THE USV ANNALS OF ECONOMICS AND PUBLIC ADMINISTRATION | VOLUME 21, ISSUE 1(33), 2021

USING GENETIC ALGORITHMS FOR A REGIONALIZATION PROBLEM

Ionut BALAN
"Ştefan cel Mare" University of Suceava, Romania
ionut.balan@usm.ro

Received 12 March 2021; Accepted 29 May 2021

Abstract:

In this paper I want to present the results obtained by using an algorithm in calculating locations for regional hospitals that could be built in Romania. The idea of regional hospitals in Romania have been discussed for some time in the local press, the final locations being established, by those responsible, in the first 8 cities, taken according to the number of inhabitants, as well as a city in the center of the country. In this approach we use a genetic algorithm, a method inspired by natural evolution, which will give us a series of alternative results to the solution proposed by officials. As in the official version, in the used method, to obtain the results we will take into account the population, but not only within the most populated localities, but within a larger number of localities from the country. The population of these localities, weighted with the distance between them and the possible solutions will contribute to the calculation of the final solution.

Key words: genetic algorithm, regional, chromosomes, mutations, cross-overs, optimization

JEL classification: I14, I15, R58,L86

1. FROM REGIONALIZATION TO REGIONAL HOSPITALS

According to the Council of Europe (v.), regionalization is the process of transferring power from the central government to the regions, for a better application of the principle of subsidiarity, within the framework of national or federal solidarity. This includes setting up, expanding or empowering the authorities and transferring competences and responsibilities to the regions. The Governance Committee and the Chamber of Regions of the Congress are responsible for legal and political issues related to regionalization in the member states of the Council of Europe. They shall examine, in particular, developments concerning the institutional and administrative organization of the regions, their competences and their financial autonomy.

From a statistical point of view, the region can have two distinct meanings (it can refer to a geographical area) (iii.):

- at sub-national level, a subdivision of a country, at different possible levels, but most frequently at level 1 of the Nomenclature of Territorial Units for Statistics (NUTS-„Nomenclature Unites Territoriales statistiques");

- at supranational level, a region of the world such as "Latin America", "Sub-Saharian Africa", "South-East of Asia" etc., or various organizations at the level of world regions: e.g. AU, ASEAN, CAN, NAFTA, EU.

Within the European Union, there are currently 240 euro-regions (ii.), created on the basis of NUTS, a nomenclature that ensures their comparability at European level. Being a member state of the European Union, 8 development regions were created in Romania: North-West (RO11), Center (RO12), North-East (RO21), South-East (RO22), South-Muntenia (RO31), Bucharest-Ilfov (RO32), South-West Oltenia (RO41) and West (RO42). Starting from this structure, the representatives of the Romanian Ministry of Health launched the idea of building a larger hospital located in the RO32 region (the most developed in Romania, with a higher population density) and 8 other regional hospitals to serve population from other regions in Romania. These hospitals would be located in Iaşi (RO21), Cluj (RO11), Craiova (RO41), Braşov (RO12), Timişoara (RO42), Târgu-Mureş (RO12), Brăila / Galaţi (RO22) and Constanţa (RO22). A closer look at the chosen cities can be seen that 8 of the nine hospitals will be built in the first 8 cities, taken from the point of view of the number of inhabitants, the only city that does not enter in this top being Târgu Mureş. Health regionalization has been adopted by several countries to improve people's access to health services (Rashidian et al., 2014; Lin and Lee, 2021). The World Health Organization (WHO) defines it as the rational distribution of medical services throughout the territory, ensuring that services and facilities provide all levels of care (primary, secondary and tertiary) with easy access to the population and cost-effective care (Ramos et al., 2020). However, in some cases, this decision may lead to a feeling of inequality in the geographical distribution of future constructions, but also other inconveniences with direct or indirect results in many areas (Chavehpour et al.' 2019):

- the RO31 region has no hospital assigned, this being due to the location of a hospital in the RO32 region, the latter being in the center of the RO31 region, thus attracting its inhabitants;
- to the most populated region (RO21) it is assign only one hospital, while two other regions with a lower population have two hospitals each;
- the chosen cities are overcrowded, they already have hospital units to serve the local population;
- we reach the increase of the gaps between the cities from the point of view of the units assigned to the sanitary system, which would negatively influence the economy, education, population density, etc.

Next, we will present the chosen method for finding locations for the 9 proposed hospitals, locations that would remove the shortcomings presented above, followed by the presentation of input data for the chosen method, the results obtained and their comparison with those proposed by the authorities, but also the conclusions I drew from this research.

2. USED METHOD

In the 1950s and the 1960s several computer scientists independently studied evolutionary systems with the idea that evolution could be used as an optimization tool for engineering problems. The idea in all these systems was to evolve a population of candidate solutions to a given problem, using operators inspired by natural genetic variation and natural selection (Mitchell, 1996).

