An Improved genetic algorithm for solving TSP

Shuaihan Tao*

School of management, Tianjin University of Technology, Tianjin, China

*Corresponding author: shuaihantao@tjut.edu.cn

Abstract. An improved genetic algorithm is proposed to solve the traveling salesman problem. On the basis of traditional genetic algorithm, greedy algorithm is introduced to initialize the population. The genetic parameters are adjusted adaptively from two aspects of genetic evolution algebra and individual fitness function value, which can speed up the optimization speed and prevent the optimization from falling into local optimum. In mutation operation, adaptive algorithm is used to select mutation operator to improve the quality of variation and search effect of the algorithm; after individual evolution, one-way evolutionary reversal operation is introduced to improve the chance of inheriting parent and parent genes, and improve the ability of algorithm to search for optimal solution.

1. Introduction
Traveling salesman problem (TSP) is a well-known NP hard problem in combinatorial optimization. It has a wide range of engineering applications and real life background, such as printed circuit drilling, aircraft route arrangement, highway network construction. All these practical problems can be transformed into TSP problems to solve. Therefore, how to solve TSP problems quickly and effectively has high practical value, Gao Hai chang et al. [1] described ant colony algorithm, genetic algorithm, simulated annealing algorithm, tabu search, neural network, particle swarm optimization algorithm, immune algorithm, etc, In reference [2], a fast ant colony algorithm for TSP is proposed. In reference [3], the combination of ant colony algorithm and immune algorithm is applied to TSP. In this paper, the genetic algorithm of intelligent algorithm is analyzed, and its application in TSP is improved to enhance its optimization ability, The improved genetic algorithm can be more reliable in the case of small population size.

2. Traveling Salesman Problem
Traveling salesman problem (TSP) can be described as: given the distance between n cities, an existing salesman who starts from a certain city must visit these n cities, and each city can only visit once, In graph theory, suppose there is a graph \( G = (V, e) \), where \( V \) is the vertex set, \( e \) is the edge set, and \( d_{ij} \) is a distance matrix composed of the distance between vertex \( i \) and vertex \( j \). The traveling salesman problem is to find a circuit with the shortest distance that passes through all vertices and each vertex only passes once.

3. Genetic algorithm
Genetic algorithm (GA) is a computational model simulating the natural selection and genetic mechanism of Darwinian biological evolution theory. It is a method to search the optimal solution by
simulating the natural evolution process. The algorithm has the functions of self-organization, self-adaptation, self-learning and population evolution, Have a strong ability to solve problems, in many areas have been applied.

3.1. Basic operation process
The basic operation process of genetic algorithm is as follows:
(1) Initialization: set the evolutionary algebra counter \( t = 0 \), set the maximum evolutionary algebra \( T \), randomly generate \( m \) individuals as the initial group \( P(0) \).
(2) Individual evaluation: calculate the fitness of each individual in population \( P(T) \).
(3) Selection operation: the selection operator is applied to the population. The purpose of selection is to directly inherit the optimized individuals to the next generation, or to generate new individuals through pairing crossover and then to the next generation. The selection operation is based on the fitness evaluation of individuals in the population.
(4) Crossover operation: the crossover operator is applied to the population. Crossover refers to the operation of replacing and reorganizing the partial structure of two parent individuals to generate new individuals. Crossover operator plays a key role in genetic algorithm.
(5) Mutation operation: the mutation operator is applied to the population. That is to change the gene values on some loci of the individual string in the population. After selection, crossover and mutation operation, the population \( P(T) \) obtains the next generation population \( P(T+1) \).
(6) Termination condition judgment: if \( t = T \), the individual with the maximum fitness obtained in the evolution process is taken as the output of the optimal solution, and the calculation is terminated.

4. Improved genetic algorithm for solving TSP
The efficiency of standard genetic algorithm is not high, and it is easy to converge in the local optimal solution. In order to improve the search speed and efficiency, and better solve the TSP problem, the author improves the standard genetic algorithm.

4.1. Decision of end algebra in genetic algorithm

4.1.1. Information Entropy. The concept of entropy comes from thermodynamics, which reflects the degree of micro chaos of the system.
Definition 1 the mathematical model of single symbol discrete memoryless source is as follows:

\[
\begin{bmatrix}
  x_1 & x_2 & \ldots & x_i & \ldots & x_n \\
  p(x_1) & p(x_2) & \ldots & p(x_i) & \ldots & p(x_n)
\end{bmatrix}
\]  (1)

Where \( 0 \leq p(x_i) \leq 1 (i = 0, 1, 2, \ldots, n) \) and \( \sum_i p(x_i) = 1 \)

The mathematical expectation of self-information quantity of discrete messages (i.e. probability weighted statistical average) of each discrete message is the average information of the source, generally referred to as the information entropy of the source, also known as the source entropy and Shannon entropy. Sometimes it is called unconditional entropy or entropy function, or entropy function, or entropy, which is called \( H(x) \), and

\[
H(x) = -c \sum_i p(x_i) \log p(x_i)
\]  (2)
Where: C is a positive integer, generally c = 1

In the search process of genetic algorithm, each generation population can be regarded as a single symbol discrete memoryless information source. According to equation (1), the information entropy of the current generation can be calculated. Once the information entropy is determined, the corresponding information source will be determined [4-5]. Only under the action of mutation operator, the genetic algorithm will eventually converge to the most promising individual in the initial population. No matter what kind of distribution the initial population gives, its limit distribution is uniform under the repeated action of mutation operator. When the population information entropy is small, genetic algorithm does not have the ability of evolution, so it is difficult to find a better solution. Because of the existence of mutation operator, the population information entropy is 0, which has randomness.

