Research on agricultural machinery operation path optimization based on traveling salesman problem (TSP)

Huichuan Liao*
School of Information Engineering, East China Jiaotong University, Jiangxi, Nanchang 330013, China

*Corresponding author. Email: 995292@qq.com

Abstract. Based on the characteristics and properties of TSP, a new heuristic algorithm is designed. Firstly, the nodes on the graph are divided into regions, and then the minimum spanning tree is found in the region. Then, the nodes larger than 2 are broken and reconnected into trees with 2 degrees except the start and end points. Finally, the start and end points of adjacent regions are connected to get the solution of the problem. Two examples show that this heuristic algorithm can effectively avoid the round trip of several points between different regions. The algorithm is applied to the optimization of agricultural machinery operation path, which provides a reference for the daily operation of agricultural machinery households.

1. Introduction
So far, many scholars have proposed various new algorithms for the traveling salesman problem, but it is impossible to list a general polynomial solution for the traveling salesman problem. Therefore, it is necessary to find a relative optimization algorithm. If we use enumeration method to solve the problem, we need to solve n! times combination in n cities, then find the best result to achieve the purpose of solving; with the increase of n value, the combination times increase exponentially, so the calculation efficiency is very low.

At present, the main algorithms for solving traveling salesman problem can be divided into exact algorithm and heuristic algorithm. However, due to the limited scope of exact algorithm, heuristic algorithm is usually used. Although the optimal solution is usually not obtained, the heuristic algorithm is efficient, easy to use, and easy to improve. At present, the branch and bound method is the only accurate algorithm in common use. Other accurate algorithms (such as linear programming algorithm and dynamic programming algorithm) usually need to rely on experience and the scope of solution is very limited, so they are generally not suitable. Heuristic algorithms include nearest neighbor algorithm (NN), simulated annealing algorithm (SA), genetic algorithm (GA), ant colony optimization (ACO), neural network algorithm (NN) and so on, which are widely used to solve TSP; But there is no algorithm to design the solution process from the perspective of practical problems. Based on the analysis of some characteristics of agricultural machinery operation path and the characteristics of traveling salesman problem, an algorithm has been designed in this paper, which is convenient to use and can get relative optimization solution.

2. Traditional farm machinery operation methods
At present, there are two kinds of common agricultural machinery operation planning methods: one is the order operation method, that is, according to the order priority submitted by farmers with operation
requirements. It follows the principle of order queuing, so that all farmers are in a relatively fair state. The planning principle of order type operation is simple, and it does not need professionals to optimize the path, only according to the order priority, which saves a lot of preliminary work. However, the order operation mode is not the optimal operation mode, because it has many shortcomings and deficiencies, mainly as follows: 1) the probability of order operation method being the optimal method of field operation is very low, so it will increase the idle running and other time of the machine, resulting in the increase of fuel consumption and the waste of power; 2) It is not good for the projects that need to catch up with the agricultural time. Because of the extra idling and other time, it may even delay the agricultural time and reduce the crop yield and income.

The second is the nearest operation method, that is, after each operation on the farmland of the previous farmer, the farmer will work on the farmland of the nearest farmer's home, and so on, and finally return to the farm machinery farmer's home. This operation method is consistent with the nearest neighbor algorithm in solving traveling salesman problem. This algorithm means that when solving the problem, it always makes the best choice at present. This kind of nearest neighbor operation method is much more reasonable than the order operation method. However, for large-scale agricultural machinery households, it is necessary to systematically optimize and design single operation to find out the optimal operation method. Therefore, it is necessary to design a new algorithm according to the agricultural machinery operation path optimization problem.

3. Algorithm overview and operation flow

3.1. Algorithm conception and definition

Because solving the traveling salesman problem is actually solving the minimum Hamilton cycle problem, we can consider reducing the dimension of the larger traveling salesman problem model first, so that a large problem can be transformed into several relatively simple small problems.

