The Minimum Steiner Tree Problem Based on Genetic Algorithm

Zhi-hao CHEN1,*, Wei-gen HOU2 and Yun DONG3

1Dept. of Basic Courses, Jiangsu Vocational College of Agriculture and Forestry, Jurong, China
2Dept. of College of Mathematical Science and Engineering, Anhui University of Technology, Maanshan, China
3Dept. of Teacher Education, Maanshan Teachers College, Maanshan, China

*Corresponding author

Keywords: Genetic algorithm, Minimum spanning tree, Minimum Steiner tree, Communication network.

Abstract. Genetic Algorithm is a method of searching the optimal solution by using computer to simulate Darwin's biological evolution theory based on natural selection and genetic principles of biological evolution process. This paper aims to solve the problem of the minimum Steiner tree by using genetic algorithm. First, it introduced the concept of genetic algorithm, minimum spanning tree and minimum Steiner tree briefly. Then, it described the application of genetic algorithm in the problem of minimum Steiner tree in detail and gave one method. Finally, this paper has attained the whole approximate root through solving one emulated experimentation emulated about communication network. The solution given in this paper is also applicable to other problems. Its versatility is better. Because of the narrowing of the search range, the search speed is faster and the number of iterations is less.

Introduction

A I (Artificial Intelligence) is known as one of the three most advanced technologies in the world. Evolutionary computation is one of these methods, It includes four typical computing methods: genetic algorithm (GA), genetic programming, evolutionary strategy and evolutionary programming. GA is relatively mature, and is a widely used method at present. It is a computational model that simulates Darwin's natural selection, natural elimination, and survival of the fittest. Genetic algorithms are usually implemented by computer simulation of genetic processes, and often used to search for optimization problems in computational mathematics.[1]

The Minimum Spanning Tree and Minimum Steiner Tree

In a weighted connected graph \( G = (V, E) \), if the subgraph \( G' \) contains all the vertices and some edges in the \( G \) and does not form a loop, then the \( G' \) is called the spanning tree of the graph \( G \). Among them, the spanning tree with minimal weight is called the minimum spanning tree. For solving the minimum spanning tree algorithm is the prim algorithm and Kruskal algorithm which is both the most famous and the most classic algorithm. The latter will be used in this article. If some points can be added in addition to the known vertices, a number of minimum spanning trees can be obtained. And the tree with the smallest total weight in all minimum spanning trees is the minimum steiner tree. The added point is called the steiner point.[2]

**Proposition 1** For the known points \( v_1, v_2, \ldots, v_n \), the minimum steiner spanning tree always exists.[3]

**Proposition 2** For the known points \( v_1, v_2, \ldots, v_n \), assuming that the number of steiner points in the minimum steiner tree \( T^* \) is \( s \), then \( s \leq n - 2 \).[4]
Proposition 3 The ratio of the weights of the minimum steiner tree $T^*$ and the minimum spanning tree $T_0$ is called the steiner ratio, and the steiner ratio is satisfied $\frac{\sqrt{2}}{2} \leq \frac{o(T^*)}{o(T_0)} \leq 1$.\[5]\]

Proposition 4 The points $P(x_1, y_1), P_2(x_2, y_2), \ldots, P_n(x_n, y_n)$ on the plane is known, assuming $x_u = \max\{x_1, x_2, \ldots, x_n\}, x_d = \min\{x_1, x_2, \ldots, x_n\},$

$y_u = \max\{y_1, y_2, \ldots, y_n\}, y_d = \min\{y_1, y_2, \ldots, y_n\},$

then all steiner points are certainly included in the point set $\{(x, y) | x_d \leq x \leq x_u, y_d \leq y \leq y_u\}$.\[6]\]

These four propositions indicate: (1) the minimum spanning tree is existent; (2) the number of steiner points is not more than the number of the known points two of the fixed point number minus two; (3) compared with the minimum spanning tree, the route designed with the minimum steiner tree can be reduced by about 13.4%; (4) the range of coordinate values of steiner points.\[7]\]

The Application of Genetic Algorithm in the Minimum Steiner Tree

The Problem of Wired Communication Network

The location of the 16 communication stations on the same level is known. The problem is laying optical cables between these sites and making the least cost. Each location is a vertex of the graph $G = (V, E)$. The set of vertices is $V = \{P_1, P_2, \ldots, P_{16}\}$. The set of edges is $E = \{P_i P_j, i = 1, 2, \ldots, 16, j = 1, 2, \ldots, 16, i \neq j\}$. The linear distance between each place is equal to the weight of the edge.

