An Approach to the Travelling Salesman Problem using Genetic Algorithm

Manabjyoti Choudhury **Dr. Laba Kumar Thakuria**
Department of Computer Science & Engineering,
Assam down town University, Guwahati 781026, India.

ABSTRACT

In this paper we have presented a solution for the travelling salesman problem using genetic algorithm. The solution provides a maximal approximation of the problem along with cost reduction. The solution also presents a method to find the nearly optimized solution for these types of optimization problems using a new crossover technique that produces a high quality solution to the TSP. Later, the paper presents a comparison of the effectiveness of the crossover operator with some traditional crossover operators.
Keywords: *Travelling Salesman Problem, Genetic Algorithm, Selection, Mutation, Crossover.*

1. INTRODUCTION

The Travelling Salesman problem (TSP)

The Travelling Salesman Problem (TSP) is a long standing problem in combinatorial optimization. It is an NP-hard problem that cannot be solved exactly in polynomial time and is an important problem in operation research and theoretical computer science. Though the problem is computationally difficult, yet researchers have developed many heuristic and exact algorithms to solve it. The aim of the problem is to find a tour with the least distance travelled through a set of n vertices so that each vertex is visited exactly once.

The Travelling Salesman Problem came to light when a German salesman, BF Voigt, authored a book on how to be a successful Travelling salesman in 1832. Even though by a different name, he addressed the TSP by suggesting that the most important feature of the scheduling of a tour is to travel to all the possible places without visiting any place twice.

The problem can be classified into two categories, namely – symmetric TSP and asymmetric TSP based on the basis of the structure of the cost matrix that contains the distances between the cities as the matrix elements. If the distances between two cities are equal in both directions then it is symmetric TSP, thus forming an undirected graph. However, in the case of asymmetric TSP there may not be paths in both directions between the cities, thus forming a directed graph.

For an n number of cities there are (n − 1)! possible solutions in the asymmetric TSP. One or more of them offers the minimum cost. On the other hand, for the symmetric TSP, there are (n − 1)!/2 possible solutions. These solutions have reverse cyclic permutations having the same total cost. In both the cases, the number of possible solutions grows to be very large even for a reasonably huge n so that an exhaustive search is not viable.

Genetic algorithm (GA)

Genetic algorithm (GA) is used as a search technique to find approximate solutions to combinatorial optimization problems. It is based on the theory of natural evolution and the concept of the survival of the fittest. Genetic algorithms provide better optimization solutions as compared to other conventional methods. The biggest advantage of GA is that even if one doesn't know how to solve the problem, one only needs to be able to evaluate the quality of the generated solutions and through iteration one can obtain the best solution. The technique is to first estimate the probable solutions and then combine the fittest ones among them to create a new generation of solutions. The new solutions are better than their parents in solving the problem. A random mutation element is also included in the technique. The method of the genetic algorithm process involves the following concepts:

1. Encoding: In order to ensure that each possible solution has a unique encoding style, an appropriate encoding is adopted for the solution to the problem.

2. Evaluation: The primary population of the possible solutions is selected randomly.

Then the fitness level of each of the individual solutions is estimated. This fitness value provides an indication of the fitness level of the solutions in solving the problem. Thus, the optimality of the solution can be compared with the other solutions in the population.

3. **Crossover**: In crossover, two individuals are recombined to produce an offspring. The fitness value is used to estimate the solutions' probability of crossover. The offspring form the next generation of solutions.

4. **Mutation**: In mutation, a mutation point and some of the individuals are randomly chosen to be mutated.

5. **Decoding:** A new generation of solutions is produced and this process is repeated in an iterative manner until some stopping condition is met. Then the solution which is closest to the optimum is decoded and the method terminates.

2. USING GENETIC ALGORITHM TO SOLVE THE TSP

Genetic algorithm provides a comparatively faster solution for the TSP. In less than a minute, one can find an almost perfect solution for a tour of 100 cities using genetic algorithm.

The steps for using genetic algorithm involve first to create a group of many random tours. This is called a population. Then the algorithm employs a greedy initial population that tends to link the neighbouring cities. Secondly, it picks two tours from the population and combines them to produce two new child tours. These child tours are expected to be better than their parents. Then, in order to make the children look distinguished from each other, mutation is performed on them. Then the mutated child tours are inserted into the population. The population remains unchanged in size. In this manner, new children tours are repeatedly produced until the desired goal is reached. The efficiency of the solution is determined by factors such as speed and population size. The best solution is selected by comparing these factors for all the children. Comparing and determining the best solution becomes complex with the increase in the size of the population.

