Research Article

An Improved Genetic Algorithm for the Large-Scale Rural Highway Network Layout

Changxi Ma, Cunrui Ma, Qing Ye, Ruichun He, and Jieyan Song

School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

Correspondence should be addressed to Changxi Ma; machangxi@mail.lzjtu.cn

For the layout problem of rural highway network, which is often characterized by a cluster of geographically dispersed nodes, neither the Prim algorithm nor the Kruskal algorithm can be readily applied, because the calculating speed and accuracy are by no means satisfactory. Rather than these two polynomial algorithms and the traditional genetic algorithm, this paper proposes an improved genetic algorithm. It encodes the minimum spanning trees of large-scale rural highway network layout with Prufer array, a method which can reduce the length of chromosome; it decodes Prufer array by using an efficient algorithm with time complexity $o(n)$ and adopting the single transposition method and orthoposition exchange method, substitutes for traditional crossover and mutation operations, which can effectively overcome the prematurity of genetic algorithm. Computer simulation tests and case study confirm that the improved genetic algorithm is better than the traditional one.

1. Introduction

Nowadays not only in urban road network have highways played an important role, but also they have become an indispensable part in rural road network. They bring geographically dispersed townships and villages together and thus render local residents better able to enjoy their life by time-saving travels. Moreover, rural highway network can be beneficial to the trade and logistics of agricultural produce, which will undoubtedly contribute to the development of regional economy [1, 2].

Over the past decades, researchers have paid great attention to layout problem of the rural highway network. Catharinus proposed the principle of sustainable land use planning by creating a development space for rural highway network layout [3]. In order to meet the demand in the future, it is necessary to perfect the rural highway network. Meanwhile, based on the adverse effects of the network which conflicted with sustainable principle, a new layout approach is urgently needed to meet the requirements on accessibility as well as environmental sustainability. Therefore, the spatial concept of the traffic-calmed rural area was put forward to solve

this dilemma by Catharinus. Through analyzing the layout network optimization design method of county and village, Li et al. asserted that network circuitous rates and accessibility between nodes should be firstly taken as the basic goal when the initial program of the rural highway network layout was determined [4]. Under the constraints of reasonable scale of development in the county and village highways, the contribution rate would be added to original large network, and the optimization of the network layout would finally be obtained. Zhou and Sun introduced a rural highway network layout model by using the 0-1 integer programming [5]. Wang et al. combined the node importance degree and area location into a new system and proposed the dynamic network layout theory [6]. Guo et al. studied the real situation of rural highway network based on policies, experience and layout theories, and proposed a node layout method, which took the optimal road network as the objective and the accessibility of network nodes as the control indicators. This method provided a good reference for townships and villages highway network layout [7]. Singh viewed a good highway system as the most important factor in infrastructures in rural areas, which would improve the accessibility of rural areas and

contribute to rural development in general, and he put forward a network layout method based on accessibility and GIS technology [8]. Scapparra and Church established a model to design the rural highway network by maximizing the highway connectivity between villages and the transportation efficiency [9].

During the process of road network layout, the key step is to find the minimum spanning tree for the whole network, which provides a scientific basis for the road network layout [10–13]. Therefore, how to quickly and effectively obtain the minimum spanning tree of the road network becomes the most important task.

As for unconstrained minimum spanning tree problems, there exist effective polynomial algorithms such as Prim algorithm and Kruskal algorithm. However, in practice most constrained minimum spanning tree problems are the intractable NP-complete problems, which means that polynomial time solution does not exist currently; therefore, the genetic algorithm seems to be a good strategy.

How to encode the feasible solution and how to design the genetic operators are the two main aspects and key steps needed to be considered when genetic algorithm is constructed. So to solve the general constrained minimum spanning tree problem by utilizing genetic algorithm, the key problem is how to encode the tree, deal with the degree of constraint, generate the initial population, and design genetic operators to avoid the premature convergence phenomenon [14, 15]. As for the minimum spanning tree problem, on one hand, if the graph is a dense graph or complete graph with many nodes, encoding according to edges will increase the length of chromosome, which will further greatly reduce the decoding efficiency; on the other hand, to transform degree constrained problems to the problems without degree constraint by punishing the tree not meeting the degree constraint will inevitably produce a large number of infeasible solutions in the process of randomly generating initial population as well as crossover and mutation operations, and because of the use of crossover operator, the premature convergence is apt to appear. Therefore, this paper adopts Prufer array to encode the spanning tree and utilizes the new designed single transposition operator and reverse operator (instead of standard crossover operator and mutation operator), which greatly reduces the chance to generate infeasible solutions and improves the efficiency of the Genetic Algorithm; furthermore, the offspring individual derives from only one parent ensured isolation between individuals. Without exchanging information with each other, the population diversity is effectively maintained to avoid the premature convergence phenomenon.

