

# Solving TSP Problems with Hybrid Estimation of Distribution Algorithms

Xiaoxia Zhang and Yunyong Ma

College of Software Engineering, University of Science and Technology Liaoning, China  
aszhangxx@163.com

**Abstract.** In this paper, a hybrid Estimation of Distribution Algorithms is proposed to solve traveling salesman problem, and a greedy algorithm is used to improve the quality of the initial population. It sets up a Bayes probabilistic model of the TSP. The roulette method is adopted to generate the new population. In order to prevent falling into local optimum, the mutation and limit were proposed to enhance the exploitation ability. At the same time, three new neighborhood search strategies and the second element optimization method are presented to enhance the ability of the local search. The simulation results and comparisons based on benchmarks validate the efficiency of the proposed algorithm.

**Keywords:** Estimation of distribution algorithm, Traveling salesman problem, Neighborhood search.

## 1 Introduction

Traveling Salesman Problem (TSP) is a classic route problem. The complexity of the problem grows exponentially with the expansion of TSP problem's scale, so TSP problem is a typical NP-Hard problem. In recent years, many intelligence computations have been used to solve the problem, such as ant colony algorithm [1], genetic algorithm [2], particle swarm optimization [3], artificial bees colony algorithm [4], estimation of distribution algorithms [5], etc.

Estimation of Distribution Algorithms (EDA) is developed from the basis of Genetic Algorithm (GA). As a relatively new paradigm, EDA is widely used in NP-Hard problems' research. According to its complexity and sampling method, it can be roughly divided into dependency-free, bivariate dependencies, and multivariate dependencies [6]. EDA is evolved from the genetic algorithm. The algorithm is different from genetic algorithm's complex operators such as the crossover and mutation. In the EDA, a probability model of the most promising area is built by statistical information based on the searching experience, and then the probability model is used for sampling to generate the new individuals. In such an iterative way, the optimal solution is obtained by evolving eventually. At present Estimation of Distribution Algorithm has been widely applied to industries, such as feature selection, engineering optimization [7], machine learning [8], flow-shop [9], multidimensional knapsack problem (MKP) [10], etc.

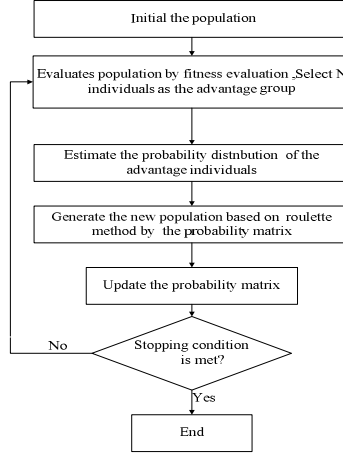
In the algorithms, greedy algorithm is used to obtain a relatively optimal solution, which enhances the quality of the initial solutions. In the literature [11], two elements optimization algorithm (2 - optimization, 2 - OPT) is introduced into the traveling salesman problem which enhances the capability of the search apparently. At the same time, three new neighborhood search strategy are proposed in this paper to improve the local search capability with little damage to the building blocks. In order to improve the global search ability of the algorithm, the negative feedback is introduced in this algorithm to control the change of the probability. The mutation operator is built into the proposed strategy to maintain population diversity and improve the ability of global exploration.

The advantage of this algorithm is put forward some new neighborhood search strategies which can improve the local search speed and balance the exploration and the exploitation abilities very well. The second element optimization method is also presented to enhance the quality of the local optimal solution. Negative feedback strategy and mutation operations are introduced to guarantee the global search ability of the algorithm, which can prevent the algorithm trapping in local optimal solution. And simulation results show that the algorithm has higher search capabilities, and the iterative algebra is significantly reduced.

## 2 Estimation of Distribution Algorithms

The notion of Estimation of Distribution Algorithms burnout was advocated in 1996, around 2000, the algorithm was developed rapidly. In recent years, EDA has been discussed as an important project in the world's famous academic conferences in the field of evolutionary computation, such as ACM SIGEVO, IEEE CEC [12]. EDA abandons the cross in the genetic algorithm (GA) such as restructuring operation. EDA builds the probability model for the next generation. The general framework of the EDA is illustrated in Fig. 1.

The critical step of the under procedure is to estimate the probability distribution [13]. The probability model used to describe the distribution of the solution space in each generation. The updating process reflects the evolutionary trend of the population. The probability distribution and updating process are both the most critical step in the Estimation of Distribution Algorithms. With respecting to several types of problem, a proper probability model and a suitable updating mechanism should be well developed to estimate the underlying probability distribution. However, more attention was paid to global exploration while the searching ability of Estimation of Distribution algorithm is relatively limited in the EDA. So, an effective EDA should balance the exploration and the exploitation abilities.