3. RELATED WORK

In the work proposed by Aybars Uğur in the Genetic Algorithm Based Solution For TSP On A Sphere [1], the Euclidean TSP which is an NP hard problem has been related with determining the shortest tour through a given set of points in the d-dimensional Euclidean space. Fozia Hanif Khan has proposed a new representation method of chromosomes using the binary matrix and the new fittest criteria for finding the optimal solution for TSP in Solving TSP Problem By Using Genetic Algorithm [2]. A new crossover operator namely, Sequential Constructive crossover has been developed by Zakir H. Ahmed. This operator generates high quality solutions to the Travelling Salesman Problem (TSP) in Genetic Algorithm for the TSP using Sequential Constructive Crossover Operator [3]. An improved genetic algorithm has been proposed by Omar M.Sallabi. A new crossover operation, population reformulates operation, multi- mutation operation; partial local optimal mutation operation and rearrangement operation have been used in this algorithm, to solve the Travelling Salesman Problem in [4]. Angel Goñi Moreno has explained how to solve a fully connected N-City travelling salesman problem (TSP) using a genetic algorithm. A crossover operator to use in the simulation of a genetic algorithm (GA) with DNA has been presented in Solving Travelling Salesman Problem in A Simulation Of Genetic Algorithms With DNA[7]. Abdollah Homaifer, Shanguchuan Guan, Guna r E. Li epins have provided a substantial proof that genetic algorithms (GAs) work for the Travelling Salesman Problem (TSP). The discussed method is based on an adjacency matrix representation of the TSP. It allows the GA to manipulate the edges while using the conventional crossover technque.

4. METHODOLOGY

Initially, genetic algorithm creates a population of strings randomly. This population is called the gene pool. Then three operators are employed to produce a new and better population of off springs successively. The first operator is *reproduction*. Here the strings are copied to the next generation with some probability based on their objective function value. The second operator is *crossover*.

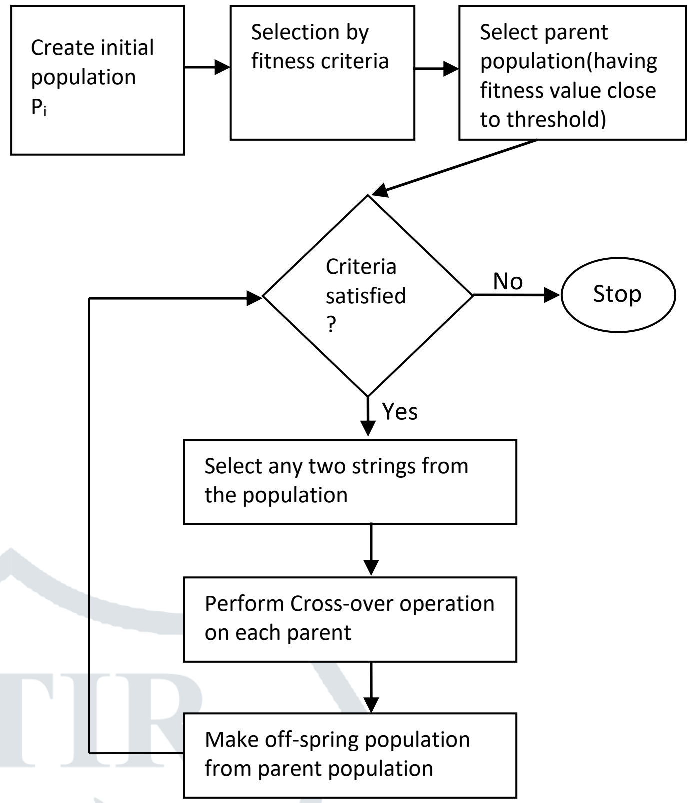
Here randomly selected pairs of strings are mated with each other to create new strings. The third operator is *mutation*. It is the occasional random alteration of the values of a string. The crossover and reproduction operations function as the most powerful process in the GA search mechanism. Mutation diversifies the search space. It also provides protection from loss of the genetic material that can be occur due to reproduction and crossover. Thus, the possibility of applying mutation is set very low. On the other hand, the probability of crossover is set very high.

Steps of the Algorithm

1. Randomly create the initial population of the individual strings of the given TSP problem and construct a path matrix to represent the costs of the paths between the cities.
2. Now, using the fitness criteria measure, F(x) = 1/x, assign the fitness to each chromosome in the population. Here, x represents the total cost of the string.

The selection criterion is determined by the value of the string. If it is close to some threshold value, then

3. Construct a new population of off springs from the existing chromosomes in the parent population by applying the crossover operator.
4. If necessary, mutate the resultant off springs.
The new population has a higher fitness value than its parent population.
5. Repeat steps 3 and 4 until an optimal solution to the problem is obtained.



