# Research and Implementation of Spatial Database Classification Query Algorithm

1st Chong Xu
*BIM Computing Research Center*
*Shenyang jianzhu University*
Shenyang, China
32500103@qq.com

2nd Xin Liu
*Human Resources Department*
*Shenyang jianzhu University*
Shenyang, China
6620274@qq.com

3rd Wei Wang
*BIM Computing Research Center*
*Shenyang jianzhu University*
Shenyang, China
519836646@qq.com

4th Yong Shi
*BIM Computing Research Center*
*Shenyang jianzhu University*
Shenyang, China
674365154@qq.com

5th Ning Tao
*BIM Computing Research Center*
*Shenyang jianzhu University*
Shenyang, China
23010150@qq.com

***Abstract*—spatial database refers to the sum of the geospatial data related to application stored by GIS on computer physical storage medium. Due to the low efficiency of the traditional query method and the single query information, it is an important aspect of spatial data query research that how to excavate information with specific meaning quickly and effectively in huge data. In order to more deeply excavate the information of the spatial database, the spatial data query conducted uses range queries, K-nearest neighbor algorithm queries, optimization of anti-nearest neighbor query algorithms, and realization of a large number of classified spatial data queries. The results obtained by using the inverse nearest neighbor query algorithm are that in all sets belonging to a class of points, a certain position is used as a collection of points from its nearest location, and the algorithm is verified by data experiments. The feasibility of the optimal query is achieved.**

***Keywords—spatial database, K nearest neighbor algorithm, Anti-nearest neighbor query algorithm***

## I. INTRODUCTION

The rapid development of computer technology and data collection technology has made it possible to collect and store information from a wider range and at an unimaginable speed, and it is hoped that it can be analysed at a higher level in order to make better use of these data. Due to the large amount of data in the spatial database, it has the characteristics of high accessibility, complex spatial data model, and joint management of attribute data and spatial data. The result is that users have a wealth of data resources but can not use them effectively. Spatial database technology has gradually[1] become the mainstream technology of spatial data management instead of traditional document management. Nowadays, the research based on spatial database has become one of the most abundant branches of applied research technology in the field of computer science and technology[2]. It is an important aspect of spatial data research that how to excavate meaningful information in huge data.

In the Visual C++ 6.0 environment, the range query, the K-nearest classification algorithm query, and the anti-proximity query(RNN) algorithm are used to query the processed data[3]. The scope query is to find a certain type of point within a certain range around a known point. The K-nearest classification algorithm is to find several data points that are the smallest distance from the known point in the entire data set. The anti-neighbor query algorithm is to find out where a data set is closest to the known point.

## II. CLASSIC QUERY ALGORITHM

### *A. Range Query Algorithm*

Specifies a point and a range in the space, centered on the specified point in the layer of the spatial feature object, with the range set as the radius, and the point within the range as the result. The quadrangle a is the specified point. With the center of the point, a certain range is the radius. The point that falls into the circle displayed in the graph within that range is the result of the query in that range.

### *B. The Nearest Neighbor Search Algorithm*

The The most recent neighbor query issue was proposed by Kunth in 1973, namely the post office issue. It can be described as: Given the set S of N points in the N-dimensional space, this N point is stored in a data structure, so that for any query point Q in the space, a large nearest neighbor can be effectively found., That is, a point P is found in S, which is the closest to Q. The nearest number can be one or more KNN.

According to the above definition, K nearest neighbor queries can be defined.

Definition 2.1 K nearest neighbor query: Given a set of points S in an N-dimensional space and a query point Q, the definition of Q's K nearest neighbor query can be described as:

KNN(Q) ={P1, P2, ..., PK}

  DOI: 10.25236/iccpb.2018.005

Where Pi ∈ S and I = 1,2, ... K, there is D(Q, Pi) ≤ D(Q, P) for <UNK> P ∈ S and P <UNK> KNN(Q).

K nearest neighbor algorithm(KNN): Which class is the circle to be given, a triangle or a square? If K = 3, because the proportion of triangles is 2/3, circle a will be assigned to the triangle class. If K = 5, since the square ratio is 3/5, circle a is given a square class.

### *C. Anti-Nearest Neighbor Query Algorithm*

For data and S and query point Q, the nearest neighbor of each point P in the data set S is calculated in advance, and then each point is corresponding to an adjacent circle, and the R-tree index structure of these circles is further established., Called the RNN-tree, using this index structure, it is easy to find Q's nearest neighbor. However, because the RNN-tree was established to counter the nearest neighbor query, rather than the nearest neighbor query, Flip Korn and S. Muthukrishnan had to establish another R-tree to perform the nearest neighbor query and other spatial queries.