**Fig. 1.** The general framework of the EDA

## 2.1 Initial Population

For ease of description, this article uses decimal-strings encode. Every individual of the population is a solution of the TSP. Suppose that there are  $L$  cities, let  $\pi_i$  be a natural number,  $\Pi(\pi_1, \pi_2, \dots, \pi_L)$  be a city permutation. To ensure the diversity of initial population, we generate  $n$  city permutations as initial population randomly.

## 2.2 Advantage Group

According to the information about two cities' distance information, calculate the fitness of individuals in the population. Advantage group is the best of  $p$  individuals by the fitness.

## 2.3 Probability Model

Establish a frequency matrix  $k$ , let  $k_{i,j}$  be the frequency in the advantage group from city  $i$  to city  $j$ .

$$k_{i,j} = \frac{1}{L} \sum_{k=1}^L \delta(x_k) \quad (1)$$

The value of  $\delta(x_k)$  is decided by whether city  $i$  to city  $j$  is appeared in the route, if appeared  $\delta(x_k)$  is 1, otherwise 0. When the probability model is constructed, Literature 14 uses the machine learning Heb rule to optimization the probability model:

$$p_{i,j} = \frac{1-\partial}{L} + \partial * k_{i,j} \quad (2)$$

Because the initial group is generated randomly, the probability model of the initial value set to be  $1 / L$ ,  $\partial \in (0, 1)$ , and  $\partial$  is the learning efficiency. The bigger  $\partial$  is, the faster learning rate is, and the greater the parent's influence on children and vice versa.

## 2.4 Generate the New Population

Bayes' theorem [15] is an important theorem in statistical inference which belongs to the conditional probability problem. According to a probability model as a prior probability is constructed, we can estimate the left posterior distribution of the occurrence of the event. Bayesian has widely applied to solve many combinatory problems now, such as shop scheduling problem [16] etc. In this paper, it is used to solve the TSP problem.

According to the matrix of probability model, taking the roulette wheel to obtain new species, using Bayesian probability model for the next point of probability distribution:

$$p(B_i | A) = \frac{p(A | B_i) p(B_i)}{\sum_{j=1}^{j=L} p(A | B_j) p(B_j)} \quad (3)$$

$1 \leq i \leq L$ ,  $A$  represents the set of reached cities,  $B_i$  means the next to reached city is  $B_i$ . Roulette method is used to produce the next point until all cities are chosen, according to the probability distribution. The basic process can be expressed in the type [17]:

$$p(\pi_1, \pi_2, \dots, \pi_L) = p(\pi_1) * p(\pi_2 | \pi_1) * \dots * p(\pi_L | \pi_1, \pi_2, \dots, \pi_{L-1}) \quad (4)$$

Algorithms produce some routes as the new initial population based on the same way. In order to keep the search speed, increase the algorithm's global search ability, reduce the loss of the quality of individual, paper will retain the best individual in this generation, which Adds directly to the next generation of initial population and takes part in the evolution of next generation.

Determine the convergence of the new population. If the population is convergent, algorithm is end, otherwise proceeds to next step.

## 2.5 Update the Probability Model

When the populations evolve to  $t + 1$  generation, the probability model is as following:

$$p^{t+1}_{i,j} = (1-\partial) p^t_{i,j} + \partial * k_{i,j} \quad (5)$$

$p^t$  shows the population when the first generation  $t$  matrix, similarly  $p^{t+1}$  says the population in the first generation  $t + 1$ . Population will be optimized completely by updating the probability model constantly and iterative optimization.

### 3 Local Search Strategy

#### 3.1 Greedy Solution

In order to improve the quality of the initial population, in this paper, the greedy algorithm is applied to the generation of initial population. Generate a point randomly as a starting point and then choose the nearest city as the second, and so on until you walk all the cities. Add the routes generated by the greedy algorithm to the initial population.

#### 3.2 Neighborhood Search

In order to improve the local search ability of the algorithm, the paper puts forward three kinds of neighborhood search mode:

1. Because the new individual is produced by the conditional probability and every two connected cities' distance is not restricted with direction, therefore if we exchange the order of the adjacent two cities, there is small effect on groups .This article chooses one position and exchanges its next adjacent city randomly. such as  $p$  (1,4,5,3,2,6), the selected location is 3, the exchanged individuals are  $p$  (1,4,3,5,2,6).
2. Select a city randomly and places it between any two cities, such as  $p$  (1,4,5,3,2,6), the selected location is 3, then it will be inserted into randomly the rest route (1,2,4,5,6).
3. Regard two adjacent cities as a module. Randomly select two cities, respectively for strategy 2 operation .Such as  $p$  (1,4,5,3,2,6), the selected location 3, the city 5 and 3 are randomly inserted into the rest route(1,4,2,6).