The rest of the paper is organized as follows. Section 2 mathematically formulates the layout problem of rural highway network, followed by the chromosome encoding method in Section 3 and chromosome decoding method and the fitness function in Section 4. The genetic operators are designed in Section 5. Computer simulation tests and a case study about Longguan village are proposed in Sections 6 and 7, respectively. Section 8, the final part, presents our conclusion.

2. Mathematical Model

Let $G = (X, E, W)$ be a weighted undirected graph, where $X = \{1, 2, \ldots, k\}$ is the set of vertices, E is the set of edges, and $W = [W_1, W_2, W_3]$ is the weight vector. Here $W_1 = (w_{ij}^1)_{k \times k}$ specifically refers to distances, where $w_{ij}^1 = w_{ji}^1$ and $w_{ii}^1 = 0$; W_2 and W_3 refer to time and cost, respectively, each of which has the same properties as W_1. We further denote by d_i ($i = 1, 2, \ldots, n$) the constrained degree of each vertex. Thus based on the minimum spanning tree, the layout problem of rural highway network can be formulated as follows:

$$\min \quad Z = \left[\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}^1 x_{ij}, \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}^2 x_{ij}, \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}^3 x_{ij} \right] \quad (1)$$

$$\text{s.t.} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = n - 1 \quad (2)$$

$$\sum_{i,j \in S} x_{ij} \leq S - 1, \quad \forall S \subset X, \ S \neq \phi \quad (3)$$

$$\sum_{j=1}^{n} x_{ij} \leq d_i, \quad i \in X \quad (4)$$

$$x_{ij} \in \{0, 1\} \quad i, j \in X, \quad (5)$$

where constraints (2) and (3) ensure that the solution should be at least a spanning tree, whereas constraint (4) shows that it is degree-constrained. Constraint (5) defines x_{ij} as the decision variable where $x_{ij} = 1$ if the edge (i, j) is in the optimal tree and $x_{ij} = 0$ otherwise.

3. Chromosome Encoding Method

Generally speaking, there are three broad categories of encoding methods: edge, vertex, and edge-vertex encoding. Here, the vertex encoding method is employed. According to the Cayley theorem, there exist exactly n^{n-2} distinct trees on n vertices. Prufer gave the simple proof of Cayley's formula by establishing a one-to-one correspondence between the set of spanning trees and $n - 2$ digit integer number with each integer between 1 and n. The $n - 2$ digit number later is known as Prufer number.

Here comes the Prufer number encoding procedure: find the leaf node with the smallest label in a labeled tree; then, add the node linking the leaf node to the Prufer array and delete this leaf node. Repeat the above steps until there are two nodes left; here, we obtain the Prufer array which encodes the labeled tree. From the above encoding rules, we know that Prufer array explicitly conveys the information about degree of every node. In practical program, chromosome (i.e., Prufer array) is randomly generated according to node number and label. If there are N nodes, the chromosome encoding will be formed by $N - 2$ natural numbers between 1 and N.

4. Chromosome Decoding and the Fitness Function

Any Prufer array uniquely corresponds to a labeled tree. Firstly, label all the nodes as the reserved nodes; then, orderly scan the number in the Prufer array. When we scan node u, the kth number, which means there exists a leaf node linking node u, and we delete u, and then find node v which has the smallest label among the reserved nodes and would not occur behind position k, link node u to node v and delete node v. Repeat the steps until the last two nodes which are reserved and connect these two nodes into an edge; then, a graph with $n - 1$ edges corresponding to a tree is formed.

The time complexity of the decoding algorithm based on the Prufer array is $O(n \log n)$, and to improve the decoding efficiency of the Prufer array, it is necessary to raise the efficiency in searching for the leaf nodes, as when the leaf nodes are without backtracking in the search, the decoding efficiency of Prufer array can be improved. Here a time-complexity of $O(n)$ of decoding algorithm based on Prufer array is given.