In order to apply genetic algorithm to solve any optimization problem, one needs to find a way for encoding solutions as feasible chromosomes so that the crossover of feasible chromosomes produce more feasible solutions to the problem. The method of encoding the solutions varies from one problem to another. In order to solve the Travelling Salesman Problem, the solution is typically represented by the length of the chromosome which represents the number of nodes in the problem. Every gene of the chromosomes gets a unique name so as to avoid redundancy of the nodes. Basically, there are two methods of representing a tour in the Travelling Salesman Problem – the adjacency representation and the path representation. Let us consider the path representation for a tour, where the labels of the nodes are listed. For instance, let {1, 2, 3, 4, 5} be the labels of the nodes of a 5 node instance. Now, a tour {1 3 4 2 5 1} can be represented as {1, 3, 4, 2, 5}.

Fitness function

Genetic algorithms are used for maximization problems where the fitness function is same as the objective function. Whereas, for minimization problems, the fitness function can be defined as

F (x) = 1/ f (x), where f (x) is the objective function. As the Travelling Salesman Problem is a minimization problem, we can consider this fitness function, where f(x) yields the cost (or value) of the tour represented by a chromosome.

The Selection Process:

During the selection process, the chromosomes are copied into the next generation with a probability associated with their fitness value. By handing over to the next generation a section of the better fit chromosomes, the selection process proceeds in accordance with Darwin's theory of survival of the fittest.

In this paper we have used the Elitism method of selection. In this method, the best chromosome (or a few best chromosomes) is copied to the new population. The rest of the process is performed in the traditional way. This method can very rapidly increase the performance of genetic algorithm as it prohibits losing the best found solution.

Crossover Operator:

The solution space for the problem is created by producing new chromosomes from old ones. The most important search process is crossover. Initially, a pair of parent chromosomes is randomly selected from the population. Then the crossover point is selected along their common length. Then we have used a method of sequential constructive crossover operator before the crossover point. Then the genetic information after the crossover point of the two parent strings has been swapped. If a gene has been already copied into the new off-spring, then it means that the gene has been replaced by an unvisited gene and thus, creating two new children. The algorithm for this new crossover technique proceeds as follows:

Step 1. Start with the node p(p being the first node in the parents P1 and P2).

Step 2. Sequentially search both of the parent chromosomes and consider the first legitimate node that appeared after the node 1 in both P1 and P2. Assume that node x and node y are found in P1 and P2 respectively. Consider the crossover point is selected after the 2nd node in both parents P1 and P2.

Step 3. If Cpx < Cpy, then go for node x as the next node and concatenate it to the partially constructed offspring chromosome. Otherwise, select node y.

Step 4. If node x is selected as the next string in the partially constructed offspring chromosome, then copy the rest of the genes from parent P2, otherwise copy them from P1.

Step 5. Suppose a gene has already been copied into the off-spring then replace that gene by an unvisited gene.

This crossover technique has been illustrated by an example shown as a cost matrix in Table 1.

Consider a pair of selected chromosomes P1: (1, 3, 6, 4, 5, 7, 2) and P2: (1, 5, 4, 2, 6, 3, 7) with values 371 and 342 respectively.

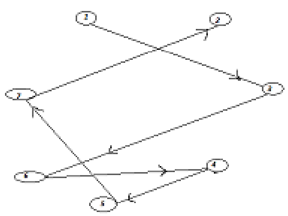
The cost matrix

We have selected the first node among the parents P1 and P2 as node 1. Then the valid nodes after node 1 are node 3 and node 5 respectively in both the parents P1 and P2. Next, we have considered the crossover point after node 3 and node 5 in parents P1 and P2 respectively. Then we have calculated the value of cost of the path as C13=98 and C15=34. As C15 < C13, so node 5 has been accepted as the next node in the partially constructed chromosome.

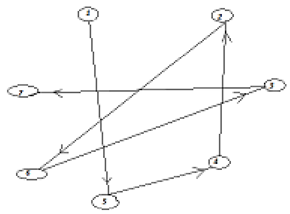
Node	1	2	3	4	5	6	7
1	100	74	98	8	34	62	7
2	50	100	85	45	87	28	19
3	49	4	100	15	27	34	27
4	19	44	10	100	58	42	48
5	85	62	32	64	100	75	71
6	35	52	88	30	20	100	51
7	57	30	42	66	51	59	100

Table1. Cost matrix of chromosomes P1 and P2.

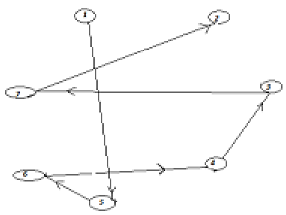
This node 5 has been selected from the parent P2. Hence, we have added the rest of the bits from the parent P1. Now the partially constructed chromosome is (1,5, 6,4,5,7,2). As node 5 has already been copied into the off-spring, so we have replaced node 5 with the unvisited node 3. Thus, the complete off-spring is (1,5,6,4,3,7,2). The value of this off-spring is 256 and it is less than the values of both the parent chromosomes. The crossover has been illustrated in the following figure. The parents are shown as (a) and (b) while the off-spring is (c).