All the three strategies take effect when the population evolution is to a certain degree and the best individual is not optimized in a certain generation. Frist and second strategies are used for all the population. Third strategy is used for the optimal individual. If the individual fitness is increased after three strategies operations, reserved the new individual, abandoned the original individual, otherwise unchanged.

In order to improve the local search speed, the algorithm also introduces 2-OPT algorithm. The general framework of 2-OPT is illustrated in Fig. 2.

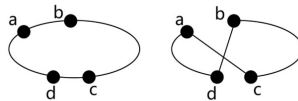


Fig. 2. 2-OPT operation

2 - OPT neighborhood search method is a kind of common neighborhood search heuristic method, which disconnected two cities' cables, and cross linking cities. If the fitness of the individual is greater than the original individual, retain new individual and abandon the original individual, otherwise unchanged.

## 4 Global Search Strategy

### 4.1 Mutation Operation

As the evolution of the population, the algorithm is likely to be trapped in local optimal solution, and can't get the global optimal solution. The essay introduces mutation operation to ensure the population diversity and to improve the global search ability of the algorithm. The basic idea of Mutation is to select a percentage of the individuals in the initial solution randomly and exchange the positions of the two cities randomly.

In the initial stages of the evolution, population diversity is very good, and almost doesn't need a mutation, so the mutation only in when the best individual is not optimized in a certain generation to take effect.

### 4.2 Negative Feedback Operation

With the evolution of the population, polarization phenomenon occurs in probability matrix. It is easy to lose the information which lead to reduce of the global search ability.

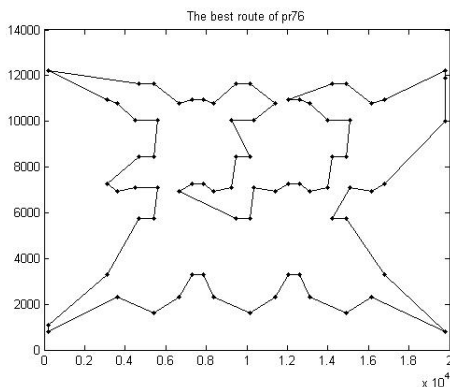
Compare the probability value with the learning rate, if two adjacent cities' probability is closes to  $\partial/L$ , reduces the learning rate, and by the same token, when one of the two adjacency cities probability is very small, increases the learning rate.

In order to keep the population diversity, this article also limits the scope of probability, and the probability controls within a certain range, so the probability is controlled under a certain scope.

## 5 Experiments

In order to verify the effectiveness of the algorithm, we code the procedure on Visual C++ 2012 on Lenovo G480 processor and use the well-known data sets TSPLIB for testing.

As a result of the algorithm running time limit, this paper set up an algorithm obtained that when the best solution provided by TSPLIB or run to maximum the algebra operation stop. All of the following tests' largest number of iterations is set to 5000 generations.  $\partial=0.15$ . The Minimum probability is  $0.0001*\partial/L$ . The size of initial population is 100. Advantage of population's size is 30. The largest probability is  $0.9999*\partial/L$ , if the optimal value is not evolution for 30generations, local search strategy and variation operation will be take effect.



**Fig. 3.** The best route of pr76

HEDA find the best route of pr76, which was shown in fig3.

Make gr24 as the test example, and the algorithm (HEDA) runs 20 times, the results was shown in table 1. In order to verify the effectiveness of the algorithm, the algorithm is compared with other algorithms in this paper, which include EHBSA algorithm, OX algorithm and eER algorithm from the literature 14; the test data of algorithm (MMIC) and improved algorithm (IMMIC) from the literature 17. The simulation results of the gr48 were shown in table 2. This algorithm is same to gr24 problem in Settings. Table 3 shows the results of the algorithm to solve the some TSP problems, and every case was run 20 times.

In the three tables, Scale means the scale of initial population, Times means the times of simulation run. Best means the best solution provided by TSPLIB. We ran the program 20 times, took the average and the best, Op is the best, Average is the average. Algebra means the number of iterations while Op is equal to Best. The pr136 case' best route is not found

**Table 1.** gr24 test results

Algorithm	Scale	Average	Times
EHBSA/WO	60	1281	10
OX	60	1345	10
eER	60	1299	10
MMIC	2000	1284	20
IMMIC	300	1273.3	20
HEDA	100	1272	20

**Table 2.** gr48 test results

Algorithm	Scale	Average	Times
EHBSA/WO	60	5212	10
OX	60	5527	10
eER	60	5653	10
MMIC	10000	5234.8	20
IMMIC	1000	5065.4	20
HEDA	100	5051	20