If there exist N nodes, we define Prufer $[N - 1]$ to store Prufer array, Degree $[N + 1]$ to store the degrees of node, and edge $[N]$ $[2]$ to store the edges of the labeled tree. Here the degree of leaf node is defined as zero and stored from the second position in Degree $[N + 1]$. After adding the current edge (u, v) to edge set defined as edge, we have Prufer $[u]$ and Degree $[v]$ minus one, and as a result Degree $[v]$ becomes -1, which has no influence on next searching for the node with the smallest degree. Only when Degree $[u]$ is valued as zero and u is smaller than v, this will cause backtracking in the searching for the node with the smallest degree. As long as u in the vacant array element Degree is stored $[0]$, we can avoid causing backtracking in searching for the node with the smallest degree by making a simple judgment. Obviously, the decoding algorithm based on Prufer array produces $O(n)$ time complexity.

Because the objective function of the model is always positive, the individual fitness function can be set as $f(x)$. Consider

$$f(x) = \begin{cases} M - F(x), & \text{if } F(x) < M, \\ 0, & \text{if } F(x) \geq M, \end{cases} \quad (6)$$

where $F(x) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^1 x_{ij} + \beta \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^2 x_{ij} + \gamma \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^3 x_{ij}$, $\alpha + \beta + \gamma = 1$, and M is a large number.

5. Genetic Operators Designing

The traditional genetic algorithm produces new offspring chromosomes by means of crossover and mutation operations. Yet as for the minimum spanning tree problem, crossover and mutation operators tend to produce enormous infeasible individuals [16, 17]. Furthermore, crossover operator is more likely to induce premature convergence. In order to improve the efficiency of the genetic algorithm searching minimum spanning trees, we design new single transposition operator and reversal operator instead of the standard genetic algorithm crossover and mutation operator combined with
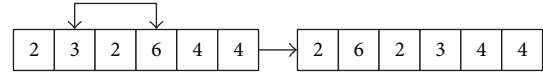

Figure 1: A typical single transposition operation.

characteristics of the spanning tree graph theory. In the application of the new algorithm, each offspring is derived from only one parent in the offspring generation process. New individuals inherit different traits from the single parent through the operations with a certain probability such as transposition, reversing, and mutation which can not only keep offspring individuals feasible, but also improve the search capability in the solution space.

5.1. Selection Operator. Each individual fitness percentage of the population is expressed in a pie chart which is similar to a gambling wheel. Each chromosome corresponds to a small piece of the pie. Area of the piece is proportionate to chromosome fitness; that is, the larger the fitness is, the greater the area occupies. To select a chromosome means to rotate the wheel and throw a ball into it until it stops; then, select the chromosome corresponding to where the ball stops. In practice, chromosomes will first be sorted based on objective value; then, the wheel will be divided according to population size, and its fitness will be labeled by a certain distribution; at last number between 0 and 1 is generated randomly, and the selective probability can be determined based on the random number. Besides, we use elite rule that will reserve best individual of the previous generation into the next generation. Thus, the best individual of the next generation certainly will not be worse than the previous generation.

5.2. Crossover Operator. We adopt single transposition method to replace the traditional crossover operation. The specific method is reversing the parental chromosome gene (i.e., Prufer array) according to two positions randomly selected. Figure 1 shows a typical single transposition operation. Obviously such operation does not produce infeasible individuals.

5.3. Mutation Operator. The orthoposition exchange method is adopted to implement mutation operation. Firstly, we randomly select two different gene positions in the chromosome and determine the beginning and ending positions of orthoposition exchange segment based on the order of the two gene positions. Then, if the gene between the beginning and ending position number is even, exchange every odd number gene in orthoposition exchange part with its right neighbour, the even number gene, and also if the gene number is odd, exchange every odd number gene except the last odd number gene in orthoposition exchange segment with its right even number gene. Figure 2 shows the typical orthoposition exchange operation. Obviously such operation does not produce infeasible individuals.

