Indexed in
Scopus Compendex and Geobase Elsevier,
Geo-Ref Information Services-USA, List B of Scientific
Journals, Poland, Directory of Research Journals

International Journal of Earth Sciences and Engineering

www.cafetinnova.org

ISSN 0974-5904, Volume 09, No. 06

December2016,P.P.2538-2543

Analysis and Application of the Shortest Path Algorithm based on Geographic Information System

HAIFENG HU AND FEIFEI WANG
College of Information Technology, PingDingshan University, Pingdingshan, Henan 467000, China
***Email:** hhf1421@163.com*

Abstract: Data structure in GIS has been studied, thoughts of various kinds of shortest patch algorithm, its applied data structure and features of various kinds of data structure have been analyzed in this paper; development of shortest path algorithm in real time and parallel has been discussed, spatial distribution features of GIS have been studied, shortages of traditional shortest path algorithm in this specific application have been analyzed by combining the specific situation of seeking for the shortest path in GIS road traffic, one improvement algorithm which makes full use of the spatial distribution characteristics in GIS has been proposed and one new search idea, which is direction first search has been proposed for the path search of traditional Dijkstra algorithm. Finally, the improved algorithm has been realized through programming and the validity of this algorithm has been verified through experiment.

Keywords: *shortest path, geographic information system, GIS, Dijkstra algorithm*

1. Introduction

In current GIS application field, there are a lot of researches and applications in shortest path search problem, in which the efficiency problem of shortest path search algorithm draws the common attention and needs to be solved urgently in practical application. The adopted algorithm is mainly from Dijkstra algorithm in graph theory domain and the result of this algorithm operation is the shortest path of one vertex to all its vertexes. Dijkstra algorithm adapts to the changes of network topology and performance is stable, therefore, it has been applied extensively in the selection of computer network topology path and GIS. For example, one important mobile service location service (Mobile GIS) development from America can help users check location information and find ways without asking others; in the functional module of “asking directions”, it adopts Dijkstra algorithm to search the shortest path. The advantages of Dijkstra algorithm also include simple program design and strong generality. But it is not suitable to apply traditional Dijkstra algorithm to the shortest path search directly [1]. First, Dijsktra algorithm is not for two specific points, while in the shortest path searching of GIS, it usually needs to search the shortest or seemingly shortest path between two special points, therefore, the efficiency is low; in addition, traditional Dijkstra algorithm adopts adjacency matrix data structure and takes up a lot of space, which wastes the computer resources seriously and not suitable for the practical situation of great nodal points in GIS. More importantly, traditional Dijkstra algorithm is a general algorithm solving the shortest path of source point in plane to all its points and in the adoption, it does not consider about the spatial distribution feature of objects to handle [2]. Therefore, in the significance of generality and completeness, there is no doubt that Dijkstra algorithm is very excellent, but based on GIS spatial distribution feature and shortest path search in GIS, it is completely necessary to make further improvement for the algorithm [3] and its efficiency.

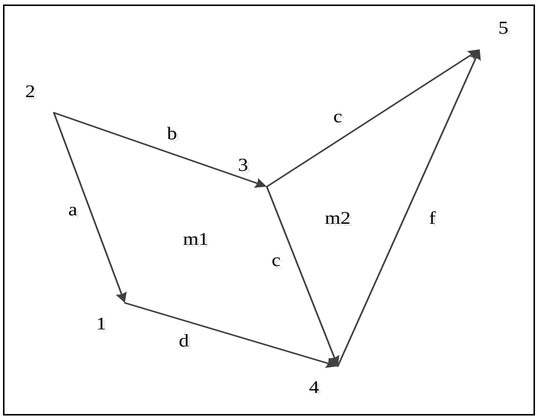


Figure 1 *a simple map*

2. Data model of GIS

Establishing geologic data model is an important part of geographic information system. One of the cores of database design is to design a good data model for describing the relationship between data content and data. Common data model of database is hierarchy, mesh and relationship. In the following, it will take a simple example (Figure 1) to explain the data organization forms and characteristics of former three data models [4-5].

2.1 Hierarchical model

Hierarchical model is directed tree used to record type node. Main characteristics: except for root node, any node has only parent node. One tree record corresponds to several sub records and one sub record only corresponds to one parent record. Figure 2 is the hierarchical model of Figure 1.