**Table 3.** The TSP test results

The case	Best	Op	Average	Algebra
Burm14	3323	3323	3324.4	100
Gr24	1272	1272	1272	190
Gr48	5046	5046	5051	466
Pr76	108159	108159	108880	1000
Pr136	96772	9700	9720	—

Through the test data, we can get the following conclusion:

- (i) Table 1 and table 2 show that the algorithm effective is very good in solving gr24 gr48 problem. The Average is better than other algorithms, which suggest that HEDA has the higher searching ability.
- (ii) The algebra to obtain best route can be seen from table 3. We can see that algorithm has a high search speed, less number of iterations. This suggests that the neighborhood strategy can be very good to speed up the search speed and improve the quality of evolution.
- (iii) According to the three table above, we can see that most of the testing example' best solution can be found and can maintain the stability. This suggests that the algorithm has high global search ability.

## Conclusions

In this paper, an effective estimation of distribution algorithm based on Bayes probability model is proposed for solving the TSP problem. For local exploitation, three effective Variable Neighborhood Search (VNS) were incorporated into EDA. For global exploration, we design mutation operation and negative feedback operation to improve the diversity of population. The simulation results and comparisons with some existing algorithms demonstrated the effectiveness of the HEDA. The future work is to design EDA based algorithm for the massive TSP problem.



## References

1. Xu, C., Chang, H.-Y., Xu, J.: Novel Ant Colony Optimization Algorithm with Estimation of Distribution. *Computer Science* 32 (2010)
2. Liang, Q.J., Shu, J., Fan, X.: TSP Modeling Method Based on Genetic Algorithm. *Computer Engineering* 37 (2011), Wang, D., Wu, X.-B., Mao, X.-C.: Improved Hybrid Particle Swarm Optimization Algorithm for Solving TSP. *Computer Engineering* 34 (2008)
3. Hu, Z., Zhao, M.: Simulation on Traveling Salesman Problem (TSP) Based on Artificial Bees Colony Algorithm. *Transactions of Beijing Institute of Technology* 29 (2009)
4. Huang, B.-Z., Xiao, J.: Solving traveling salesman problem with improved MIMIC algorithm. *Computer Engineering and Design* 31 (2010)
5. Zhou, S.-D., Sun, Z.-Q.: A Survey on Estimation of Distribution Algorithms. *Acta Automatica Sinica* (2007)
6. Simionescu, P.A., Beale, D.G., Dozier, G.V.: Teeth-number synthesis of a multispeed planetary transmission using an estimation of distribution algorithm. *Journal of Mechanical Design* 128 (2006)
7. Joaquin, R., Roberto, S.: Improving the discovery component of classifier systems by the application of estimation of distribution algorithms. In: *Proceedings of Students Sessions, ACAI 1999, Chania, Greece* (1999)
8. Wang, L., Wang, S., Xu, Y., et al.: A bi-population based estimation of distribution algorithm for the flexible job-shop scheduling problem. *Computers & Industrial Engineering* 62 (2012)
9. Wang, L., Wang, S.-Y., Xu, Y.: An effective hybrid EDA-based algorithm for solving multidimensional knapsack problem. *Expert Systems with Applications* 39 (2012)
10. Wang, X., Li, Y.-X.: A Solution to Traveling Salesman Problem by Using Local Evolutionary Algorithm. *Computer Engineering* 32 (2006)
11. He, X.-J., Zeng, J.-C.: Solving TSP Problems with Estimation of Distribution Algorithm Based on Superiority Pattern Junction. *PTA&AT* 24 (2011)
12. Sheng, J., Xie, S.-Q.: *Probability and mathematical statistics*. Higher Education Press, BeiJing (2008)
13. He, X.-J., Zeng, J.-C.: Solving flexible job-shop scheduling problems with Bayesian statistical inference-based estimation of distribution algorithm. *Systems Engineering Theory & Practice* 32 (2012)
14. Hauschild, M., Pelikan, M.: An introduction and survey of estimation of distribution algorithms. *Swarm and Evolutionary Computation* 1 (2011)
15. Hao, C.-W., Gao, H.-M.: Modified Decimal MIMIC Algorithm for TSP. *Computer Science* 39 (2012)
16. Xiao-Juan, H., Jian-Chao, Z.: Solving flexible job-shop scheduling problems with Bayesian statistical inference-based estimation of distribution algorithm. *Systems Engineering Theory & Practice* 32, 380–388 (2012)
17. Hauschild, M., Pelikan, M.: An introduction and survey of estimation of distribution algorithms. *Swarm and Evolutionary Computation* 1, 111–128 (2011)
18. Cheng-Wei, H., Hui-Min, G.: Modified Decimal MIMIC Algorithm for TSP. *Computer Science* 39, 233–236 (2012)