4.1.2. Determination of cut-off algebra of genetic algorithm. The cut-off algebra of genetic algorithm can be judged when the information entropy reaches a certain minimum. If the population variance is large and the information entropy is large, the individuals in the population may be scattered in all corners of the search space, which is suitable for the initial population. If the population variance is large and the information entropy is small, the individuals in the population are scattered in a larger space, but the individual diversity is weak. The population variance is small and the information entropy is large, which indicates that the population individuals are concentrated in a small space, but the individual diversity is strong. After several generations of evolution, the population converges and the variance and information entropy become small. The cut-off algebra of genetic algorithm can be judged when the variance and information entropy are less than a certain threshold.

4.2. Operators for TSP

4.2.1. Coding method. Path representation: this method is logical and can be easily and naturally corresponding to the gene coding of the journey. If the initial gene is fixed, the coding method has uniqueness, otherwise, it is not unique. This method is most logical, passing through each node and does not generate sub loop [6]. This paper mainly uses path representation method to code.

4.2.2. Improved crossover operator. The design of crossover operators for traditional GA to solve TSP problems includes partial matching crossover (PM), sequential crossover (ox), oxlike crossover, CX, a sexual crossover etc. There are two disadvantages: gene repetition, the inability of the parent excellent gene offspring to retain, and the coding conditions need to be changed. The main purpose of genetic crossover is to inherit the excellent genes of the parent generation as much as possible. In order to keep the excellent genes of the parent, the greedy crossover operator can be used to modify the crossover operator. The basic idea of greedy crossover operator is: First, select the first city of the father, then compare the remaining cities in the two parents, select the city close to continue the city travel. If the city already appears in the travel, select the city of the other parent, if both cities have appeared, The greedy crossover operator makes full use of the local information of chromosome to guide genetic evolutionary search. The cyclic greedy intersection (GX) can be described in the following forms: suppose there are n cities 1, 2 …N; chromosome coding is represented by path, and the parents to be crossed are:

\[
\begin{align*}
\chi_1 &= (I_{11}, I_{12}, \ldots, I_{1n}) \\
\chi_2 &= (I_{21}, I_{22}, \ldots, I_{2n})
\end{align*}
\]
According to the greedy algorithm, the above chromosomes are regarded as a ring, and the crossover steps are designed as follows:

- Step 1 set the current city as R1 and add R1 to the offspring
- Step 2 compares the two parent individuals adjacent to R1 and selects the nearest child
- Step 3 if the expanded city exists in the sub-individuals, the expanded city can be determined as the next nearest city. Thus, if all adjacent cities exist in the sub-individuals, a city that has not been expanded can be randomly added to the sub-individuals
- Step 4 repeat step 2 and step 3 continuously, and the offspring can be generated completely.

4.2.3. Design mutation operator. When TSP is solved, any operator with local search ability is its mutation operator. Compared with crossover operator, it is more flexible to design mutation operator.

- Reverse mutation: randomly select two points in the string, and insert the substrings in the two points into the original position in reverse order.
- Insertion mutation: randomly select a code from the string and insert it into the middle of the randomly selected insertion point.
- Exchange mutation: randomly select two points in a string and exchange their values.
- Point mutation: mutation only mutates some bits of a string with a certain probability (usually very small)

5. Experimental analysis
Based on the above analysis, the improved genetic algorithm for TSP problem is designed and implemented by C# language. Several examples in TSPLIB test database are selected for simulation analysis. The experimental environment is: Windows XP system, 2G memory, Microsoft visual studio platform.

| method                          | Results after 0.2S |
|---------------------------------|-------------------|
| Tradition                       | 825.47            |
| 4.1                             | 508.89            |
| 4.1+4.2.1                       | 482.53            |
| 4.1+4.2.1+4.2.2                 | 474.94            |
| 4.1+4.2.1+4.2.2 +4.2.3          | 423.91            |

The experiment of Oliver 30 problem is carried out. In the same operation time (0.2 s), the results of traditional methods and the implementation of various improvement strategies are obtained respectively, as shown in Table 2. From table 2, the improved genetic algorithm can converge to the optimal solution in a short time.

6. Conclusion
In this paper, an improved scheme of genetic algorithm is proposed, which adds initialization heuristic information and improves greedy crossover operator. According to the diversity of the population and the distribution of individuals in the population, a method to determine the cut-off algebra of genetic algorithm is found; If heuristic information is introduced in group initialization, the accuracy and convergence of the problem can be improved; after the greedy crossover operator is improved, the excellent genes of the parents can be better retained, the feasibility of the offspring can be guaranteed, and the crossover operation times and optimization ability of the genetic algorithm can be improved.

References
[1] Gao H C, Feng B Q, Zhu L. Reviews of the meta-heuristic algorithm for TSP[J]. Control and Decision, 2006, 21(3):241-247.
[2] Ji J Z, Huang Z, Liu C N. A fast ant colony optimization algorithm for traveling salesman problems[J].J of Computer Research and Development, 2009, 46(6):968-978.
[3] Liu Z H, Zhang Y J, Zhang J, et al. Using combination of ant algorithm and immune algorithm to solve TSP[J]. Control and Decision, 2010, 25(5): 695-700.

[4] Mu Yanling, Li Xuewu, Gao Runquan, Genetic algorithm for parallel implementation of TSP problem[J]. Beijing Union University, 2008(4): 99-1021. (in Chinese)

[5] Yang Huafen, Wei Yan. Problem solving TSP improved genetic algorithm[J]. Chongqing Institute of Technology: Natural Science, 2009(7): 113-115. (in Chinese)

[6] Ao you Yun, Chi Hong Qin. An algorithm for TSP Based on genetic algorithm[J]. Computer and Digital Engineering, 2008 (5): 75-77.