# Travelling Salesman Problem Using Genetic Algorithms (A MINOR PROJECT REPORT)

Submitted By

Arijit Ray, Univ. Roll No.-21301217105, Univ. Reg. No.- 172131010016

Bhagyashree Maity, Univ. Roll No.-21301217098, Univ. Reg. No.- 172131010023

Jayashree Mahapatra, Univ. Roll No.-21301217, Univ. Reg. No.- 172131010047

Sankha Misra, Univ. Roll No.-21301217041, Univ. Reg. No.- 17213101080

*Under the Supervision of* 

# Prof. Arnab Kole Assistant Professor

In partial fulfillment for the award of the degree of

Bachelor of Computer Application



The Heritage Academy Maulana Abul Kalam Azad University of Technology

# **ACKNOWLEDGEMENT**

We would take the opportunity to thank *Prof. Gour Banerjee*, *Principal*, *The Heritage Academy* for allowing us to form a group of 4 people and for supporting us with the necessary facilities to make our project worth.

We are thankful to *Prof. Arnab Kole* (Assistant Professor, BCA) our Project Guide who constantly supported us, and Prof. Avik Mitra (Assistant Professor, BCA), the Project Coordinator for providing and clarifying the administrative formalities related to project proceedings. Their words and instructions have assisted us to excel in our project.

We thank all our other faculty members and technical assistants at The Heritage Academy for paying a significant role during the development of the project. Last but not the least, we thank all our friends for their cooperation and encouragement.

Signature:		
	(Arijit Ray)	
Signature:		
	(Bhagyashree Maity)	
Signature:		
21g.iu.u.u	(Jayashree Mahapatra)	
Signature:		
5151141410.	(Sankha Misra)	_

# The Heritage Academy **Kolkata**



# **PROJECT CERTIFICATE**

This is to certify that the following students:

	Name of the students	Roll No.	Registration No.
1.	ARIJIT RAY	21301217105	172131010016
2.	BHAGYASHREE MAITY	21301217098	172131010023
3.	JAYASHREE MAHAPATRA	21301217074	172131010047
4.	SANKHA MISRA	21301217041	172131010080

of 3rd Year 1st Semester in BCA(H) have successfully completed their Minor Project Work on

# TRAVELLING SALESMAN PROBLEM USING GENETIC **ALGORITHMS**

towards partial fulfillment of Bachelor of Computer Applications from Maulana Abul Kalam Azad University of Technology, West Bengal in the year 2019.

Prof. Dr.Gour Banerjee Prof. Aranb Kole Principal The Heritage Academy

Project Guide Assistant Professor Department of BCA The Heritage Academy

The Heritage Academy

Prof. Avik Mitra Project Coordinator Assistant Professor Department of BCA

# **ABSTRACT**

Meta-heuristic Search Algorithms are very efficient optimization techniques when it comes to problems that require large solution space. Instead of giving exact solution, it tries to produce an approximate one that takes less time. These algorithms are problem independent and hence can be applied to a wider variety of problems. Some popular meta-heuristic algorithms in the literature are Genetic Algorithms, Simulated Annealing, Ant Colony Optimization etc.

In this project, we have tried to harness the Genetic Algorithms to solve the travelling salesman problem on ATT48 dataset from TSPLIB and tried to find a solution that has a minimum cost of around 21,000 approx.

# **CONTENTS**

1	Introduction	1
2.	Genetic Algorithms	2
3.	Proposed Algorithm	5
4.	Parameters and Configuration	6
5.	Result and Discussion	7
6.	Reference	8
7.	Appendix	9
List	of Figures:	
1	Fig 1: A Roulette Wheel	3
2	Fig 2: Demo for One-Point Crossover	4
3	Fig 3: Demo for Two-Point Crossover	4
4	Fig 4: Screenshot of an average case	18
5	Fig 4: Screenshot of the best result yielded	19
List	of Tables:	
1	Table 1: Parameters used for GA	6

#### INTRODUCTION

The Travelling Salesman Problem is a classical optimization problem in the field of Computer Science and Operation Research. The formulation of this problem dates back to the 19<sup>th</sup> century where it was created as a recreational puzzle based on finding a Hamiltonian cycle. According to the problem statement, a salesman must travel between N cities in such a way that it visits each city only once and return to its starting city with the catch that the solution must create a tour with the shortest route possible.