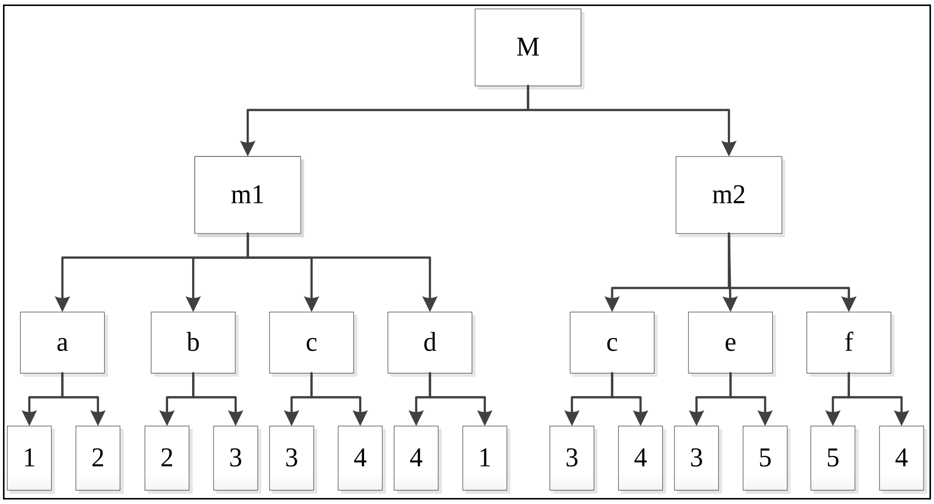


Figure 2 *Hierarchy model*

2.2 Network model

In GIS, if adopt hierarchical model for the topology relationship among common points, lines and physical elements, rendezvous redundancy increases, which brings about difficulty to topology checking [6-8]. Network model takes record type as node, which has two major differences with hierarchical model:

(1) One sub node can have two or more parent nodes;
(2) Any two nodes can have two or more relationships; Figure 3 is the network model of Figure 1. Each square in the Figure is called as a node, representing one physical element. Each physical element is expressed with one record and connection among different elements is connected with internet. One node can have many parent nodes or one node may be the sub node of many networks [9]. In GIS, graphic data generally adopts topology data model, similar to network model. The difference is that topology model adopts target identification and network connects the pointer.

2.3 Relationship model

Relationship model is a digital model, which summarizes the logic structure of data into a two-dimensional table 1 meeting certain conditions. One entity is composed of several relationships and the set of relationship table 2 composes relationship model. The following relationship table 3 is the relationship between polygon, boundary and node listed in Figure 1. The main advantage of relationship model is consistent description. Relationship among objects is that the data itself expresses the relationship among them through public value implicitly. Relationship model uses relational algebra and relational operation to operate data, the structure is simple and flexible, revising and updating data is convenient, easy to maintain and understand.

Table 1 *Boundary relation*

polygon	Edge number	Side length
I	a	30
I	b	23
I	c	16
I	d	24
II	c	16
II	e	15
II	f	18

Table 2 *Boundary and node relation*

Edge number	The starting point	The end point
a	1	2
b	2	3
c	3	5
d	4	1
e	3	5
f	4	4

Table 3*node coordinate relation*

Edge number	x	y
1	26.5	12.3
2	25.1	42.3
3	32.1	45.6
4	42.3	35.2
5	65.3	42.6

3. Analysis the shortest path algorithm

Original Dijkstra algorithm is used to solve the shortest path in network from one source point to all the other points in the Figure 3. This algorithm divides network nodes into unlabeled node, temporary label node and permanent labeled node [10]. First, initialize all nodes in network as unlabeled nodes and nodes in search process and connected to the shortest path node are temporary labeled node [11-13]. Each cycle is to search the node with the shortest distance to source point from temporary labeled points as permanent labeled node, terminate algorithm till all the nodes become permanent labeled nodes. In original Dijkstra algorithm, temporary labeled nodes are saved disorderly, which becomes the bottleneck of Dijkstra algorithm. Because when search for the nodes with the shortest path in temporary labeled nodes every time, it needs to traverse all temporary labeled nodes. The solution is to order temporary labeled nodes based on the shortest path, each search path does not need to transverse all or few temporary labeled points, this is one of the important starting points of various kinds of optimization algorithms based on original Dijkstra algorithm; another important optimization channel is trying to decrease the number of temporary labeled nodes in shortest path analysis, which is trying to search from the starting point along the direction with the fastest speed to the terminal and then reach the target nodes as soon as possible (the improvement algorithm proposed in this paper is studied and improved from this perspective) [14-18]; in addition, in practical application, make corresponding improvement for path algorithm based on different corresponding paths of specific situations, for example, solution in GIS may possibly solve the shortest path between fixed starting point and end point, but not the shortest path of this point to all points.