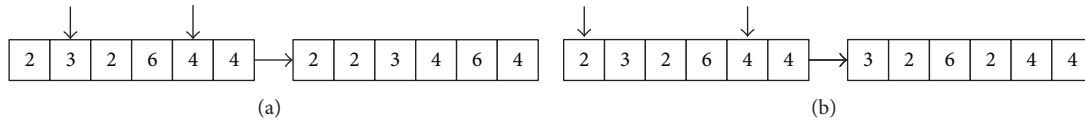The basic procedure of the improved genetic algorithm is showed in Figure 3.

| 2 | 3 | 2 | 6 | 4 | 4 | → | 2 | 2 | 3 | 4 | 6 | 4 | | 2 | 3 | 2 | 6 | 4 | 4 | → | 3 | 2 | 6 | 2 | 4 | 4 |

(a)   (b)

FIGURE 2: The orthoposition exchange method.
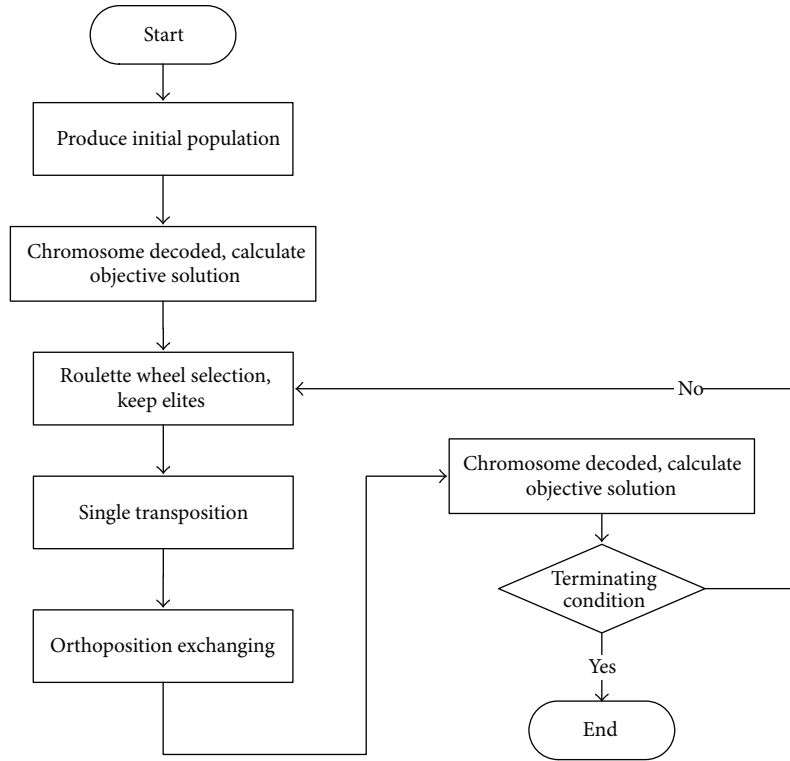

FIGURE 3: The flow chart of improved genetic algorithm.

6. A Computer Simulation Test

In this section, we perform a computer simulation test to evaluate the improved genetic algorithm proposed above. As shown in Table 1, a weighted matrix of nine nodes is randomly generated by computer. The key step for this rural highways network planning is to find out the minimum spanning tree.

The simulation test has been implemented on a Pentium Dual-Core 2.0 GHz PC, running on Windows XP and in Visual C++6.0 programming language. Here, we set the population size as 30, maximal generation 100, node number 9, crossover rate 0.8, and mutation rate 0.1, $\alpha = \beta = \gamma = 1/3$. When degree of each node is less than 3, by decoding the Prufer array we can obtain the corresponding edges as follows: $[(6, 4) (1, 6) (3, 1) (5, 3) (2, 7) (8, 2) (5, 8) (9, 5)]$, with the objective value 1744.

Figure 4 shows the evolutionary progress of the improved genetic algorithm. From Figure 4, we can see that there is a good convergence for the improved genetic algorithm. Table 2 compares the simulation results obtained using the improved genetic algorithm and the traditional one. From Table 2, it is safe to conclude that the improved genetic
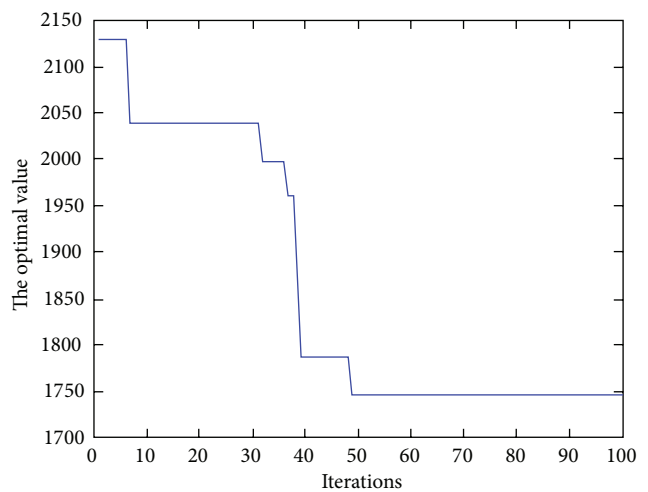

FIGURE 4: The evolution curve.

algorithm is better than the traditional one, especially for the large-scale minimum spanning tree problem.

