ARTIFICIAL INTELLIGENCE

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
UNIVERSITY OF WEST ATTICA

DEPARTMENT OF INFORMATION AND COMPUTER ENGINEERING

REAL GENETIC ALGORITHM APPLICATION

STUDENT DETAILS

NAME: ATHANASIOU VASILEIOS EVANGELOS
STUDENT ID: 19390005
SEMESTER: 7th
STUDENT STATUS : UNDERGRADUATE
STUDY PROGRAM : UNIWA
THEORY INSTRUCTOR: MASTOROKOSTAS PARIS
LAB INSTRUCTOR : TSELENTI PANAGIOTA
DELIVERY DATE : 1/23/2023

ARTIFICIAL INTELLIGENCE

STUDENT PHOTO:

ARTIFICIAL INTELLIGENCE

CONTENTS

1. The world of the problem

A mobile telephony company has at its disposal the exact locations of 25 settlements in the prefecture of Halkidiki, whose coordinates are presented in detail in Table 1.1 and the locations of the settlements on the map, in Figure 1.1

The problem concerns the installation location of a single relay station, so that it is cumulatively the shortest possible distance from all the settlements in the area, in a straight line without taking into account the topography of the area.

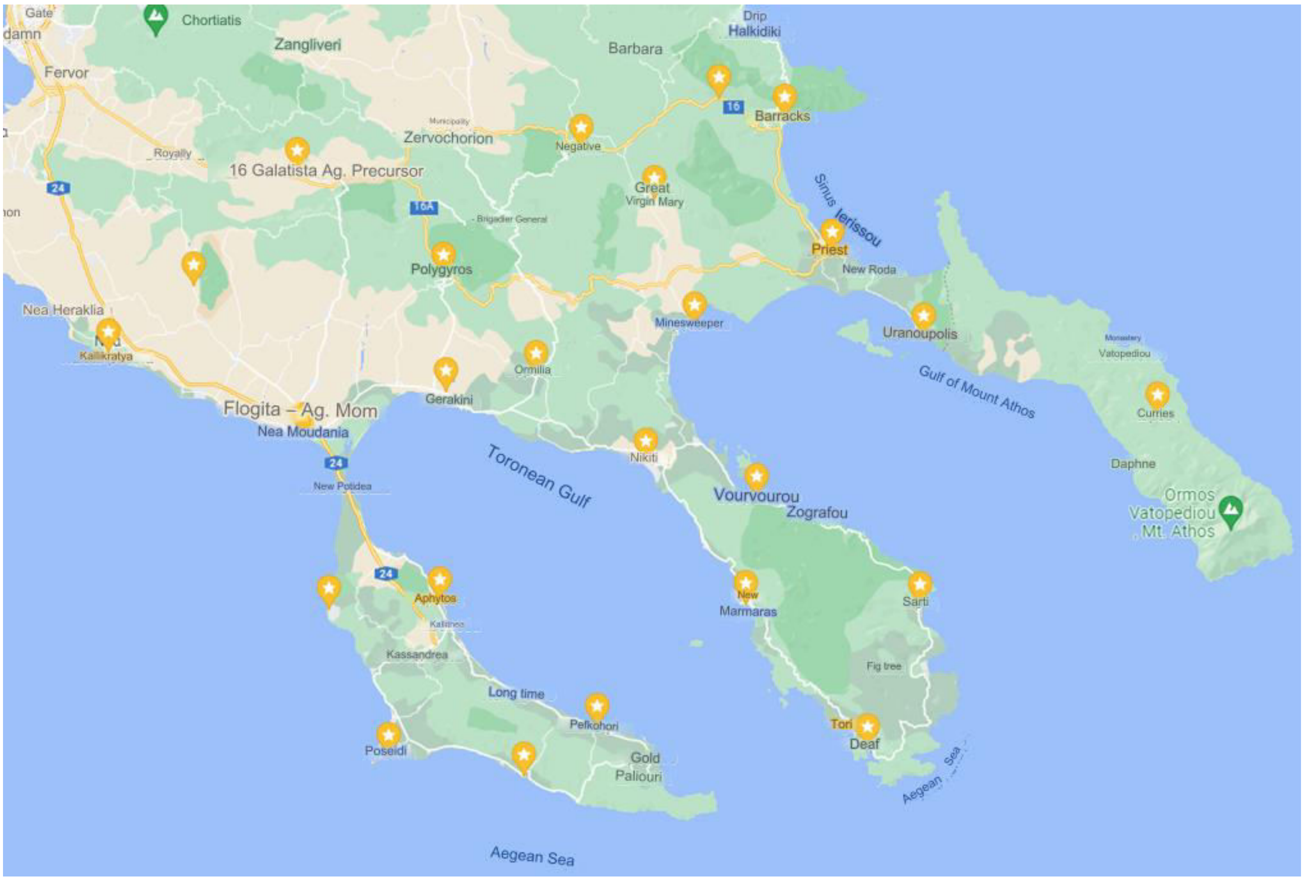

Figure 1.1 The locations of the 25 settlements on the map

Table 1. 1

N/A	City/town/village	Geogr. width	Geogr. length
1	Arnaia	40.487386	23.596722
2	Aphytos	40.097854	23.437297
3	Vourvourou	40.187549	23.794397
4	Galatista	40.467818	23.276774
5	Yerakini	40.277924	23.444131
6	Ierissos	40.398178	23.878921
7	Karyes	40.257182	24.245642
8	Porto Koufo	39.970574	23.919930
9	Megali Panagia	40.444233	23.680327
10	Kallikrateia	40.312016	23.063361
11	Nea Moudania	40.239382	23.283929
12	Nea Skioni	39.946193	23.531817
13	Neos Marmaras	40.095044	23.782260
14	Nikiti	40.217262	23.669855
15	Ormylia	40.293163	23.545274
16	Ouranoupoli	40.325774	23.982267
17	Petralona	40.369312	23.159980
18	Pefkohori	39.988786	23.614820
19	Polygyros	40.377528	23.441502
20	Poseidi	39.963767	23.380025
21	Pyrgadikia	40.334955	23.724034
22	Sani	40.091066	23.312003
23	Sarti	40.093456	23.978882
24	Stagira	40.529959	23.751533
25	Stratoni	40.513395	23.826336

The experiment is repeated with the company having at its disposal the exact locations of 12 settlements in the prefecture of Halkidiki, whose coordinates are presented in detail in Table 1.2 and the locations of the settlements on the map, in Figure 1.2