<table>
<thead>
<tr>
<th>place</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>coordinate</td>
<td>(392, 492)</td>
<td>(248, 376)</td>
<td>(712, 328)</td>
<td>(776, 240)</td>
</tr>
<tr>
<td>place</td>
<td>$P_5$</td>
<td>$P_6$</td>
<td>$P_7$</td>
<td>$P_8$</td>
</tr>
<tr>
<td>coordinate</td>
<td>(88, 240)</td>
<td>(440, 204)</td>
<td>(280, 44)</td>
<td>(520, -44)</td>
</tr>
<tr>
<td>place</td>
<td>$P_9$</td>
<td>$P_{10}$</td>
<td>$P_{11}$</td>
<td>$P_{12}$</td>
</tr>
<tr>
<td>coordinate</td>
<td>(350, 435)</td>
<td>(200, 400)</td>
<td>(700, 300)</td>
<td>(750, 200)</td>
</tr>
<tr>
<td>place</td>
<td>$P_{13}$</td>
<td>$P_{14}$</td>
<td>$P_{15}$</td>
<td>$P_{16}$</td>
</tr>
<tr>
<td>coordinate</td>
<td>(50, 280)</td>
<td>(340, 200)</td>
<td>(250, 84)</td>
<td>(530, -24)</td>
</tr>
</tbody>
</table>

Because the distance between the two locations is directly proportional to the cost of laying the optical cable, the problem to solve is how to design the laying cable line, making the network connected and the total length of the shortest route. In order to save the cost of laying cables, it is allowed to add $s$ new sites $v_1^*, v_2^*, \ldots, v_s^*$ outside the known locations, so as to get the smallest weight tree $T^*$ in the graph $G_{p_s}$, which is the minimum steiner tree problem in practical application.

The Design of Algorithm

Genetic algorithm is an approximate method for solving the minimum steiner tree problem. It consists of 4 main elements: chromosome coding method, individual fitness evaluation, genetic operators and operational factors. In this paper, the genetic algorithm is used to solve the minimum steiner tree problem. The design of the components is as follows:

(1) Encoding and decoding

Here is a binary encoding, which is a string consisting of 0 and 1 of all the individuals, every individual corresponding to a group of unknown points. In order to limit the search scope to the
rectangle box, we need to shrink the rectangle area to the square of the unit area with \([0,1] \times [0,1]\) size, and adjust the corresponding points at the same time.[8]

The binary encoding structure of each individual is as follows:

\[
\begin{array}{cccccccc}
* & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * \\
\end{array}
\]

The first point to be solved

The second point to be solved

The \(d_i\) point to be solved

here \(*\) is 0 or 1.

Each individual is a binary string that is randomly generated, and the length is \(w_i = d_s \times p \cdot d_y\) is the number of Steiner points to be solved, set by yourself when running the program. \(P\) is a binary encoding corresponding to a length of the point to be solved. If it is required to be exact to \(10^{-k}\), then \(p = d_s + d_y\), in which \(d_s = \left\lceil \log_2 (10^k \times c_s) \right\rceil, d_y = \left\lceil \log_2 (10^k \times c_y) \right\rceil\).

For each individual's decoding, we first need to segment the individual's coding according to the individual structure (such as the upper figure), and then convert every coding to decimal digit.

\[
\begin{array}{cccc}
(x_1',y_1'), & (x_2',y_2'), & \ldots, & (x_d',y_d') \\
\end{array}
\]

The first point to be solved

The second point to be solved

The \(d_i\) point to be solved

then convert the decimal numbers to the point coordinates \((x_1,y_1), (x_2,y_2), \ldots, (x_d,y_d)\), \(x_i = c_x \times x_i' + x_d, y_i = c_y \times y_i' + y_d, i = 1,2,\ldots,d_s\), in which \(x_i = x_i' = x_{d,s}, y_i = y_i' = y_d\).

