

Solving TSP using Genetic Algorithms - Case of Kosovo

AVNI REXHEPI, AGNI DIKA, ADNAN MAXHUNI

Department of computerics

Faculty of Electrical and Computer Engineering

University of Prishtina

Kodra e Diellit, p.n. 10000 Prishtina

KOSOVO

avni.rexhepi@uni-pr.edu, agni.dika@uni-pr.edu, adnan.maxhuni@uni-pr.edu

<http://www.uni-pr.edu/fiek>

Abstract: In this paper we describe the use of a Genetic Algorithm to solve the TSP problem for Kosovo Cities/Municipalities, where one has to minimize the travelling distance between the cities (locations). Since the TSP problem is a NP problem, approaches such as dynamic programming, backtracking, branch and bound, etc. are not very useful for solving it. For problems traditionally thought of as computationally infeasible such as the TSP, Genetic Algorithms prove to be the best approach in obtaining solutions.

John Holland's book "Adaptation in Natural and Artificial Systems" (1975, 1992) showed how the evolutionary process can be applied to solve a wide variety of problems using a highly parallel technique that is called the "Genetic Algorithm" (GA). The genetic algorithm transforms a population (set) of individual objects, each with an associated fitness value, into a new generation of the population using the Darwinian principle of reproduction and survival of the fittest and analogs of naturally occurring genetic operations such as crossover (sexual recombination) and mutation. Each individual in the population represents a possible solution to a given problem. The genetic algorithm attempts to find a very good (or best) solution to the problem by genetically breeding the population of individuals over a series of generations.

Kosova is a relatively small country, in aspect of territory size, which is 10887 Km². Besides capital Prishtina, there are six other bigger cities and about 23 smaller cities, while the rest of locations are villages. Basically, there are 30 municipalities, with 1467 localities, in total, with about two milion residents. Recently, seven localities are proposed to be recognised as municipalities and are in the process of transformation.

For solving the problem, we designed an application in C#, which will use a genetic algorithm to find the best possible solution for TSP Kosova problem.

The first part of this paper is a brief presentation Kosovo and Kosova municipalities.

Key-Words: - Genetic Algorithms, TSP, C#, Kosovo.

1 Introduction

Kosovo is the newest country in Europe and as per size of the territory, belongs to the small countries (169 in a world rank for territory size). It has 10887 Sq Km. Kosovo is located in Southeast Europe, in Balkan Peninsula, in geographic coordinates 42 35 N and 21 00 E, between Albania, Montenegro, Serbia and Macedonia,. The total border line is 702 km, with border countries: Albania 112 km, Macedonia 159 km, Montenegro 79 km and Serbia 352km.

There are 7 big cities in Kosovo, with capital Prishtina and 30 municipalities in total. Most of the roads are paved. There are around 2000 km of roadways and 94% of them are paved. Only 2% are international highways, while 66% are national and 32% are regional ways.



Fig.1 – Kosovo position in Balkan Peninsula

In this paper we describe the Traveling Salesman Problem (TSP) for Kosovo municipalities, using the genetic algorithm.

2 Traveling Salesman Problem

TSP is a problem, where traveling salesman wants to visit each of a set of cities exactly once, starting from hometown and returning to his hometown. His problem is to find the shortest route for such a trip. TSP has a model character in many branches of Mathematics, Computer Science, Operations Research, etc. Linear programming, heuristics and branch and bound which are main components for the most successful approaches to hard combinatorial optimization problems, were first formulated for the TSP and used to solve practical problem instances in 1954 by Dantzig, Fulkerson and Johnson.

When the theory of NP-completeness developed, the TSP was one of the first problems to be proven NP-hard by Karp in 1972. New algorithmic techniques have first been developed for or at least have been applied to the TSP to show their effectiveness. Such examples are branch and bound, Lagrangean relaxation, Lin-Karneyhan type methods, simulated annealing, etc [3].