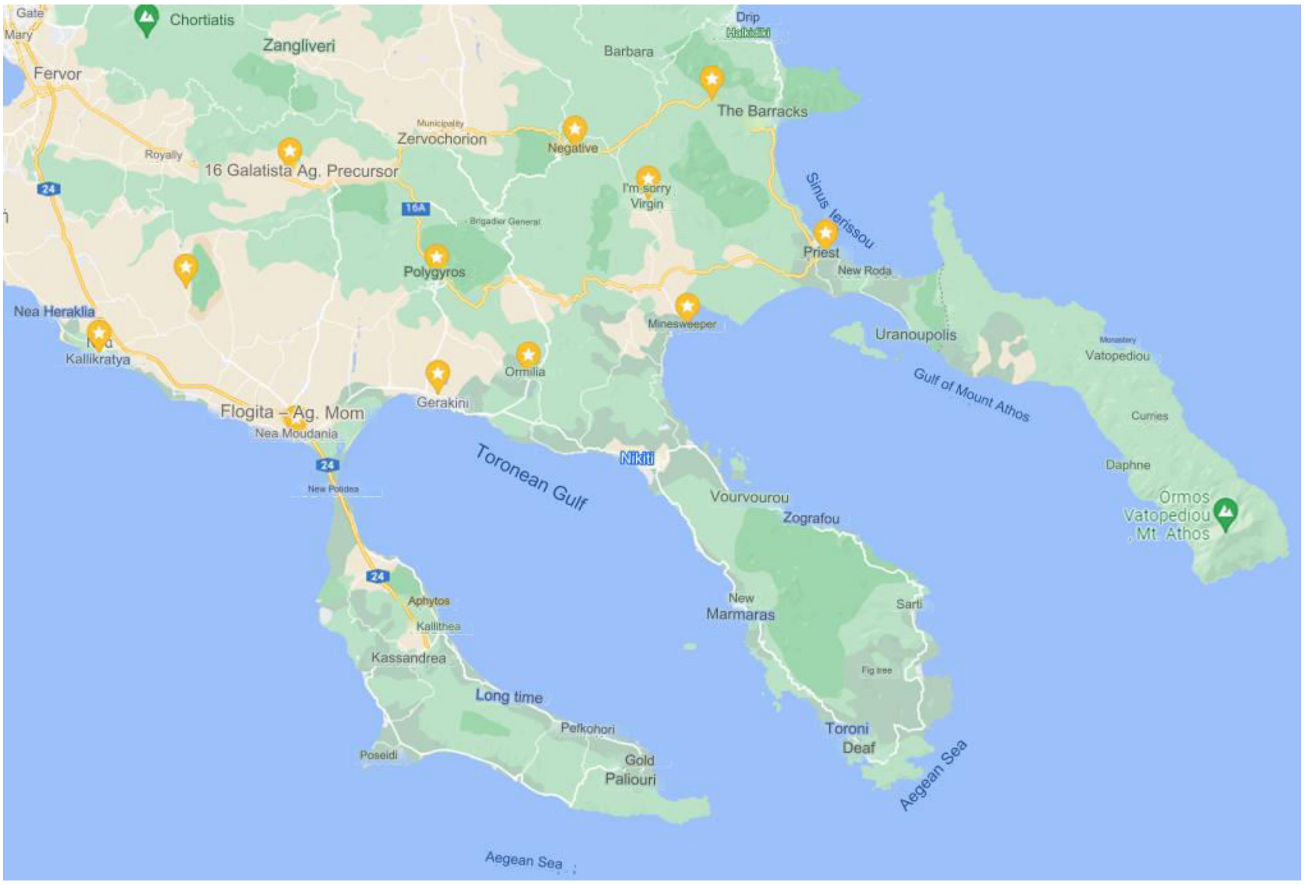

Figure 1.2 The locations of the 12 settlements on the map

Table 1.2

N/A	City/town/village	Geogr. width	Geogr. length
1	Arnaia	40.487386	23.596722
2	Galatista	40.467818	23.276774
3	Yerakini	40.277924	23.444131
4	Ierissos	40.398178	23.878921
5	Megali Panagia	40.444233	23.680327
6	Kallikrateia	40.312016	23.063361
7	Nea Moudania	40.239382	23.283929
8	Ormylia	40.293163	23.545274
9	Petralona	40.369312	23.159980
10	Polygyros	40.377528	23.441502
11	Pyrgadikia	40.334955	23.724034
12	Stagira	40.529959	23.751533

2. The genetic algorithm in application for the 25 settlements

Table 2.1 shows the results of running a simple genetic algorithm, where each line corresponds to a set of parameters. The coordinates of the corresponding region on the map constitute the chromosome, where this chromosome includes two genes, that is, two real values that capture the latitude and longitude of the region. The data for the execution of the algorithm are presented in Table 1.1

The "Population size" column shows the number of possible chromosomes (coordinates of a random region within the map that may be the solution to the problem) that the algorithm randomly selects as a solution to the problem.

The "Crossover Probability" column shows the probability of a gene exchange between two chromosomes producing child chromosomes, that is, the probability of a coordinate exchange between two regions within the map.

The "Mutation probability" column shows the mutation probability of a gene , that is, the probability of modifying the numerical value of a coordinate of a point.

The "Point with minimum distance" column shows the point in the form of coordinates (latitude , longitude), which has the minimum sum of the Euclidean distance ($D = | X - A1 | + \dots + | X - A25 |$) from the 25 settlements.

The "Minimum distance" column displays the minimum sum of the Euclidean distance ($D = | X - A1 | + \dots + | X - A25 |$) that the point with coordinates in the corresponding line of the "Point with minimum distance" column has, of the 25 settlements.