In this way, the binary coded individual can be decoded to get all the coordinates of the points to be solved. [9]

(2) Selection of initial group

When the population size is \(M\), according to the required computational accuracy, \(M\) binary strings between 0 and 1 is randomly generated. Then the initial population is constructed. These \(M\) binary strings that do not repeat each other form a single individual. Each individual can decode the coordinates of the points to be solved according to the (1) method.

(3) Evaluation of seeds

For a spanning tree corresponding to an individual, the standard of judging its merits is to see the value of its target function. We take the minimum spanning tree length of the graph with the known points and the points to be solved as the vertex as the target function value. Of course, this value is the smaller the better.[10]

(4) Selection

In this paper, the two methods of wheel selection and optimal preservation strategy are combined to select the individual.

First, the wheel is designed according to the value of the objective function of the individual. Suppose that the objective functions of \(M\) individuals are \(\omega(T_0), \omega(T_1), \ldots, \omega(T_M)\). The maximum value is \(\omega(T^*) = \max \{\omega(T_0), \omega(T_1), \ldots, \omega(T_M)\}\). The difference value is \(\Delta \omega_i = \omega(T^*) - \omega(T_i), i = 1,2,\ldots,M\). Obviously, one of all \(\Delta \omega_i\) must be equal to 0.

Then the sum of the previous \(k\) values is calculated one by one \(S_k = \sum_{i=1}^{k} \Delta \omega_i, k = 1,2,\ldots,M\). Then calculate the endpoints of each interval \(p_k = \frac{S_k}{M}, k = 1,2,\ldots,M\). What needs to be pointed out is

\[
0 \leq p_1 \leq p_2 \leq \ldots \leq p_M = 1.
\]

These \(M\) endpoints divide the interval \([0,1]\) into \(M\) intervals \([0, p_1], [p_1, p_2], [p_2, p_3], \ldots, [p_{M-1}, p_M]\). The \(M\) interval corresponds to \(M\) individual and the value of
its target function $\omega(T_{m0})$. The length of the $M$ interval is $\Delta p_m = p_m - p_{m-1}, (i = 1, 2, ..., M)$. The length of the interval directly determines the size of the sector that he occupies on the wheel. The longer the interval, the greater the range it occupies on the wheel; the shorter the interval, the smaller the range it occupies on the wheel.

Then, the number of $M$ numbers $r$ is randomly generated on the interval $[0,1]$. If $r$ falls within the $M$ interval, the individual that corresponds to it will be retained. The smaller the objective function value is, the shorter the corresponding interval is, and the smaller the occupied area on the roulette, then the individual is not easy to be selected; otherwise, it is easier to be selected. After repeating this step for many times, individuals with higher fitness will be more likely to be selected and retained.

(5) Hybridization
This paper is a single point cross operation, and the intersection point is randomly generated, and the cross probability is $p_c = 0.03$. Two parent individuals exchange strings at random crossover positions to achieve cross operation. All the offspring produced by the cross operation are put into the group, and the parents who have crossed the operation are abandoned.[11]

(6) Mutation
In this paper, one point mutation operation is used, and the mutation probability is 0.03. The mutation location of the gene is generated randomly, and the number of genes on the gene to perform the mutation from 0 to 1 or from 1 to 0.[12]

Experimental Analysis and Conclusion
The above solution is programmed through Mathematica11, and runs through the windows7 system. For the 16 sites given above, we have tested several times that the best effect is when the number of Steiner points is $d_s = 3$. By using this method, ten times are solved randomly. And the algorithm can converge to the approximate optimal solution with an average iteration of about 10 times. The optimal approximate solution of the minimum Steiner tree is 1651.854, while the result of the minimum spanning tree is 1704.7. Compared with the minimum spanning tree, the total path length of the minimum Steiner tree is shortened by 3.1%. The Steiner point is:

$$(217.961, 289.078), (289.577, 206.626), (579.942, 100.617)$$

![Figure 1. The Minimum Steiner tree.](image)

Acknowledgement
This research was financially supported by the key project of the natural science research of Anhui Provincial Department of Education --“Research on the Limit Theory of Hidden Markov Chain and its application”.
References