Travelling salesman is a difficult problem to solve. The easiest but the most expensive solution is to simply try all the possible solutions. But for N number of cities the number of solutions will (N-1) be factorial. This means that for only 10 cities there are over 180 thousand combinations. This problem requires special approaches such as Genetic Algorithms, Simulated Annealing, Ant-colony optimisation, etc.

The problem was first formulated in 1930 and is one of the most intensively studied problems in optimization. It is used as a benchmark for many optimization methods. Even though the problem is computationally difficult, a large number of heuristics and exact algorithms are known, so that some instances with tens of thousands of cities can be solved completely and even problems with millions of cities can be approximated within a small fraction of 1%.

We have used Genetic Algorithm approach to solve this problem. Genetic Algorithm is a metaheuristic search algorithm. Genetic Algorithm gives optimum results in cases where the search spaces are very huge. It gives a approximate solution which is in general closer to the real solution. If we had to try every possible solution, the number of possible solutions for our dataset is (48-1) factorial. If we assume, that each possible solution takes 1ms, then to compute all of the possible solutions which will take (47 factorial)ms, which is greater than a decade.

The TSP has several applications even in its purest formulation, such as planning, logistics and even manufacturing of microchips. Slightly modified, it appears as a sub-problem in many areas, such as DNA sequencing. In these applications, the concept city represents, for example, customers, soldering points, or DNA fragments, and the concept distance represents travelling times or cost, or a similarity measure between DNA fragments. The TSP also appears in astronomy, as astronomers observing many sources will want to minimize the time spent moving the telescope between the sources. In many applications, additional constraints such as limited resources or time windows may be imposed.

#### ABOUT THE DATASET

Name: ATT48 TSP Dataset

Source: TSPLIB

• Contains Aerial Co-ordinates of 48 capital of the US

# **GENETIC ALGORITHMS**

A **genetic algorithm** is a search method that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation.

The process of natural selection starts with the selection of fittest individuals from a population. They produce offspring which inherit the characteristics of the parents and will be added to the next generation. If parents have better fitness, their offspring will be better than parents and have a better chance at surviving. This process keeps on iterating and at the end, a generation with the fittest individuals will be found.

There are five phases in a genetic algorithm.

- 1. Initial population
- 2. Fitness function
- 3. Selection
- 4. Crossover
- 5. Mutation

In terms of computation, Genetic Algorithms (GA) is a maximization technique with multiple probable maxima inside the search space. Hence, it may not give always give us the exact solution but will return an approximate one which will be close to the exact solution.

#### **BASIC GA PROCEDURE:**

- Choose initial population
- Evaluate the fitness of each individual in the population
- Repeat
  - Select best-ranking individuals to reproduce.
  - o Breed new generation through crossover and mutation (genetic operations) and give birth to offspring
  - o Evaluate the individual finesses of the offspring
  - o Replace worst ranked part of population with offspring
- Until <terminating condition>

#### BASIC TERMINOLGIES ASSOCIATED WITH GENETIC ALGORITHMS:

#### • Chromosomes:

A chromosome is a randomly generated one dimensional array of bits, digits/suitable characters that represents a solution in the search space of the problem.

#### • Decoding Functions:

These functions map a particular chromosome to its solution in the search space.

#### Fitness Functions:

Fitness Functions are objective functions that take a candidate solution represented as a chromosome from the search space and evaluates them. As Gas are maximization problems, high fitness values results in better solutions.

#### • Population Size:

A cluster of chromosomes are called the population. It is a hyper-parameter usually adjusted by the algorithm designer. The initial population is generated at random.

# Mating Pool:

Set of selected chromosomes that will produce new off-springs.

#### • GA Operators:

The whole GA technique has 3 primary operations:

- 1. **Selection:** Selection operator picks chromosomes from the existing population to create a mating pool. There are 2 types of widely used selection techniques:
  - **Roulette Wheel:** In this, the chances of the chromosomes to be selected for mating pool are correlated to its fitness function. Imagine a roulette wheel. Now calculate sum of all chromosome finesses in population as  $sum\_fitness$   $(sum\_fitness = \sum_{i=0}^{popsize} f(chr[i]);$  where popsize is population size; chr is chromosomes; f(x) is the fitness function).