The values of the "Point with minimum distance" and "Minimum distance" columns for each set of parameters are the average of the solutions and distances from 10 runs of the algorithm. This is because genetic algorithms do not give the same results in every run.

Table 2.1

N/A	Population size	Probability of crossover	Mutation probability	Point with minimum distance	Minimum distance
1	10	0.1	0.05	40.25 9471 , 23.616675	7.5 79460
2	10	0.5	0.1	40.258915, 23.616772	7.579374
3	20	0.1	0.05	40.259728, 23.616370	7.579319
4	20	0.5	0.1	40.259423, 23.616924	7.579306
5	50	0.1	0.05	40.259556, 23.616567	7.579276
6	50	0.5	0.1	40.259565, 23.616549	7.579270
7	100	0.1	0.05	40.259412, 23.616587	7.579269
8	100	0.5	0.1	40.259498, 23.616661	7.579265
9	1000	0.4	0.1	40.259465, 23.616770	7.579260

3. The final state of the problem for the 25 settlements

The installation location of the station is calculated from the average of all solutions (Point with minimum distance) highlighted through the results of Table 2.1, the genetic algorithm. The installation point of the station is:

$$\text{AVG Width} = \frac{(\text{Width1} + \text{Width2} + \ldots + \text{Width9})}{9} \rightarrow$$

$$\text{AVG Width} = \frac{(40.259471 + 40.258915 + \ldots + 40.259465)}{9} \rightarrow$$

$$\text{AVG Width} = 40.259448$$

$$\text{AVG Length} = \frac{(\text{Length1} + \text{Length2} + \ldots + \text{Length9})}{9} \rightarrow$$

$$\text{AVG Length} = \frac{(23.616775 + 23.616772 + \ldots + 23.616770)}{9} \rightarrow$$

$$\text{AVG Length} = 23.616653$$

Therefore, the coordinates of the point X with which the relay station can be installed, in order to be the minimum distance from all the settlements, without taking into account the topography of the area (in a straight line) are:

X (40.259448, 23.616653)

Figure 3.1 shows the exact location of the station on the map, while Figure 3.2 shows the locations of the points calculated by the genetic algorithm in Table 2.1, as well as the X coordinates of the station.