It is necessary to describe some definitions of the algorithm before the operation of the algorithm process. Degree of node: the number of connecting lines on any node; \( C_{ij} \): the lines that are not connected again after solving the minimum spanning tree process \((i, j \in V)\); \( S_{ij} \): the lines that exist but need to be deleted after solving the minimum spanning tree process \((i, j \in V)\); \( K \): the weight difference between the rewiring lines and the deleted lines within the region; \( q_{ij} \): the starting and ending lines between regions \((i, j \in V)\).

3.2. Algorithm flow and steps

Due to some follow-up problems, a new algorithm is designed according to the application examples of existing heuristic algorithms, which can select the starting point and the ending point in the region first, and then avoid the selection and connection problem between regions. Firstly, all the points in the problem are divided into n regions according to the distance matrix, and the points are clustered and divided according to the clusterdata function in MATLAB. After the region is divided, the minimum spanning tree of each region is calculated, and the lines with degree greater than 2 in the minimum spanning tree of each region are deleted, so that the degree of each point in the region is 0, 1 or 2; the points with degree 0 or 1 in each region are connected, and the entry point and exit point of each region are reserved; In the process of breaking redundant line segments and reconnecting line segments, the principle of equation (1) should be followed, so that after deleting and connecting each region, there are only two points with degree 1, and the degree of the remaining points is 2; finally, connect the region and the points with degree 1 to form a closed loop, so that the degree of each node is 2, and select the smallest solution to minimize \( \sum q_{ij} \).

\[
\text{Min } K = \sum C_{ij} - \sum S_{ij} \tag{1}
\]

It can be proved from the minimum spanning tree property that if there are no points to be broken between regions, there must be two nodes with degree 1 in the region as the starting point and the ending point, while the degree of other nodes must be 2.
Through the above description and operation steps of the algorithm, the detailed flow chart is obtained as follows.
Step 1: Partition all nodes;
Step 2: Solve the minimum spanning tree of each area;
Step 3: Is it a node with degree greater than 2 in the region? If it is No, goto step 6;
Step 4: Delete the redundant segments of the node so that the degree is 0 or 1;
Step 5: The connection degree is 0 or 1 node, and the region start and end degree is 1;
Step 6: Connect the starting point and ending point of each region;
Step 7: The Hamiltonian cycle is obtained;

4. Algorithm demonstration and comparison
The demonstration of the algorithm takes rand50 in TSPLIB data set as an example to implement step 1 and step 2. The results are shown in Fig. 1. Step 3 is performed on each area to make all the points in Figure 1 have the degree of 0, 1 or 2. The result is shown in Fig. 2.

Finally, step 4 and step 5 are performed. Note that equation (1) needs to be satisfied during step 4, and $\sum q_{ij}$ is minimized during step 5. Finally, the degree of 50 nodes is 2. Fig. 3 is Hamilton cycle.

After calculation, the solution obtained by using this heuristic algorithm is 6127, and the optimal solution given by TSPLIB database is 5553. If the optimal solution is standard 100%, the mature method is used. The comparison is shown in Table 1.

The above experiments show that this heuristic algorithm has certain advantages in solving TSP problems, and it needs to be solved by comparing the distance between cities in China to analyze the results. At present, the linear distance between 20 cities, such as Beijing, Shanghai and Chengdu, is not listed here because of the large amount of data. If the enumeration method is used to solve the problem, the optimal solution can only be obtained by about $20! \approx 1.126 \times 10^{17}$ operations.

Therefore, the above heuristic algorithm is used to solve the problem, and compared with the mature algorithm, as shown in Table 2 (the branch and bound method to get the relative optimal solution is 100%).