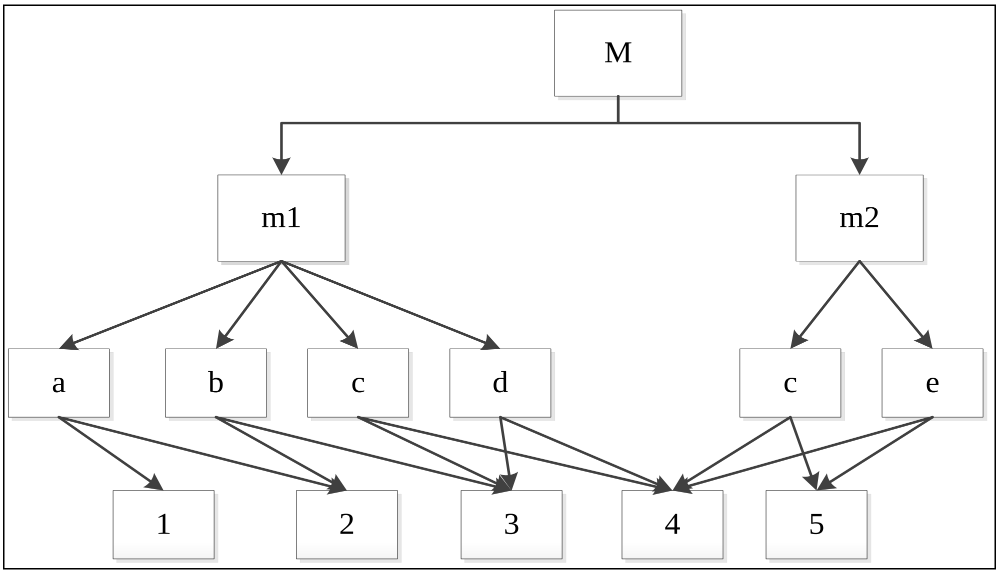
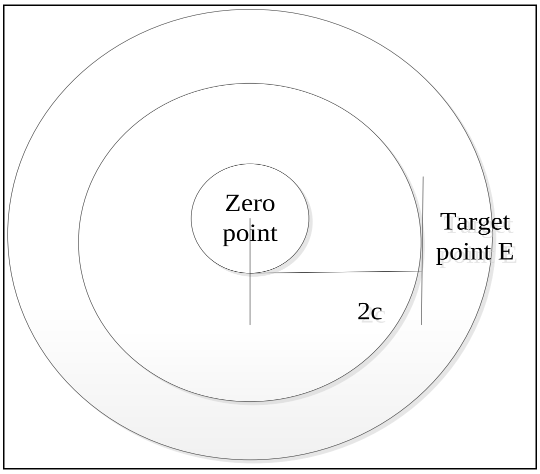


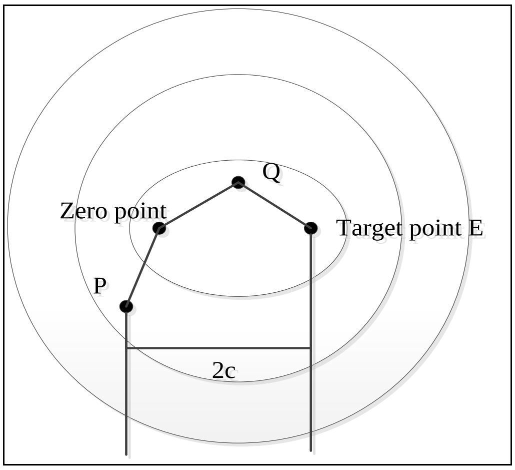
Figure 3 *Network model*

3.1 Linear optimization Dijkstra algorithm idea

Comparison diagram of searching process between linear optimization Dijkstra algorithm and original Dijkstra algorithm is as shown in Figure 4.