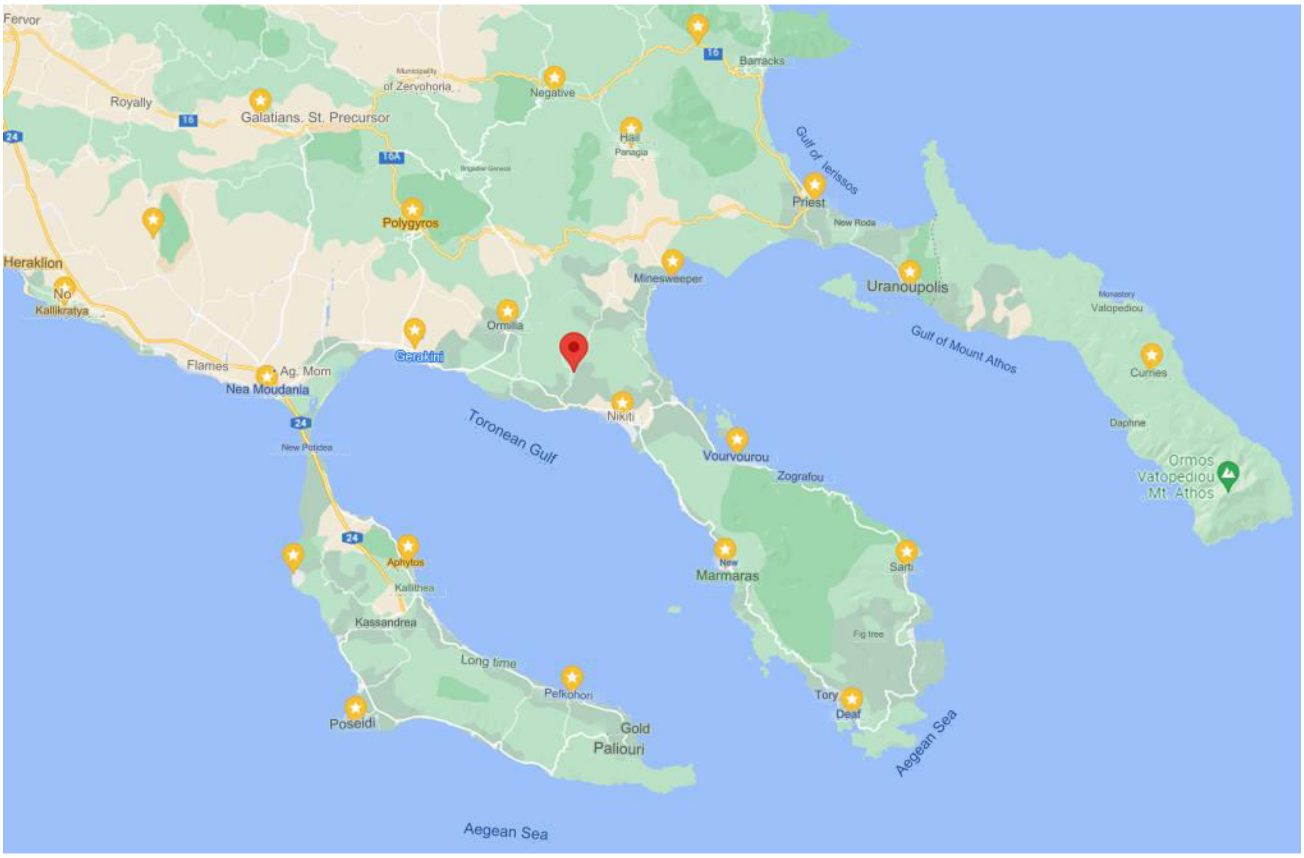

Figure 3.1 The location of relay station X for the 25 settlements

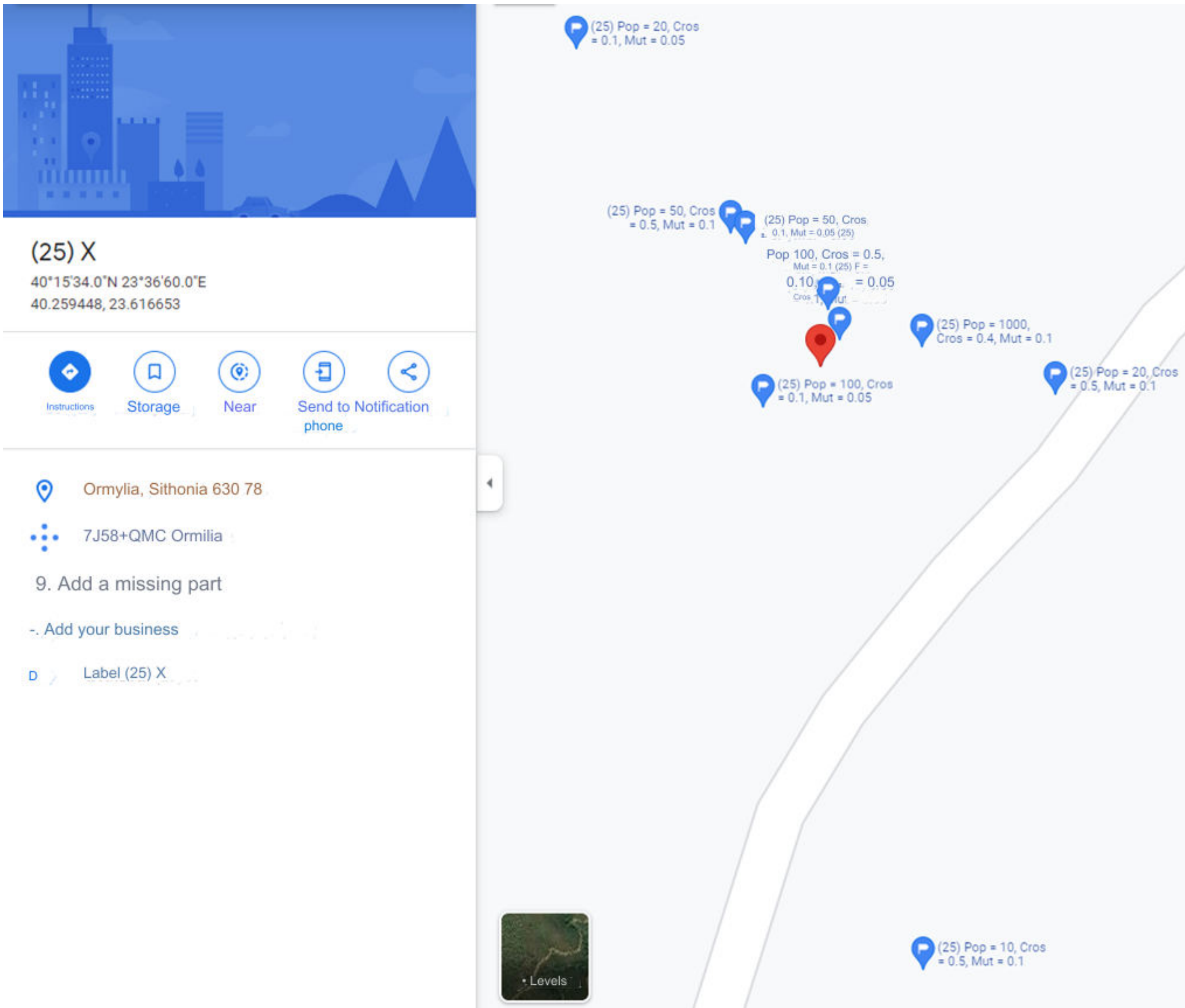

Figure 3.2 The locations of the points of Table 2.1 calculated by the genetic algorithm and the exact location of station X.

4. Commenting on the solutions of the genetic algorithm

The points calculated by the genetic algorithm and listed in Table 2.1, as can be seen in Figure 3.2, are very close to the final point calculated to install relay station X. The annotations target the meaning of the terminologies "Exploitation of the best solutions (exploitation)" and "Exploration of the solution space (exploration)" through the results of the execution of the genetic algorithm. Finally, a commentary is also provided for the intermediate solutions (in terms of the distance to the final station installation point) of Table 2.1 (lines 3 – 9).

Exploitation of the best solutions (exploitation)

The point which is the least distance from the station is the point of its line 1

Table 2.1, with parameter set " M population size : 10", "Crossover probability : 0.1", "Mutation probability : 0.05". The point has coordinates "40.259471, 23.616775" and the minimum sum of the Euclidean distance (D = | X – A 1 | + … + | X – A 25 |) from the 25 settlements is "7.579460". This point is the average of the solutions from 10 runs of the algorithm. The size of the initial population is 10 which means that the possible solutions of the problem to be solved are limited to these few solutions which are randomly selected according to their performance and the probability of their selection by the forced roulette method. Obviously, the chance of crossing over and the chance of mutation are very small with the consequence that the parent chromosomes have a lower chance of producing children of the next generation that have different genetic material. In other words, with a "small" population size and "small" mutation and crossover probabilities, the possible solutions that the algorithm calculates are those that are closest to the final state (which are also the best), thus limiting the exploration of points in the space of solutions that are not the best. The result is the "greed" of the algorithm to rely only on the best solutions, ignoring the rest, which justifies that the point with this set of parameters is the least distance from station X (final state).