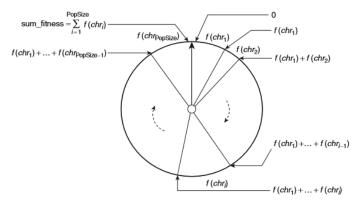


Fig 1: A Roulette Wheel [1]

This sum fitness is the total circumference of the wheel with partitions equal to *popsize* and the area of partition for each chromosomes is equivalent to its corresponding fitness value. Now, a random number is generated in between 0 and *sum\_fitness*, say *d*. The chromosomes within whose limit d falls is taken in for the mating pool. This process is repeated *popsize times* to create a mating pool.

■ *Tournament Selection:* In this strategy, 2 chromosomes are random selected from the current population and selected for the mating pool on the basis of their fitness score i.e. whichever has the higher fitness score gets into the mating pool. The procedure runs for *popsize* times.

- 2. **Crossover:** In this operation, 2 chromosomes are randomly selected and some of their segments are swapped to give 2 new off springs. The swapping operation is done by deciding whether they would go for crossover. For this, a threshold probability  $p_c$  is predefined by the algorithm designer. Now, a random number is generated in between 0 and 1, say r. If  $r \le p_c$ , the crossover occurs. The crossover point is generated between 0 and *chrlength* (*chrlength* is the length of the chromosomes). The crossover occurs by swapping the segments of the crossover point. There are 2 crossover techniques:
  - **a.** *One-Point Crossover:* In this technique, only one crossover point p is generated between 0 and *chrlength* and the bits of the parent chromosomes are swapped from  $p^{th}$  bit till *chrlength*

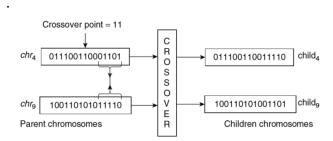


Fig 2: Demo for One-Point Crossover [2]

**b.** Two-Point Crossover: In this technique, 2 crossover points  $p_1$  and  $p_2$  is generated between 0 and *chrlength* and the bits of the parent chromosomes are swapped from  $p_1$ <sup>th</sup> bit till  $p_2$ <sup>th</sup>bit.

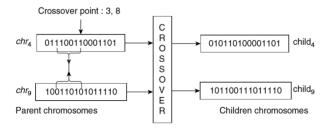


Fig 3: Demo for Two-Point Crossover [3]

- 3. **Mutation:** The mutation operation is implemented to create a small change the chromosomes so that is explored unbiased. A threshold probability  $p_m$  is predefined by the algorithm designer for deciding whether the mutation operation is to be executed or not. Now a random number r between 0 and 1 is generated. If  $r \le p_m$ , Mutation occurs. The mutation probability is usually kept low (say  $p_m < 0.1$ ) because mutation is a permanent change in the gene and hence can deviate from the solution.
- **Termination Condition:** Termination can be the no. of iterations or generations of the GA.

# PROPOSED ALGORITHM

#### TRAVELING SALESMAN PROBLEM USING GENETIC ALGORITHMS

#### **ALGORITHM:**

- Step-1: Read in the dataset and calculate the aerial distance of all the 48 cities from each other, initialising them into a cost matrix.
- Step-2: Covert the cost matrix into reward matrix.
- Step-3: Initialising 48 binary chromosomes of 38 bit each as initial population
- Step-4: Repeat Step 4 to Step 7 till the number of iterations mentioned by the user
- Step-5: Apply selection operation (Tournament/Roulette Wheel) to create a new mating pool
- Step-6: Apply *crossover* and *mutation* operators on the mating pool to generate the *new* population
- Step-7: Calculate fitness score of the new population. Store the elite chromosome i.e. the chromosomes with high fitness value of the new population.
- Step-8: Replace the current population by the new population.
- Step-9: Return the *best* solution of the current population

# PARMETERS AND CONFIGURATION

#### **PARAMETERS:**

Parameters are customizable units of an algorithm which when modified can impact the performance of the algorithm. The algorithm designer is responsible for formulating and alterations of the parameter as deemed necessary.