(a)The original Dijkstra algorithm



(b) Linear optimal Dijkstra algorithm

Figure 4 *Dijkstra algorithm*

It can be seen from Figure 4 (a) that the searching process of original Dijkstra algorithm is similar to a series of concentric circles taking source point as the center of the circle. The searching process does not consider about the direction or location where the end point locates. In the searching process of source point, the probability of being search between other nodes and end point is the same; the searching process of linear optimization Dijkstra algorithm is similar to a series of concentric ellipse taking source point and end point as the focus. Because the selection principle of permanent labeled point is: the smallest sum between the shortest distance of current node to source point and linear distance from this node to end point is taken as permanent labeled point. Therefore, its searching process is more prone to end point, the speed of searching process reaching end point is obviously quicker than original Dijkstra algorithm and the searched nodes is less then original Dijkstra algorithm.

As shown in Figure 4(b), set Sp=2, Sq= 5, p E = 8, q E =2. Start from source point to make original Dijkstra algorithm searching. First, select nodes p and q as temporary labeled points. Because the length of Sp is smaller than that of Sq, so select node p as permanent labeled node; and then take p as the new source point and then make the next round searching, while the node p will be eliminated in the end. If adopt linear optimization Dijkstra algorithm, it is because Sp+ Ep=10, Sq+ Eq =7, so node q is selected as the permanent label point. As shown in Figure 5, L1 is the length of the shortest path of temporary labeled point P1 to source point, Ln is the shortest path length of temporary labeled point Pn to source point, D1 is the linear distance of node P1 to end point, Dn is the linear distance of node Pn to end point, line between P1 and Pn is C. Selection principle of permanent labeled point of original Dijkstra algorithm is: if L1>Ln, select Pn as permanent labeled node; If L1<Ln, and then select P1 as permanent labeled node.

Selection principle of permanent labeled point of linear optimization Dijkstra algorithm is: if L1+D1>Ln+Dn, and then select Pn as permanent labeled node; if L1+D1<Ln+ Dn, and then select P1 as permanent labeled node.

One premise condition of the selection principle of permanent labeled node of linear optimization Dijkstra algorithm is that L1 is the shortest path of P1 to source point, which is that the shortest path value of P1 will not be revised again, which is that known: L1+D1<Ln+ Dn, Prove: Ln+C>L1. Ln+Dn>L1+D1, so Ln+Dn-D1>L1; because C>D-D1, Ln+C>L1. Finished

The key of linear optimization Dijkstra algorithm superior to original Dijkstra algorithm lies in that its maximum searching range has been decreased significantly and then the searching speed has been improved; Figure 6 displays the biggest searching range of two algorithms. Circle in the Figure is the biggest searching range of original Dijkstra algorithm and ellipse is the biggest searching range of linear optimization Dijsktra algorithm.

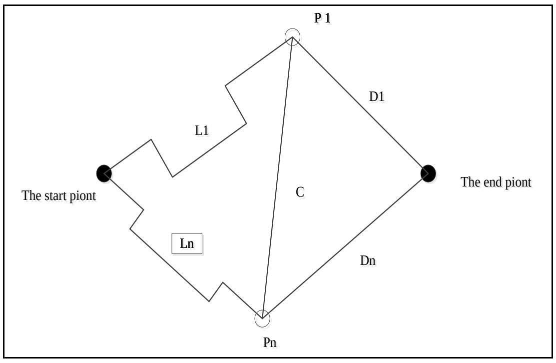


Figure 5 *Beeline optimizing Dijkstra`s shortest path algorithm*

3.2 N angle improvement in reduction algorithm-sub graph method

Based on above idea, specific descriptions of improved Dijkstra algorithm are as following:

(1) Connect source vertex and end point of network graph together into straight line;

(2) Select suitable m value and divide network graph into m sub-graphs based on above division method;

(3) Calculate each sub-graph respectively with adoption of Dijkstra algorithm.