Exploration of the solution space

The point with the maximum distance from the station is the point in line 2 of Table 2.1, with parameter set " M population size : 10", "Crossover probability : 0.5", "Mutation probability : 0.1". The point has coordinates "40.258915, 23.616772" and the minimum sum of the Euclidean distance (D = | X – A 1 | + … + | X – A 25 |) from the 25 settlements is "7.579374". This point is the average of the solutions from 10 runs of the algorithm. The size of the initial population is 10 which means that the possible solutions of the problem to be solved are limited to these few solutions which are randomly selected according to their performance and the probability of their selection by the forced roulette method. Obviously, the crossover probability and the mutation probability are increased relative to the parameter set in the 1st row of Table 2.1, which results in the parent chromosomes being more likely to produce offspring of the next generation that have different genetic material. In other words, with a "small" population size and "large" mutation and crossover probabilities, the possible solutions that the algorithm calculates are also those that are not closest to the final state (which are not the best), thus limiting dominance of the best solutions. The result is that the algorithm "spreads" in the space of solutions using solutions that are not the best, which justifies that the point with this set of parameters is also the maximum distance from station X (final state).

Intermediate solutions

With the logic of the terminology mentioned above, the results of lines "3 – 9" of Table 2.1 are also justified, with the difference that the population size is greater than 10, where this implies a larger solution space. In combination with the size of the population and the crossover and mutation probabilities, the distances of the points from the final are recorded, which are neither the minimum nor the maximum.

5. Show paradoxical solutions?

During the execution of the genetic algorithm, no paradoxical solutions (sea point or point outside the prefecture) were detected beyond some solutions of the initial population which were coordinates of sea points that cannot be considered paradoxical, as they also belong to the state space. Points outside the prefecture were not found anywhere, that is, latitude coordinate with an integer part other than 40 and longitude coordinate with an integer part other than 23. The coordinates contain 6 decimal places which means that the state space (including sea points) is infinite. Therefore, the various solutions depended solely on the decimal part of each coordinate.

6. The genetic algorithm in application for the 12 settlements

Table 6.1 shows the results of running a simple genetic algorithm, where each line corresponds to a set of parameters. The coordinates of the corresponding region on the map constitute the chromosome, where this chromosome includes two genes, that is, two real values that capture the latitude and longitude of the region. The data for running the algorithm are presented in Table 1.2.

The "Population size" column shows the number of possible chromosomes (coordinates of a random region within the map that may be the solution to the problem) that the algorithm randomly selects as a solution to the problem. The "Crossover Probability" column shows the probability of a gene exchange between two chromosomes producing child chromosomes, that is, the probability of a coordinate exchange between two regions within the map. The "Mutation probability" column shows the mutation probability of a gene, that is, the probability of modifying the numerical value of a coordinate of a point. The "Point with minimum distance" column displays the point in the form of coordinates (latitude, longitude), which has the minimum sum of the Euclidean distance ($D = | X - A\ 1 | + \ldots + | X - A\ 12 |$) of the 12 settlements.

The "Minimum distance" column displays the minimum sum of the Euclidean distance ($D = | X - A\ 1 | + \ldots + | X - A\ 12 |$) that the point with coordinates has in the corresponding line of the "Point with minimum distance" column, of the 12 settlements.

The values of the "Point with minimum distance" and "Minimum distance" columns for each parameter set are the average of the solutions and distances from 10 runs of the algorithm. This is because genetic algorithms do not give the same results in every run.

Table 6.1

N/A	Population size	Probability of crossover	Mutation probability	Point with minimum distance	Minimum distance
1	10	0.1	0.05	40.365660, 23.482883	2.780777
2	10	0.5	0.1	40.365028, 23.482983	2.780724
3	20	0.1	0.05	40.365008, 23.483659	2.780708
4	20	0.5	0.1	40.365365, 23.483319	2.780706
5	50	0.1	0.05	40.365201, 23.483477	2.780694
6	50	0.5	0.1	40.365185, 23.483256	2.780694
7	100	0.1	0.05	40.365055, 23.483296	2.780692
8	100	0.5	0.1	40.365238, 23.483311	2.780690
9	1000	0.4	0.1	40.365164, 23.483297	2.780689