Formal definition of RNN queries. Assume that D(P, Q) is the distance between two points P and Q.

Definition 2.2 Against nearest neighbor queries: Given a data set S and a query point Q, the inverse nearest neighbor query is a subset of the set RNN(Q) of S:

RNN(Q) ={P ∈ S | <UNK> R ∈ S: D(P, Q) ≤ D(P, R)}

Reverse Nearest Nearer(RNN): To determine which circles and triangles consider the middle square as the nearest square[ 2] And ... First find the square closest to yourself in all circles and triangles. For those figures that consider the middle square to be the closest square to themselves, it is the result of the inverse neighbor algorithm.

## III. Design and Implementation of Algorithm

### *A. Enquiries*

The range query algorithm implements the layer of the spatial feature object selected by the user and uses the Search Within Division function of the Cmap XLayer class in MapX. This function is specifically used to determine the layer range query for a known location centered and a known distance radius. Set the query in the new MapX layer for display.

### *B. The Nearest Neighbor Query Algorithm is Implemented*

First, an array object of the newly established data structure CMFeatureclass. Separates the feature object, the set of graphs. Calculates the distance between each graph element and a known point in this set of graphs, and sorts this distance by insertion into an array of CMFeatures classes. If this graph element can be inserted into the array by distance calculation, then the graph element and its distance from the known point are inserted into the array[4].

When all the meta sets have completed the above operation, the meta elements in the array of the CMFeatureclass are displayed in the map[5].

### *C. Implementation of the Inverse Nearest Query Algorithm*

First, the feature object, the set of graphs, is separated. According to Theorem 3.1, the "My Position" point is the origin of its plane, and the horizontal line of the point is the line L1, and the map is divided into six regions: S1, S2, and S6. Calculates the graph element that falls closest to the known point in each regional graph element set and deposits the graph element into an array of objects of the newly established data structure CMRnFeel class. It can be seen from Theorem 3.1 that there are up to 6 objects in this array. Then take the graph elements of the six objects as the center point, and find its KNN(K = 1) in the set of graph elements that satisfy the characteristics of the "position point nature" if its KNN happens to be the graph element of "my position"., Then add it to the new layer and show it on the map.

TABLE I. Table Type Styles

| Structure Name | Structure representation | Structural functions |
|---|---|---|
| RNN query result graph | CMFeature | Storage feature objects and their distance from known graphs |
| RNN query result graph | CMRnnFeature | The graph element closest to the known point, the distance from the known point, and the position of the store in the meta set when storing RNN queries |

## IV. Conclusion

Spatial database technology is the core technology of GIS data organization. It is also a combination of geographic science, mapping science, computer science and information science. Spatial database technology has gradually become the mainstream technology of spatial data management instead of traditional document management. However, due to the large amount of data in spatial databases, users have abundant data resources but can not use them effectively. This system is to use some algorithms to help users to classify a large number of spatial data query. The system not only realizes the traditional input location name query on the map, but also classifies all map data according to their nature. Users can enter a range of queries in the map of a good class, can query the location of the closest to me in the map, and can also use this system to make decisions on site selection.

## REFERENCES


[1] Liao Haojun , Han Jizhong , Fang Jinyun.All-Nearest-Neighbor Queries Processing in Spatial Databases.Journal of computer research and development,2011,48(1):17～22.

[2] Hao Zhongxiao,Wang Yudong,He Yunbin.Line Segment Nearest Neighbor Query of Spatial Database.Journal of computer research and development,2008,45(9):5～11.

[3] XIAO Hui,YANG Bisheng.An Improved KNN Search Algorithm Based on Road Network Distance.GEOMATICS AND INFORMATION SCIENCE OF WUHAN UNIVERSITY,2016,33(4):19～33.

[4] Li Gou-Hui,Luo Chang-Yin,Li Jian-Jun.Authentication of reverse k nearest neighbor query//Lecture Notes in Computer Science 9049,Berlin:Springer,2015:625-640

[5] Wu Ding-Ming,Choi B,Xu Jian-Liang,et al.Authentication of moving top-k spatial keyword queries.IEEE Transactions on Knowledge and Date Engineering,2015,27(4):922-935