TABLE 1: A weighted matrix of 9 nodes randomly generated by computer.

Nodes	1	2	3	4	5	6	7	8	9
1	0	374	105	644	886	645	941	859	588
2	374	0	889	567	828	269	372	142	736
3	105	889	0	277	431	995	608	993	955
4	644	567	277	0	495	352	653	470	618
5	886	828	431	495	0	423	432	305	261
6	645	269	995	352	423	0	196	332	306
7	941	372	608	653	432	196	0	124	472
8	859	142	993	470	305	332	124	0	965
9	588	736	955	618	261	306	472	965	0

TABLE 2: A comparison of the simulation results.

Algorithm	Optimization result	The objective value/m	Computational time/s
Traditional GA	[(6, 4) (1, 6) (3, 1) (5, 3) (2, 7) (8, 2) (5, 8) (9, 5)]	1744	68
Improved GA	[(6, 4) (1, 6) (3, 1) (5, 3) (2, 7) (8, 2) (5, 8) (9, 5)]	1744	23

7. Case Study

Yinzhou district of Ningbo city is located at the eastern coast of Zhejiang province, which is in close proximity to Ningbo-Shanghai Bridge. Yinzhou district has become one of the economic prosperity counties in China. The economic and social development goals of Yinzhou district are to grow as a modern county with economic prosperity, high-performance social management and service system, and coordinated urban and rural ecological environment through vigorously developing the processing industry and the tertiary industry. Highway network density of Yinzhou district is relatively large, and the highway service level is also appropriate. New processing industrial development zone obviously starting with a large- scale, strong development has a tremendous impact on Yinzhou urbanization. Highway traffic must promote and adapt to this new change. Based on the above analysis, highway network layout goals are determined as follows: adjust to the local skeleton network, strengthen the construction of radiation road network, and form rapid transport channel connecting southern coastal cities and East China, Central China, and Southwest China. Meanwhile, a traffic circle within half-hour from centre region or central area of Ningbo city is required. The final goad is to build a convenient, safe and economical highway network.

On the basis of the above skeleton highway network of rapid urbanization region, as to other regions of Yinzhou district, we select towns (planned as regional arterial road), central villages (assistant line for networked roads), and so forth. As real nodes are up to 55 and as virtual nodes are up to 13 for assisting external traffic hub, calculate each node importance degree, apply cluster analysis to layer, utilize graph theory to find the optimal tree to form a network supporting tree, and deploy the layout combined with main zone bit line direction, in which the lines connecting the zone's trunk line with exports of the outside have the priority to be considered.
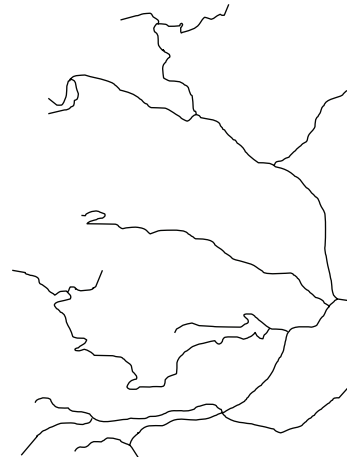

FIGURE 5: Rural highway network layout of Longguan village, Yinzhou District, China.

Yinzhou district is located at the east of Xiang hill, north of two major harbors. We select the shortest planned Baozhan road and coast line which are better connecting to eastern coast to connect S71 and Mingzhou road to build the transportation corridor from development zone and centre area of Yinzhou to eastern coast so as to take advantage of port location [18].

Using the genetic algorithm proposed in this paper, we can obtain the optimization layout as shown in Figure 5.