(4) Compare the source vertex and end point of each sub-graph with the source vertex and end point of neighboring sub-graph:

If the end point of former neighboring sub-graph is the source vertex of following sub-graph, and then connect these two sub-graphs with two shortest paths calculated by Dijkstra algorithm; moreover, it is believed that the connected path is one part of the shortest path;

If the end point of former neighboring sub-graph is not the source vertex of following sub-graph, and then revise the end point of former sub-graph, take the source vertex of following sub-graph as its new end point and then use Dijkstra algorithm to recalculate to get d_{i+1}`. At the same time, revise the source vertex of following sub-graph, replace with the end point of former sub-graph and use Dijkstra algorithm to recalculate to get $d_{i+1}$` similarly. Based on calculation results, d=min $\{d_i+d_{i+1}`, d_i`+d_{i+1}\}$ (in which d_i and d_{i+1} are the lengths of the shortest path after calculating former sub-graph and following sub-graph,) at the same time, connect two shortest paths corresponding to d together as part of the shortest path.

(5) Repeat step (4) until find the shortest path connecting source point and end point.

3.3 Realization of improved algorithm

In the process of searching the shortest path between two points, the nodes needed to be passed can't be confirmed in advance. If use array form to construct All Node and Leaf Node, the reserved space should not be smaller than the requested space in practice, which will waste a lot of inventory or exceed the limit of array. Therefore, the paper adopts pointer form which can allocate inventory dynamically to construct All Node and Leaf Node, which is to use chain stack and chain queue. There is also one advantage of selecting this method, which is that when select one need from Leaf Node stack top to embed into All Node, there is no need to allocate inventory again and only need to connect this point to the end of All Node and it can save in both inventory and operation time.

Under this big premise, it also needs to define one pointer type Nodeptr, which points to one defined data type TNode. Because the node structure of All Node and Leaf Node is basically the same, the only difference is that there is one more sub node pointer in All Node structure. To keep the unity of algorithm, this paper adopts unified structure during algorithm realization. In addition, because the coordinates of each point needs to be used in this algorithm, it needs to add X, Y value of plane coordinates in node structure, therefore, this node structure is as following:

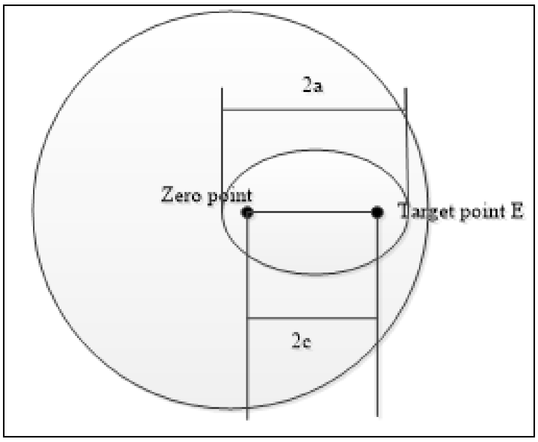


Figure 6 *the sketch map of the most seek arrange of original Dijkstra*

4. Improvement and analysis of Dijkstra algorithm based on GIS spatial distribution features

4.1 Example introduction

When searching for the shortest path from Chongqing Normal University to Chongqing University with the improvement algorithm proposed in this paper, it will find the path as shown by the red line in the Figure 6. When the network is very large and the searching time is extremely long, the algorithm can be terminated after finding out few paths, in which the shortest one will be equal to or very close to the real shortest path, while it is in consistent with the situation in GIS that network data is big and usually needs to be similar to the optimization.

4.2 Experimental verification of improvement algorithm

Based on above ideas, this paper realizes the simulation for the checking function of the best path under Delphi with the help of draw.

Experimental results prove the validity and accuracy of the algorithm.

The algorithm program in this paper is developed with Delphi 7 programming language; it can be embedded into Map Info Professional 7.8 based on Windows2000platform and combine with other functional modules to form more complete GIS application system.

5. Conclusions

This paper starts from the research and analysis of the shortest path algorithm based on Dijkstra, analyzes the shortages and improvement ideas of traditional shortest path algorithm in the special application by combining the spatial distribution features of GIS road traffic and specific situation of searching for the shortest path and finally realizes an improvement algorithm which makes full use of spatial distribution features in GIS through programming. Main research work and innovation points are as following:

(1) The main ideas of classic Dijkstra algorithm and its realization have been studied;
(2) Data structural in GIS application has been studied;
(3) Various kinds of the shortest path algorithm ideas, its application data structure and various kinds of features of data structure have been analyzed;
(4) Development of the shortest path algorithm in real time and parallel has been discussed;
(5) Improvement for the shortest path algorithm based in Dijkstra algorithm in GIS application has been studied emphatically, the shortest path algorithm based on Dijkstra algorithm has been improved for the practical situation of solving the shortest path in GIS, which can solve the search for the shortest path between two specific points and improve the efficiency of algorithm. This algorithm has been realized finally through programming.