7. The final state of the problem for the 12 settlements

The installation location of the station is calculated from the average of all solutions (Point with minimum distance) highlighted through the results of Table 6.1, the genetic algorithm. The installation point of the station is:

$$\text{AVG Width} = \frac{(\text{Width1} + \text{Width2} + \ldots + \text{Width9})}{9} \rightarrow$$

$$\text{AVG Width} = \frac{(40.365660 + 40.365028 + \ldots + 40.365164)}{9} \rightarrow$$

$$\text{AVG Width} = 40.365212$$

$$\text{AVG Length} = \frac{(\text{Length1} + \text{Length2} + \ldots + \text{Length9})}{9} \rightarrow$$

$$\text{AVG Length} = \frac{(23.482883 + 23.482983 + \ldots + 23.483297)}{9} \rightarrow$$

$$\text{AVG Length} = 23.483276$$

Therefore, the coordinates of the point X with which the relay station can be installed, in order to be the minimum distance from all the settlements, without taking into account the topography of the area (in a straight line) are:

X (40.365212, 23.483276)

Figure 7.1 shows the exact location of the station on the map, while Figure 7.2 shows the locations of the points calculated by the genetic algorithm in Table 6.1, as well as the X coordinates of the station.

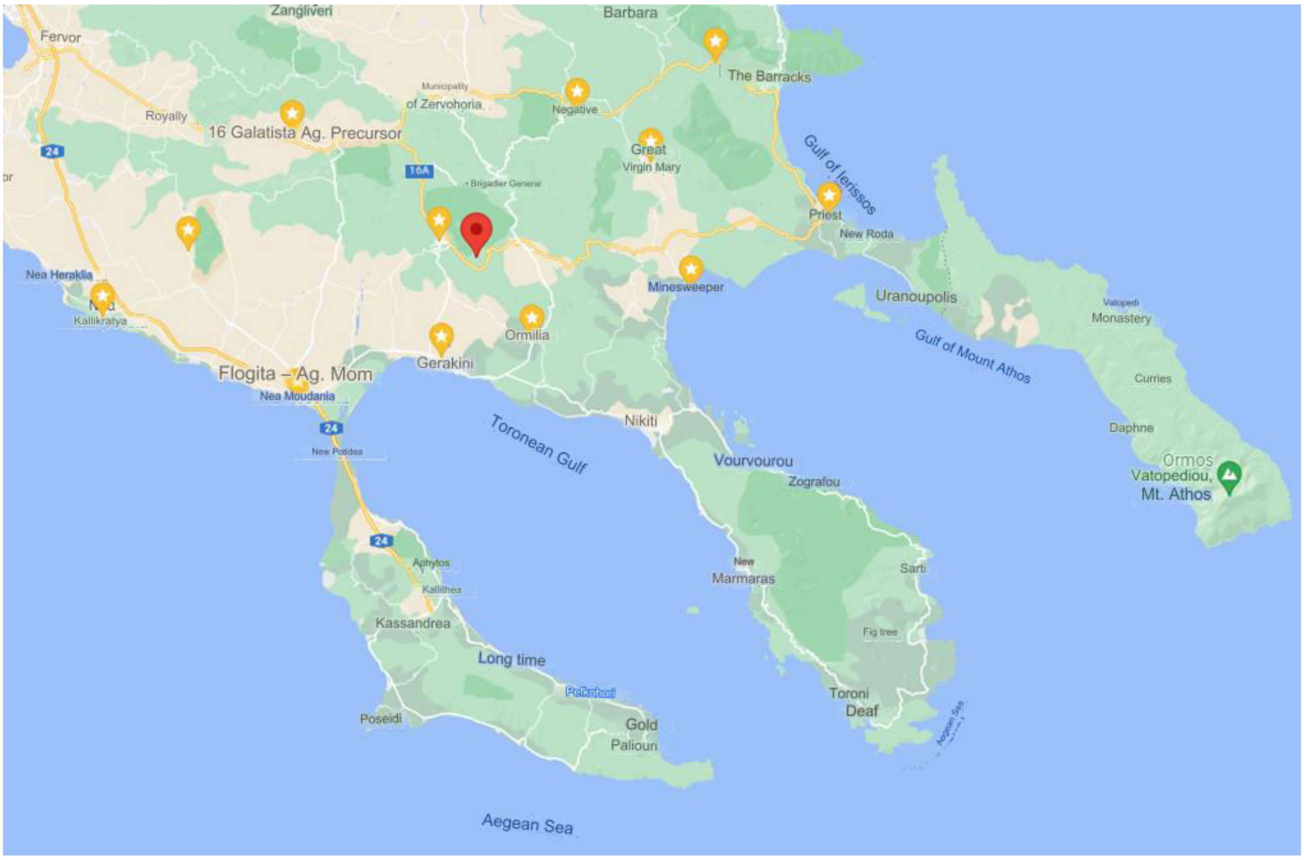

Figure 7.1 The location of relay station X for the 12 settlements

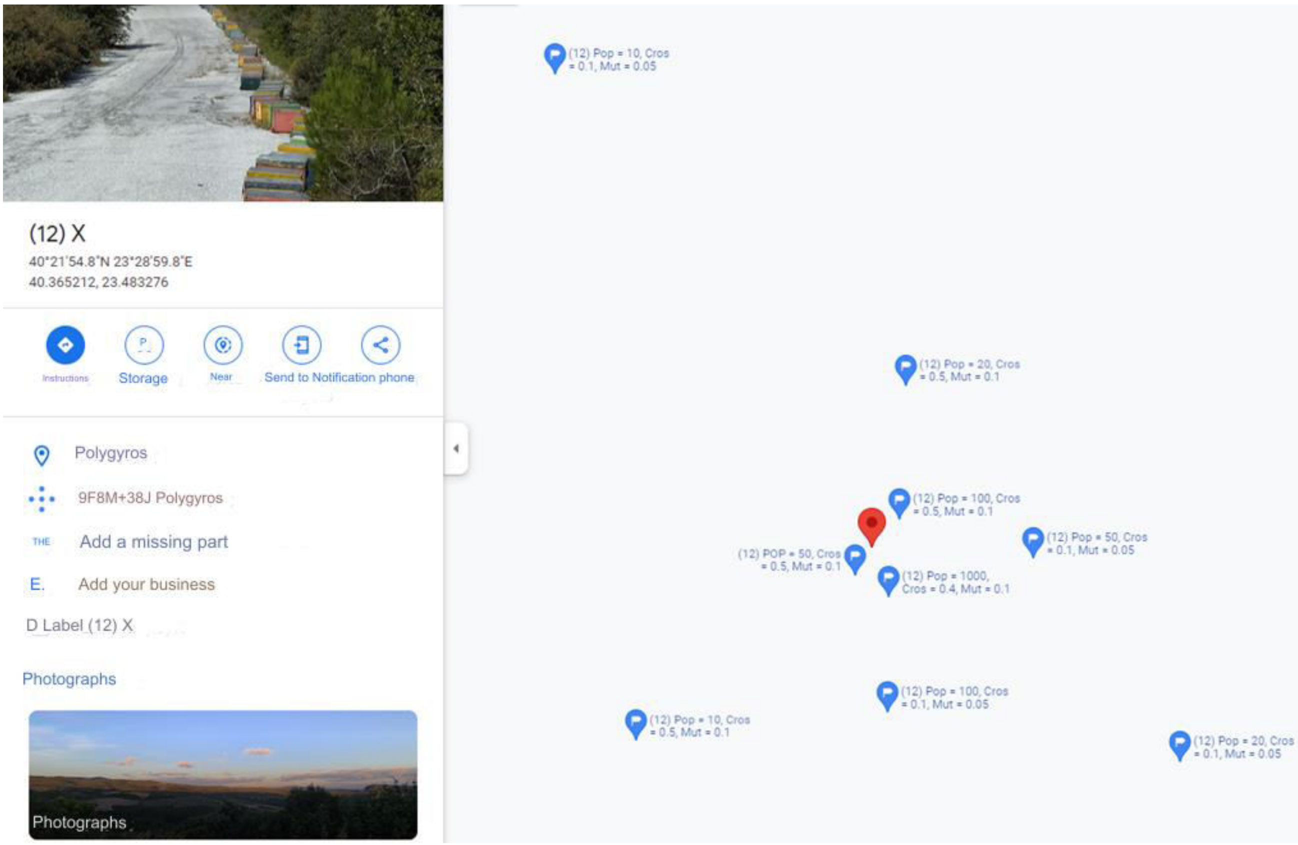

Figure 7.2 The locations of the points of Table 6.1 calculated by the genetic algorithm and the exact location of station X.

8. Differences in the solutions of the two problems?

The solutions for the problem with the 25 settlements differ from the corresponding ones for the 12 settlements in terms of the minimum distances, that is, in terms of the Euclidean distance that separates the point from the N (12 or 25) settlements. In more detail, the distance for the problem with the 25 settlements ranges from 7.579260 – 7.579460, while the corresponding one for the 12 settlements ranges from 2.780689 – 2.780777. The difference is due to the population of the settlements and has as a consequence that the solutions have a difference in terms of decimal places, as the state space differs for each problem.

Thank you for your attention.