8. Conclusion

Rural highways, regarded as an important social infrastructure, play an important role in the promotion of the social and economic development in rural areas. Traditional Prim algorithm and Kruskal algorithm have not been suitable for the large-scale rural highway network layout. In order to

solve the minimum spanning tree problem of rural highway network layout, this paper establishes the improved genetic algorithm which encodes the minimum spanning trees with Prufer array reducing the length of chromosome; besides, the proposed algorithm of time complexity is utilized to decode Prufer array, and it adopts single transposition method and orthoposition exchange method instead of the traditional crossover and mutation operators which effectively avoid prematurity. Computer simulation test and the case study confirm that the improved genetic algorithm can obtain the minimum spanning tree of the rural highway network layout quickly and effectively; even more, this algorithm can be applied in different kinds of degree constrained minimum spanning tree problems to get group or subminimum trees with a high probability in a short time, which compensates for the disadvantage of greedy algorithm being able to find one minimum spanning tree and provides more options for practical engineering program evaluation and decision-making.

Highway network layout is a very complex task; finding the minimum spanning tree is just one important aspect before designing the network, and after that the initial highway network layout program still cannot be obtained. To achieve transition from the tree to the planned road network, a deep investigating combined with the actual situation of the layout area is also needed.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] P. Xin, H. Z. Zhang, S. X. Zhao, and C. N. Wang, "Research on developing strategy of village highway of Gansu Pronvince," *Highway*, vol. 23, no. 9, pp. 76–80, 2004.

[2] K. Jiang and S. R. Tu, "Study on the planning method for mountain areas highway network in western regions," *Communications Standardization*, vol. 33, no. 7, pp. 211–214, 2006.

[3] C. F. Jaarsma, "Approaches for the planning of rural road networks according to sustainable land use planning," *Landscape and Urban Planning*, vol. 39, no. 1, pp. 47–54, 1997.

[4] X. H. Li, H. Liu, and S. F. Yang, "Method of network optimizing design in the township highway planning," *Journal of Southeast University*, vol. 27, no. 11, pp. 156–160, 1997.

[5] M. Zhou and S. D. Sun, *Genetic Algorithm Theory and Applicantion*, National Defence Industry Press of China, 1998.

[6] Y. Q. Wang, S. H. Cui, and S. B. Hao, "Study of the theory and methods for highway network in countryside," *China Journal of Highway and Transport*, vol. 10, no. 4, pp. 86–89, 2002.

[7] X. C. Guo, B. Hu, and F. J. Chen, "Study on rural highway network planning," *Journal of Highway and Transportation Research and Development*, vol. 19, no. 2, pp. 92–95, 2002.

[8] A. K. Singh, "GIS based rural highway network planning for developing countries," *Journal of Transportation Engineering*, vol. 12, no. 8, pp. 212–219, 2010.

[9] M. P. Scaparra and R. L. Church, "A GRASP and path relinking heuristic for rural road network development," *Journal of Heuristics*, vol. 11, no. 1, pp. 89–108, 2005.

[10] G. C. Yu and W. Zhou, "Road network planning optimization theory and methods," *Journal of Xi'an Highway University*, vol. 7, no. 3, pp. 45–49, 1998.

[11] C. F. Yu, "The position and function of the countryside highway in new period," *Jilin Science & Technology of Communications*, vol. 19, no. 3, pp. 69–71, 2007.

[12] Z. G. Yao, Y. Q. Wang, and W. Zhou, "The strategic thinking of the countryside highway in western china," *Comprehensive Transportation*, vol. 21, no. 3, pp. 18–20, 2005.

[13] F. Tao and J. J. Wang, "The function and significance of the countryside highway," *Transportation World*, vol. 12, no. 10, pp. 29–30, 2006.

[14] K. C. Lu and H. M. Lu, *Graph Theory and Application*, Tsinghua University Press, 1995.

[15] H.-Z. Song, "Partheno-genetic algorithm for solving the degree-constrained minimum spanning tree problem," *System Engineering Theory and Practice*, vol. 25, no. 4, pp. 61–66, 2005.

[16] G. R. Raidl and B. A. Julstrom, "Edge sets: an effective evolutionary coding of spanning trees," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 3, pp. 225–239, 2003.

[17] B. A. Julstrom, "Codings and operators in two genetic algorithms for the leaf-constrained minimum spanning tree problem," *International Journal of Applied Mathematics and Computer Science*, vol. 14, no. 3, pp. 385–396, 2004.

[18] Y. Q. Wang and W. Z. Yinzhou, *District Highway Network Planning*, Guangdong Province People Press, 2003.