Representation model is: Let $K_n=(V_n, E_n)$ be the complete undirected graph with $n=|V_n|$ nodes and $m=|E_n|=\binom{n}{2}$ edges. An edge e with endpoints i and j is also denoted by ij , or by (i,j) . We denote by R^{E_n} the space of real vectors whose components are indexed by the elements of E_n . The component of any vector $z \in R^{E_n}$ indexed by the edge $e=ij$ is denoted by z_e, z_{ij} , or $z(i,j)$.

Given an objective function $c \in R^{E_n}$, that associates a "length" c_e with every edge e of K_n , the symmetric traveling salesman problem consists of finding a Hamiltonian cycle such that its c -length (the sum of the lengths of its edges) is as small as possible.

Of special interest are the Euclidean instances of the traveling salesman problem. In these instances the nodes defining the problem correspond to points in the two-dimensional plane and the distance between two nodes is the Euclidean distance between their corresponding points. More generally, instances that satisfy the triangle inequality, i.e., $c_{ij} + c_{jk} \geq c_{ik}$ for all the three distinct i,j and k , are of particular interest.

For our case, we consider the locations of the cities/municipalities in Kosovo map as nodes of the graph. For to do this, we take their geographic coordinates and than based on that, we calculate their position in our map scaled to a smaller size, for to calculate the real positions and distances, by using the real life values for distances in kilometres between the cities.

3 Genetic Algorithms

Genetic Algorithms (GA) [1,2] are computer algorithms that search for good solutions to a problem within a large number of possible solutions. They were proposed and developed in the 1960s by John Holland, his students, and his colleagues at the University of Michigan. These computational paradigms were inspired by the mechanics of natural evolution, including survival of the fittest, reproduction, and mutation. These mechanics are well suited to resolve a variety of practical problems, including computational problems, in many fields. Some applications of GAs are optimization, automatic programming, machine learning, economics, immune systems, population genetic, and social system.

GAs have been successfully applied to many problems of business, engineering, and science. Because of their operational simplicity and wide applicability, GAs play an important role in computational optimization and operations research [6].

The genetic algorithm transforms a population (set) of individual objects, each with an associated fitness value, into a new generation of the population using the Darwinian principle of reproduction and survival of the fittest and analogs of naturally occurring genetic operations such as crossover (sexual recombination) and mutation. Each individual in the population represents a possible solution to a given problem. The genetic algorithm attempts to find a very good (or best) solution to the problem by genetically breeding the population of individuals over a series of generations.

3.1 Basic elements of GAs

Most GAs methods are based on the following elements: populations of chromosomes, selection according to fitness, crossover to produce new offspring, and random mutation of new offspring. The chromosomes in GAs represent the space of candidate solutions. Possible chromosomes encodings are binary, permutation, value, and tree encodings. GAs require a fitness function which allocates a score to each chromosome in the current population. Thus, it can calculate how well the solutions are coded and how well they solve the problem [2].

The selection process is based on fitness. Chromosomes that are evaluated with higher values (fitter) will most likely be selected to reproduce, whereas, those with low values will be discarded.

The fittest chromosomes may be selected several times, however, the number of chromosomes selected to reproduce is equal to the population size, therefore, keeping the size constant for every generation. This phase has an element of randomness just like the survival of organisms in nature. The most used selection methods, are roulette-wheel, rank selection, steady-state selection, and some others. Moreover, to increase the performance of GAs, the selection methods are enhanced by elitism. Elitism is a method, which first copies a few of the top scored chromosomes to the new population and then continues generating the rest of the population. Thus, it prevents losing the few best found solutions.

Crossover is the process of combining the bits of one chromosome with those of another to create an offspring for the next generation that inherits traits of both parents. For example, consider the following parents and a crossover point at position 3:

```
Parent 1      1 0 0 | 0 1 1 1
Parent 2      1 1 1 | 1 0 0 0
Offspring 1   1 0 0 1 0 0 0
Offspring 2   1 1 1 0 1 1 1
```

In this example, Offspring 1 inherits bits in position 1, 2, and 3 from the left side of the crossover point from Parent 1 and the rest from the right side of the crossover point from Parent 2. Similarly, Offspring 2 inherits bits in position 1, 2, and 3 from Parent 2 and the rest from Parent 1.