In our project, we have considered the following as the necessary parameters of the GA process:

- Number of Iterations
- Crossover Probability
- Mutation Probability

#### **CONFIGURATIONS:**

The project was created in DEV C++ (5.11) IDE and run on TDM-GCC 4.9.2 (64 bit Release)

Operating System: Windows 10 Home 64 bit(10.0, Build:18362)

Processor: Intel Core i5-8250U CPU @ 1.6 GHz x 4

RAM: 8 GB

Also Tested and Run on Xcode Version 11.2.1 (11B500)

Operating System: macOS Mojave (Version 10.14.6)

Processor: Intel Core i5-5250U CPU @ 1.6 GHz x 2

RAM: 8 GB

# **RESULTS AND DISCUSSIONS**

#### **RESULT:**

Best results were found when threshold crossover probability and mutation probabilities were set to 0.7 and 0.05 respectively. Among the Crossover techniques, 2 point crossover gave better results. Tournament Selection was more effective than Roulette Wheel Selection in our case. As it is a Heuristic Search Algorithm, exact solutions are difficult to find. The whole process is computationally intensive

Population Size for Each Case is 50.

No. of Iterations	Crossover type	Crossover Threshold	Mutation threshold	Result
50000	2-point	0.7	0.05	21951.7
50000	1-point	0.7	0.05	23802.1
75000	2-point	0.7	0.05	22297
100000	2-point	0.7	0.05	23679
100000	2-point	0.7	0.05	22926.2

Table 1: Parameters used for GA

# **DISCUSSION:**

From multiple builds run, we found out that Optimum number of iterations is approximately 50000.

When iterations were greater than 50000, the result was not optimum in majority of the cases

In case of both 1-point and 2-point crossovers, results were similar but 2-point crossover gave results more close to the expected solution. The best result was found using 50000 iterations with 2-point crossover type and the result in 21951.7.

# CONCLUSION AND FUTURE WORKS

#### **CONCLUSION**

To solve the travelling salesman problem we have used Genetic Algorithm approach. But as we know Genetic Algorithm does not provide us with the exact solution. With every build run, we get a new result. The results are based on random generations and mutations. But, Genetic Algorithm gives us a value close to the actual result. For more accurate results, we need to consider different approaches. Comparing every possible solution is very time consuming and intensive task. Using of meta-heuristic search algorithms and other algorithms are very necessary.

#### **FUTURE SCOPE**

The travelling salesman problem has many applications in different fields. To get further better results, we need to make modifications to the proposed algorithm. We might also use different heuristic approaches or different algorithms. We might also build an hybrid algorithm using genetic algorithm and simulated annealing. Ant-colony optimization technique can also be tested for the travelling salesman problem

# **REFERENCES**

- 1. Fig 12.9, Samir Roy, Udit Chakraborty, "Introduction to soft computing Neuro-fuzzy and Genetic Algorithms", Chapter 12, Page 11, Pearson India, 2013
- 2. Fig 12.13, Samir Roy, Udit Chakraborty, "Introduction to soft computing Neuro-fuzzy and Genetic Algorithms", Chapter 12, Page 14, Pearson India, 2013
- 3. Fig 12.13, Samir Roy, Udit Chakraborty, "Introduction to soft computing Neuro-fuzzy and Genetic Algorithms", Chapter 12, Page 14, Pearson India, 2013
- 4. https://en.wikipedia.org/wiki/Travelling\_salesman\_problem (Last Accessed: 25-11-19)
- 5. https://www.geeksforgeeks.org/travelling-salesman-problem-set-1/2 (Last Accessed: 26-11-19)