Table 1. Comparison of the accuracy of various known algorithms and this algorithm

| Solution method               | Solution result | Method comparison with optimal solution /% |
|------------------------------|-----------------|--------------------------------------------|
| heuristic algorithm          | 6125            | 109.25                                     |
| Nearest neighbor algorithm   | 6456            | 120.53                                     |
| Random insertion of convex hull | 6674          | 118.23                                     |
| Two way improvement method   | 6465            | 115.78                                     |
| Branch and bound method      | 5637            | 101.54                                     |
Table 2. Comparison of solving distance problems among 20 cities in China

| Solution method                        | Solution result | Method comparison with optimal solution /% |
|----------------------------------------|-----------------|---------------------------------------------|
| heuristic algorithm                    | 7539            | 109.25                                      |
| Nearest neighbor algorithm             | 7889            | 120.53                                      |
| Random insertion of convex hull        | 6674            | 118.23                                      |
| Two way improvement method             | 6465            | 115.78                                      |
| Branch and bound method                | 5637            | 101.54                                      |

It can be seen from table 2 that when solving practical problems, the heuristic algorithm can get a solution optimized by the nearest neighbor algorithm. It can be seen from the comparison between table 1 and table 2 that the branch and bound method in the accurate algorithm is not very effective in dealing with large-scale problems because of the large amount of computation, so it often solves approximate solutions in small-scale problems. When solving practical problems, the scientific region division is carried out by using the clusterdata clustering function in MATLAB, which avoids the unnecessary increase of the total distance caused by the round trip between the two regions, and can get a more satisfactory solution when solving large-scale problems. Therefore, this heuristic algorithm has reference significance in solving the actual traveling salesman problem.

5. Practical application of the algorithm

In the algorithm flow, step 1 decomposes a big problem into several small problems to solve. In the problem of agricultural machinery operation path optimization, it can be divided according to the plots to be operated in each village, township or county.

If the operation scale is too large to return to the farm machinery households within the same day, the actual area division can be carried out according to the operation schedule, that is, each area divided is the land that needs to be operated within one day; if the distribution of operation land is not obvious, it can be divided by clustering function. When each area is determined, in the case of determining the parcel spacing, the application of the above algorithm to optimize. Preliminary calculation shows that it can save about 10% of operation cost and 15% of operation time for agricultural machinery operation.

6. Conclusion

The algorithm of traveling salesman problem discussed in this paper can effectively improve the operation efficiency by using the method of partition dimension reduction combined with minimum spanning tree, and the solution can approach the optimal solution. More importantly, the algorithm can be used to solve the practical problem of agricultural machinery operation path optimization, so it can be used as a new algorithm for large-scale traveling salesman problem, and has reference significance for practical problems. If we further study the selection of the starting point and the end point of each region, the possible solution will be closer to the optimal solution and produce better results in solving practical problems.

Acknowledgments

The work described in this paper was fully supported by the scientific research project of Jiangxi Provincial Department of Education(No. GJJ180316).

Reference

[1] Su Lijie, Nie Liyong. Real traveling salesman problem [J]. Minicomputer system, 2005,26 (4): 655.
[2] Zhang Xiaoxia, Tong Jiewei, Liu Zhe. A hybrid path reconnection algorithm for traveling salesman problem [J]. Computer Engineering, 2012, 38 (12): 122.
[3] Rao Weizhen, Jin Chun, Huang Yingyi. Hybrid algorithm of nearest neighbor and difference for solving TSP problem [J]. Theory and practice of systems engineering, 2011, 31 (8): 14-19.
[4] Huang Hongxuan. Operations research: mathematical programming. Beijing: Tsinghua University Press, 2011: 208-214.

[5] Ma Liang. Survey of algorithms for traveling salesman problem [J]. Practice and understanding of mathematics, 2000, 30 (2): 158.

[6] HAMEED IA,BOCHTIS D D, SORENSEN C G, et al. Automated generation of guidance lines for operational field planning [J]. Journal of Biosystems Engineering, 2010, 107 (4): 294-306.

[7] JIN J, TANG L. Coverage path planning on three-dimensional terrain for arable farming [J]. Journal of Field Robotics, 2011, 28 (3): 424-440.