References

[1] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *International Journal of Photogrammetry & Remote Sensing*, Volume 58. Issue 3-4, PP 239-258, 2004.

[2] G. Câmara, R. C. M. Souza, U. M. Freitas, J. Garrido, Spring: integrating remote sensing and GIS by object-oriented data modeling. *Computers & Graphic*s, Volume 20. Issue 3, PP 395-403, 2011.

[3] C. S. Jarnevich, T. R. Holcombe, D. T. Barnett, T. J.Stohlgren, J. T. Kartesz, Forecasting weed distributions using climate data: a GIS early warning tool. *Invasive Plant Science &*

Management, Volume 3. Issue 4, PP 365-375, 2010.

[4] H. Samet, Applications of spatial data structures: computer graphics, *image processing, and GIS*, Volume 5. Issue 2, PP 253-25, 1990.

[5] J. Malczewski, GIS-based land-use suitability analysis: a critical overview. *Progress in Planning*, Volume 62. Issue 1, PP 63-65, 1990.

[6] K. H. Kozak, C. H. Graham, J. J. Wiensm, Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, Volume 238. Issue 3, PP 141-148, 2008.

[7] S. L. Lee., X. Li, W. Shi, C. N. Cheung, I. Thornton, Metal contamination in urban, suburban, and country park soils of Hongkong: a study based on GIS and multivariate statistics. *Science of the Total Environment*, Volume 356. Issue 1-3, P 45-61, 2006.

[8] P. E. Osborne, J. C. Alonso, R. G. Bryant, Modeling landscape scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology*, Volume 38. Issue 2, P 458-471, 2001.

[9] B. C. Pijanowski, D. G. Brown, B. A. Shellito, G. A.Manik, Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers Environment & Urban Systems*, Volume 26. Issue 6, PP 553-575, 2002.

[10] G. Walsham, S. Sahay, GIS for district-level administration in India: problems and opportunities. *Mis Quarterly*, Volume 23. Issue 1, PP 39-65, 1999.

[11] J. Xiao, Y. Shen, J. Ge, R. Tateishi, C. Tang, Y. Liang, et al, Evaluating urban expansion and land use change in Shijiazhuang, china, by using GIS and remote sensing. *Landscape & Urban Planning*, Volume 75. Issue 1-2, PP 69-80, 2006.

[12] H. Sand, B. Zimmermann, P. Wabakken, H. Andrèn, H. C. Pedersen, Using GPS technology and GIS cluster analyses to estimate kill rates in wolf-ungulate ecosystems. *Wildlife Society Bulletin*, Volume 33. Issue 3, PP 914-925, 2005.

[13] J. B. Mena, State of the art on automatic road extraction for GIS update: a novel classification. *Pattern Recognition Letters*, Volume 24. Issue 16, PP 3037-3058, 2003.

[14] S. Arlimatti, S. Hassan, A Habbal. Minimization of Communication Cost between Controller Domains with Graph Partitioning Algorithms, *Revista de la Facultad de Ingeniería*, Volume 31, Issue 5, PP 01-08, 2016

[15] N. B. Chang, G. Parvathinathan, J. B. Breeden, Combining GIS with fuzzy multicriteria decision-making for landfill siting in a fast-growing urban region. *Journal of Environmental Management*, Volume 87. Issue 1, PP 139-153, 2008.

[16] A. Carrara, F. Guzzetti, M. Cardinali, P. Reichenbach, Use of GIS technology in the prediction and monitoring of landslide hazard. *Natural Hazards*, Volume 20. Issue 2, PP 117-135, 1999.

[17] X. H. Wang, H. F. Hu. Research on Vehicle Routing Problems Based on the Improved Pareto Ant Colony Algorithm, *Revista de la Facultad de Ingeniería*, Volume 31, Issue 5, PP 114-122, 2016

[18] J. Geoghegan, L. A. Wainger, N. E. Bockstael, Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. *Ecological Economics*, Volume 23. Issue 3, PP 251-264, 1997.