#### **APPENDIX**

```
#include<iostream>//including required header files
#include<conio.h>
#include<cstdlib>
#include<ctime>
#include<math.h>
#include<fstream>
using namespace std;
int iter;
double cdef, mdef;
double costMatrix[48][48];
double rewardMatrix[48][48];
int chromosomes[48][288];
int combinations[48][48];//stores decoded chormosomes
double fitness[48];//stores fitness of decoded chromosomes
int cityBuffer[48];//buffer to identify the cities decoded and used for each chromosomes
int elite[100000][48];//stores elite
double elscore[100000];
double cost[100000];
int elcount=0;
int tempchrom[48][288];
int coords[48][2];
int genChrom()//generating binary chromosomes
       int i,j;
       for(i=0;i<48;i++)
               for(j=0;j<288;j++)
               {
                      chromosomes[i][j]=(rand()%2);
               }
       }
}
double findDist(int i, int j)// udf to find reward between two cities
{
       double r;
       if(i==j)
               return -1.0;
       else
               r=costMatrix[i][j];
               return r;
```

```
double dist(int i, int j)
       double sum,a,b;
       a=pow((coords[i][0]-coords[i][0]),2);
       b=pow((coords[j][1]-coords[i][1]),2);
       sum = sqrt((a+b)/10.0);
       return sum;
int genCost()
       int i,j;
       double max=0.0,d=0;
       for(i=0;i<48;i++)
               for(j=i+1;j<48;j++)
               {
                       d=dist(i,j);
                       costMatrix[i][j]=d;
                       costMatrix[j][i]=d;
               }
       for(i=0;i<48;i++)
               for(j=0;j<48;j++)
                       if(max<costMatrix[i][j])</pre>
                               max=costMatrix[i][j];
               }
       cout<<endl<<"Max: "<<max<<endl;</pre>
       for(int i=0;i<48;i++)
        {
               for(int j=0; j<48; j++)
                       if(i!=j)
                              rewardMatrix[i][j]=max-costMatrix[i][j];
               }
        }
int printCost()
```

```
int i,j,c=0;
        cout<<"Generated Cost Matrix:\n \n";</pre>
        for(i=0;i<49;i++)
               for(j=0;j<49;j++)
                       if(i==0)
                       {
                               if(i==0 && j==0)
                                       cout<<" \t";
                               else
                                       cout << j-1 << "\t";
                       }
                       else
                       {
                               if(j==0)
                                       cout<<c<"\t";
                               else
                                       cout << costMatrix[i-1][j-1] << "\setminus t";
                       }
               }
               if(i>=1)
                       ++c;
               if(i==0)
                       cout<<endl;
               cout<<endl;
        }
int printReward()
{
       int i,j,c=0;
       cout<<endl<<"Reward Matrix:\n \n";</pre>
        for(i=0;i<49;i++)
        {
               for(j=0;j<49;j++)
                       if(i==0)
                       {
                               if(i==0 && j==0)
                                       cout<<" \t";
                               else
                                       cout << j-1 << "\t";
                       }
```

```
else
                              if(j==0)
                                      cout<<c<"\t";
                               else
                                      cout<<rewardMatrix[i-1][j-1]<<"\t";</pre>
               }
               if(i>=1)
                       ++c;
               if(i==0)
                       cout<<endl;
               cout<<endl;
        }
       cout<<endl;
}
double findReward(int i, int j)// udf to find reward between two cities
       double r;
       if(i==j)
               return -1.0;
       else
               r=rewardMatrix[i][j];
               return r;
        }
void elitism()
       int i,ind;
       double max=0;
       for(i=0;i<48;i++)
               if(max<fitness[i])</pre>
                       max=fitness[i];
                       ind=i;
               }
       for(i=0;i<48;i++)
               elite[elcount][i]=combinations[ind][i];
       elscore[elcount]=max;
       elcount++;
```

```
int decode() //decoding the chromosomes
       int i,j,d,c,k;
       for(k=0;k<48;k++)
       {
              c=0;
              for(i=0;i<48;i++)
               {
                      c=i*6;
                      d=0;
                      for(j=5;j>=0;j--)
                             d=d+(chromosomes[k][c+(5-j)]*pow(2,j));
                      if(d<48)
                             while(cityBuffer[d]!=0)
                                     d=(d+1)\%48;
                      }
                      else
                      {
                             d=0;
                             while(cityBuffer[d]!=0)
                                     d=(d+1)\%48;
                      cityBuffer[d]=1;
                      combinations[k][i]=d;
                      c++;
              for(i=0;i<48;i++)
                      cityBuffer[i]=0;
       }
void fitneScore() //finding and storing fitness scores of each chromosomes
       int i,j;
       double r;
       decode();
       for(i=0;i<48;i++)
       {
              r=0.0;
              for(j=0;j<47;j++)
```

```
r+=findReward(combinations[i][j],combinations[i][j+1]);
              r+=findReward(combinations[i][47],combinations[i][0]);
              fitness[i]=r;
       }
}
void selection()
       int i,j,r1,r2,select;
       for(i=0;i<48;i++)
       {
              r1=rand()%48;
              r2=rand()%48;
              while(r1==r2)
                     r1=rand()%48;
              if(fitness[r1]>fitness[r2])
                     select=r1;
              else
                     select=r2;
              for(j=0;j<288;j++)
                     tempchrom[i][j]=chromosomes[select][j];
              }
       for(i=0;i<48;i++)
              for(j=0;j<288;j++)
                     chromosomes[i][j]=tempchrom[i][j];
               }
       }
void crossover()
       int i,j,k,swap,pos1,pos2,chr1,chr2;
       double cp;
       for(i=0;i<24;i++)
              cp=(rand()%1000)/1000.0;
              if(cp<cdef)
                     chr1=rand()%48;
                     chr2=rand()%48;
                      while(chr1==chr2)
```

```
chr1=rand()%48;
                     pos1=rand()%288;
                    pos2=rand()%288;
                    while(pos1==pos2)
                           pos1=rand()%288;
                    if(pos1>pos2)
                           swap=pos1;
                           pos1=pos2;
                           pos2=swap;
                    for(k=pos1;k<=pos2;k++)
                           swap=chromosomes[chr1][i];
                           chromosomes[chr1][i]=chromosomes[chr2][i];
                           chromosomes[chr2][i]=swap;
                     }
              }
       }
}
void mutation()
{
      int i,pos;
      double mp;
      for(i=0;i<48;i++)
       {
             mp = (rand()\% 1000)/1000.0;
             if(mp<mdef)
              {
                     pos=rand()%288;
                    if(chromosomes[i][pos]==0)
                           chromosomes[i][pos]=1;
                    else
                           chromosomes[i][pos]=0;
              }
       }
}
void GeneticAlgorithm(int n,double c=0.7,double m=0.01)
      int i,j;
      iter=n;
      cdef=c;
      mdef=m;
```

```
genChrom();
       cout<<"Generation : 0"<<endl;</pre>
       fitneScore();
       elitism();
       cout<<"Elite: ";
       for(j=0;j<48;j++)
               cout<<elite[elcount-1][j]<<" ";</pre>
       cout<<":\t"<<elscore[elcount-1];</pre>
       cout<<endl<<endl;
       for(i=1;i<iter;i++)
               if(i\%500==0)
                       srand(time(NULL));
               selection();
               crossover();
               mutation();
               fitneScore();
               elitism();
        }
int main()
{
       int n;;
       double c,m,dist=0.0;
       ifstream inData("tspdistance.txt");
  for (int i=0; i<48; i++) {
       inData >> coords[i][0] >> coords[i][1]; //Read x and y coordinates at position i
       cout << coords[i][0] <<"\t"<< coords[i][1] << endl; //Output the x and y coordinates
read above
       cityBuffer[i]=0;
  inData.close();
  genCost();
  printCost();
  printReward();
  cout<<"No. of Iteration"<<endl;
  cin>>n;
  cout<<"Enter Crossover Probability"<<endl;</pre>
       cin>>c;
       cout<<"Enter Mutation Probability"<<endl;</pre>
       cin>>m;
       srand(time(NULL));
       cout<<endl;
```

