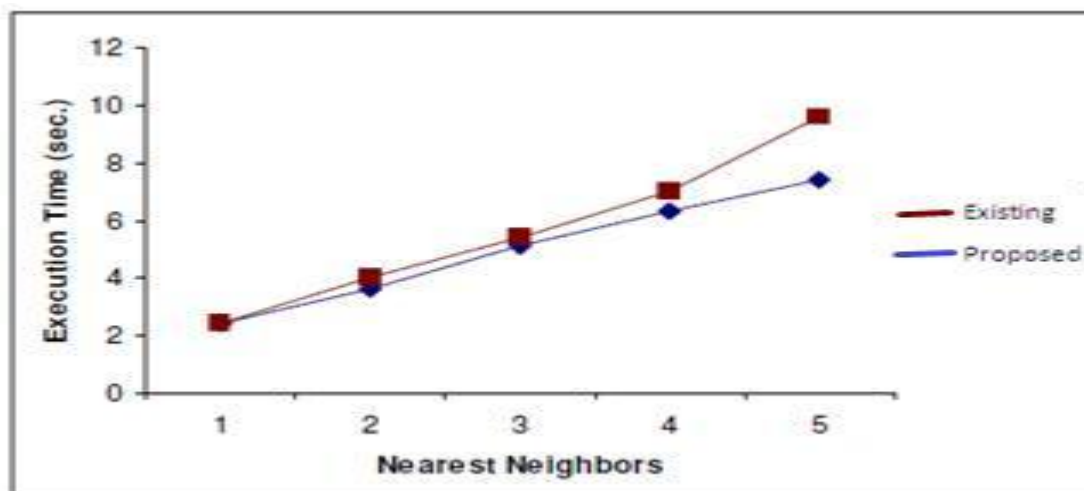
Step 11: if the level is even comparison value as an x-value of the coordinate else y-value of the coordinate.

Step 12: If the comparison value less than the value of the root node of tree search node available in the left-hand subtree of tree else there in the right hand subtree of tree.

The algorithm finishes the tree traversal means search is to complete. This algorithm saves computation time and gives better performance.

IV. RESULTS AND DISCUSSION

The proposed scheme reduce processing time to provide accurate Nearest Neighbor. Evaluation chart for rough estimates defines that there improvement in proposed technique when compared to the existing technique[1] with reduced processing time.



International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

V. CONCLUSION

In this paper, provides Approximate Nearest Neighbor using a new algorithm. Proposed scheme provides good result in order to find nearest neighbor fastly with low computation time. Result of our proposed work compared to existing work described in a discussion part of the paper.

In future, we plan to carried the work for an new algorithm to make that to improve the performance.

REFERENCES

- [1] Mohammed Otair, "Approximate k-Nearest Neighbor based Spatial Clustering" International Journal of Database Management Systems Vol.5, No.1, Feb 2013 page:97-108
- [2] Daniel Mullner, "Modern Hierarchical agglomerative clustering algorithms" sep 2011 page:1-29
- [3] NitinBhatia, Vandana, "Survey of Nearest Neighbor Techniques" International Journal of Computer Science and Information Security, Vol.8, No.2, 2010 page 302-305
- [4] SudhirSingh, Nasib Singh Gill "Enhancing The Performance of Nearest Neighbor Clustering Algorithm" International Journal of Computer Applications and information technology, Vol.3, IssueII, sep 2013 page14-19.
- [5] Jerome h.friedman, Jonlouisbentley, Raphaelarifinkel, "An algorithm for finding best matches in logarithmic expected time" ACM Transactions on Mathematical Software, Vol.3, No.3, sep 1977 page:209-226.
- [6] A.stewartfotheringham, peterA.rogeron "GIS and spatial analytical problems" Geogrphic Information Systems 1993, Vol.7, No.1 page:3-19
- [7] YufeiTao, Cheng Sheng, "Fast Nearest Neighbor search with keywords" IEEE Transactions on knowledge and data Enginnering.
- [8] Yinanjing, LingHu, Wei-Shinn ku, CyrusShahabi " Authentication of k-nearest neighbor query on road networks" IEEE Transactions on knowledge and data engineering, Vol.26, No.6, june 2014.