(a) P1 : (1, 3, 6, 4, 5, 7, 2)



(b) P2: (1, 5, ,4 ,2, 6, 3, 7)



(C) Offspring: (1, 5, 6, 4, 3, 7, 2)

5. RESULT

We have illustrated an example of a graph consisting of seven vertices representing seven cities. We have represented the graph in the matrix representation method. It has been observed that our new crossover operator is capable of calculating an approximately optimal path for TSP using genetic algorithm in comparatively less

Algorithm	Time	Complexity	Advantage Disadvantage	Disadvantage
Genetic Algorithm	Exponential	O(Kmn)	Best solution using fitness criteria	Approximate solution is obtained though not an optimal solution.
Greedy Approach	5 seconds for 15 seconds	O(log n)	Fastest	No accuracy is achieved as no best choice is considered.
Dynamic Programming	9 seconds for 15 seconds	$O(n^2 2^n)$	Global optimal solution	Expensive in terms of memory and time.

time. Our crossover technique is better in terms of cost and time as compared to sequential constructive crossover (SCX) where there is a sequential search for each gene in a chromosome. This task is time consuming and also incurs a higher cost. In our crossover method, there is only a single search for a crossover point. Thereafter, the child chromosome is developed using the single point crossover technique only. This requires less time and cost and thus provides a better optimal solution for the given TSP problem.

6. CONCLUSION AND FUTURE WORK

Genetic algorithm provides efficient solutions for solving the Travelling Salesman Problem. However it relies on how the problem is encoded and which crossover and mutation methods are used .In this paper, we have proposed a new crossover operator for a genetic algorithm approach for the Travelling Salesman Problem (TSP).

It has been observed that our proposed crossover operator (SCX) is better in terms of cost and time. We have used a local search technique to improve the solution quality. We have also set the highest probability of crossover to show the exact nature of the crossover operator. When necessary, mutation with the lowest probability has been applied. A comparative study among Greedy approach, Dynamic programming and Genetic Algorithm for solving TSP has also been presented. It appears that other instances might be solved exactly with an integration of good local search techniques to the algorithm. This is under our investigation and is categorized under future work.

7. REFERENCES

[1] Aybars Uğur, Serdar Korukoğlu, Ali Calıskan, Muhammed Cinsdikici, Ali Alp, Department of Computer Engineering, "*Genetic algorithm based solution for TSP on a sphere*",Mathematical and Computational Applications,Vol. 14, No. 3, pp. 219-228, 2009

[2] Fozia Hanif Khan, Nasiruddin Khan, SyedInayaTullah, And Shaikh Tajuddin Nizami,"*Solving TSP problem by using GeneticAlgorithm*", International Journal of Basic &Applied Sciences IJBAS Vol: 9 No: 10

[3] Zakir H. Ahmed , "*Genetic Algorithm for theTSP using Sequential Constructive Crossover Operator*" International Journal of Biometrics & Bioinformatics (IJBB) Volume (3): Issue (6)

[4] Omar M.Sallabi, "*An Improved GeneticAlgorithm to Solve the Travelling SalesmanProblem"* World Academy of Science,Engineering and Technology 52 2009

[5] Eric matel , "*Solving the travelling salesman problem using a genetic algorithm*".

[6] Faulkner, G. Talhami, H. ,Kalgoorlie Coll.,WA, "*Using biological genetic algorithm tosolve the travelling salesman problem with applications in medical engineering*",Evolutionary Computation, 1995., IEEEInternational Conference in 1995.

[7] Geetha, R.R.; Bouvanasilan,N.; Seenuvasan, V.,"A perspective view on Travelling Salesman Problem using genetic algorithm" Nature & Biologically Inspired Computing, 2009. NaBIC 2009, World Congress.

[8] Kylie Bryant ,Arthur Benjamin, Advisor,"Genetic Algorithms and the TravellingSalesman Problem", Department of Mathematics, December 2000.

[9] Adewole Philip, Akinwale Adio Taofiki,Otunbanowo Kehinde "A Genetic Algorithm for Solving Travelling Salesman Problem",(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No.1, January 2011.

[10] Abdollah Homaifer, Shanguchuan Guan, Gunar E. Li epins "Schema Analysis of theTravelling Salesman Problem Using Genetic Algorithms".

[11] Milena Karova Vassil Smarkov Stoyan Penev,"Genetic operators crossover and mutation in

[12] solving the TSP problem", International Conference on Computer Systems and Technologies - CompSysTech" 2005.

[13] Ms. Monica Sehrawat, Mr. Sukhvir Singh , "Modified Order Crossover (OX) Operator",Monica Sehrawat et al. International Journal on Computer Science and Engineering(IJCSE).