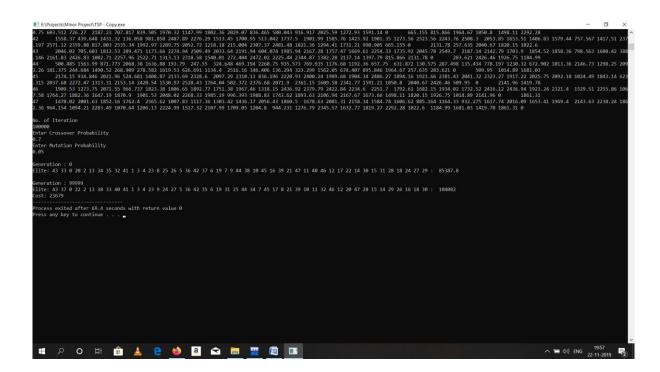


Fig 4: Screenshot of an average case

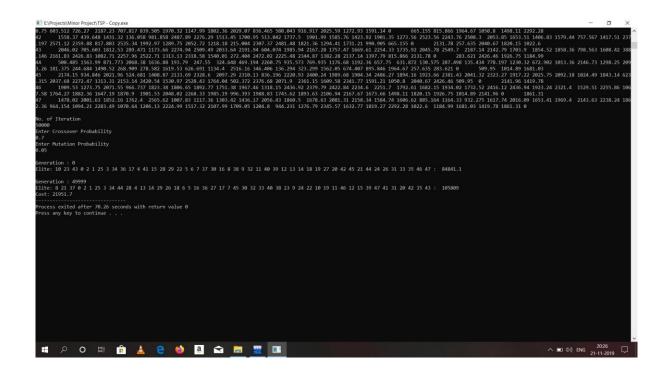


Fig 5: Screenshot of the best result yielded