Mutation is performed after crossover to prevent falling all solutions in the population into a local optimum of solved problem. Mutation changes the new offspring by flipping bits from 1 to 0 or from 0 to 1. Mutation can occur at each bit position in the string with some probability, usually very small (e.g. 0.001). For example, consider the following chromosome with mutation point at position 3:

```
Not mutated chromosome: 1 0 0 0 1 1 1
Mutated chromosome:     1 0 1 0 1 1 1
```

The 0 at position 3 flips to 1 after mutation.

So, general outline of basic GA:

1. Start: Randomly generate a population of N chromosomes.
2. Fitness: Calculate the fitness of all chromosomes.
3. Create a new population:
 - a. Selection: According to the selection method select 2 chromosomes from the population.
 - b. Crossover: Perform crossover on the 2 chromosomes selected.
 - c. Mutation: Perform mutation on the chromosomes obtained.
4. Replace: Replace the current population with the new population.

5. Test: Test whether the end condition is satisfied. If so, stop. If not, return the best solution in current population and go to Step 2.

Each iteration of this process is called generation. The genetic algorithm object determines which individuals should survive, which should reproduce, and which should die. It also records statistics and decides how long the evolution should continue. A typical genetic algorithm will run forever, so we must build functions for specifying when the algorithm should terminate. These include terminate-upon generation, in which you specify a certain number of generations for which the algorithm should run, and terminate-upon-convergence, in which you specify a value to which the best-of-generation score should converge. One can customize the termination function to use own stopping criterion and must tell the algorithm when to stop. Often the number-of generations is used as a stopping measure, but you can use goodness-of-best-solution, convergence-of-population, or any problem-specific criterion if you prefer.

There are some flavors of genetic algorithms. For example, the first is the standard 'simple genetic algorithm' described by Goldberg in his book [2]. This algorithm uses nonoverlapping populations and optional elitism. Each generation the algorithm creates an entirely new population of individuals. The second is a 'steady-state genetic algorithm' that uses overlapping populations. In this variation, you can specify how much of the population should be replaced in each generation. The third variation is the 'incremental genetic algorithm', in which each generation consists of only one or two children. The incremental genetic algorithms allow custom replacement methods to define how the new generation should be integrated into the population. So, for example, a newly generated child could replace its parent, replace a random individual in the population, or replace an individual that is most like it. The fourth type is the 'deme' genetic algorithm. This algorithm evolves multiple populations in parallel using a steady-state algorithm. Each generation the algorithm migrates some of the individuals from each population to one of the other populations.

The base genetic algorithm class contains operators and data common to most genetic algorithms. The genetic algorithm contains the statistics, replacement strategy, and parameters for running the algorithm. the population object, a container for genomes, also contains some statistics as well as selection and scaling operators.

The number of function evaluations is a good way to compare different genetic algorithms with

Nr./Rendbor	Komuna	Kodi	Numri i VendiBanimeve	Pozita Gjeografike N Shkalles/minuta	Pozita Gjeografike E Shkalles/minuta	Longitude	Latitude
1	Deçan	01	37	42.32	20.17	42.54	20.28
2	Gjakovë	02	88	42.23	20.26	42.38	20.42
3	Glogoc	03	35	42.37	20.53	42.62	20.88
4	Gjilan	04	49	42.27	21.28	42.47	21.48
5	Dragash	05	36	42.03	20.39	42.06	20.65
6	Istog	06	50	42.46	20.29	42.78	20.48
7	Kaçanik	07	31	42.13	21.15	42.23	21.26
8	Kline	08	54	42.37	20.34	42.62	20.57
9	Fushë Kosovë	09	16	42.38	21.05	42.64	21.09
10	Kamenicë	10	58	42.34	21.34	42.58	21.58
11	Mitrovicë	11	47	42.53	20.52	42.88	20.86
12	Leposaviq	12	75	43.06	20.48	43.10	20.80
13	Lipjan	13	62	42.31	21.07	42.52	21.12
14	Novobërdë	14	26	42.36	21.25	42.60	21.43
15	Obiliq	15	20	42.41	21.04	42.68	21.07
16	Rahovec	16	36	42.24	20.39	42.41	20.65
17	Pejë	17	79	42.40	20.18	42.65	20.30
18	Podujevë	18	77	42.54	21.11	42.91	21.19
19	Prishtinë	19	43	42.40	21.10	42.66	21.16
20	Prizren	20	76	42.13	20.44	42.21	20.73
21	Skenderaj	21	49	42.44	20.47	42.75	20.79
22	Shtime	22	23	42.25	21.02	42.43	21.04
23	Shterpcë	23	16	42.14	21.01	42.24	21.02
24	Suharekë	24	41	42.21	20.49	42.36	20.82
25	Ferizaj	25	45	42.22	21.10	42.37	21.17
26	Viti	26	39	42.19	21.21	42.32	21.36
27	Vushtrri	27	67	42.49	20.57	42.82	20.96
28	Zubin Potok	28	61	42.55	20.41	42.92	20.69
29	Zveçan	29	35	42.54	20.50	42.91	20.84
30	Malishevë	30	44	42.29	20.44	42.48	20.74

Table 3 – Geographical coordinates, positions

By running the application, we can calculate the shortest traveling distance between the selected cities, by finding a solution of the TSP using a genetic algorithm.

We also use two lists, where the first one will show the selected cities, while the second one will show the sequence of the cities in the found solution. Application will calculate and show the total traveling distance in kilometres.

User can set up the numbers for initial population, maximal number of generations, size of the group, percentage of mutations and number of close cities/locations (used by the algorithm, while finding closest locations).

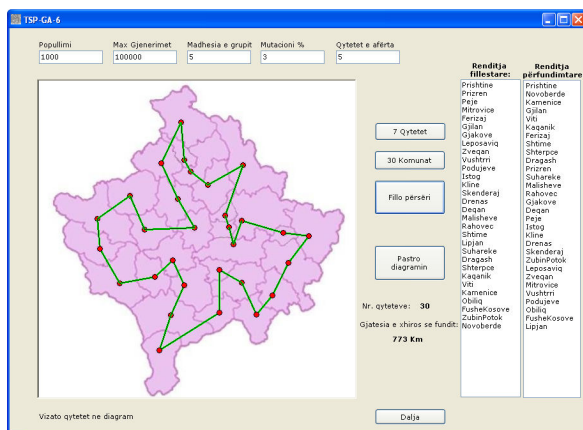


Fig.4 – TSP Solution for 30 municipalities

6 Conclusion

Finding a best solution for NP-hard problems is not possible by using standard mathematical approach. Genetic Algorithms are computer algorithms that search for good solutions to a problem within a large number of possible solutions.

Traveling Salesman Problem is a NP-Hard problem. Using a genetic algorithm, we can find a feasible solution, which will not have guaranty to be the best possible solution, but will be the good one, found in reasonable time.

- In order to find the shortest traveling distance between Kosovo cities/municipalities we use a genetic algorithm, in an application built in C#.
- We use geographical coordinates to calculate locations of the nodes/cities in the graph/image.
- We use a matrix of inter-city distances to calculate the length of the traveling distance.
- Application will find the shortest traveling distance between the selected cities.

References:

- [1] J. H. Holland, *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press, 1975.
- [2] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley, 1989.
- [3] Michael Junger, Gerhard Reinelt, Giovanni Rinaldi, *The Traveling Salesman Problem*, M.O. Ball et al, Eds. Handbooks in OR & Ms, Vol. 7, Elsevier Science, B.V. 1997
- [4] T.-L. Yu, D. E. Goldberg, and Y.-P. Chen, “A genetic algorithm design inspired by organizational theory: A pilot study of a dependency structure matrix driven genetic algorithm,” IlliGAL Report No. 2003007, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, Urbana, IL, 2003.
- [5] K. Sastry, D. E. Goldberg, and G. Kendall, “Genetic algorithms: A tutorial,” in *Introductory Tutorials in Optimization, Search and Decision Support Methodologies*, ch. 4, pp. 97–125, Springer, 2005.
- [6] Martin Pelikan, *Genetic Algorithms*, MEDAL Report No. 2010007, 2010