Genetic algorithms are a family of computational models inspired by evolution. These algorithms encode a potential solution to a specific problem using a chromosome (often expressed by an array) and apply recombination operators to these structures (chromosomes) so as to preserve critical information about them (Whitley, 1994).

Holland's 1975 book Adaptation in Natural and Artificial Systems presented the genetic algorithm as an abstraction of biological evolution and gave a theoretical framework for adaptation under the GA. Holland's GA is a method for moving from one population of "chromosomes" to another population by using some natural−inspired operators of crossover, mutation, and inversion. Each chromosome consists of "genes", each gene being an instance of a particular "allele". The selection operator chooses those chromosomes in the population that will be allowed to reproduce, and on average the fitter chromosomes produce more offspring than the less fit ones. Crossover exchanges subparts of two chromosomes; mutation randomly changes the allele values of some locations in the chromosome; and inversion reverses the order of a contiguous section of the chromosome, thus rearranging the order in which genes are arrayed (Mitchell, 1996).

For the case study presented, the chromosome has 10 genes, the first 9 genes being the city codes from the database, the tenth being used to store the value of an objective function, a value that influences the selection of chromosomes used in evolution.

The operators that I use in this algorithm are mutation and crossover. For both operators I use an elitist strategy for selecting the participating chromosomes (Balan, 2014). Thus, in the case of mutation, I randomly generate a chromosome, one of the best in the current population, as well as a gene (from 0 to 8, values that indicate the position of the gene on the chromosome). The value of that gene is replaced by another value (city id) that does not exist in the current chromosome. In figure 1 we have a graphical representation of this process.

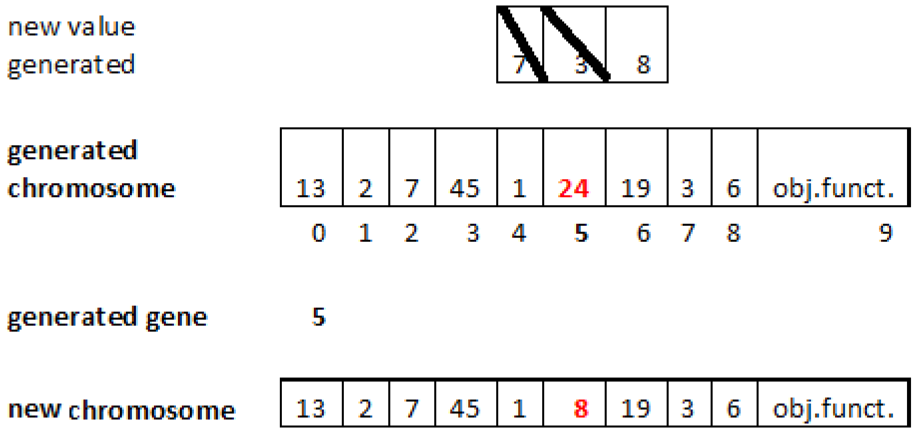

Figure 1. Mutation operation

For the crossover operator, for the selection of chromosomes I applied the same elitist strategy. After generating two chromosomes and one gene, I obtain a child chromosome as in figure 2.

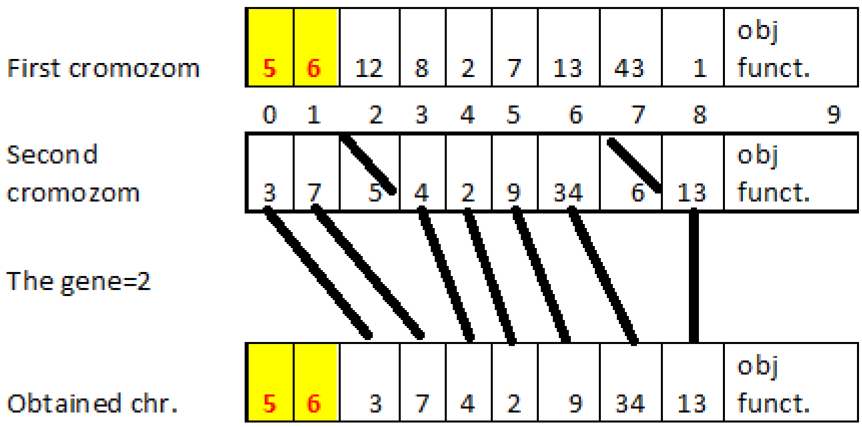

Figure 2. Crossover operation

3. COMPUTATIONAL DATA USED

The genetic algorithm used to solve the proposed problem, uses a population of 100 chromosomes, a population that has been evolving for 100 generations. In each generation I apply 20 mutations and 10 crossovers. The first 9 genes of each chromosome are city ids from Romania. The number of cities, which I consider, is 52, being cities with a population of over 40,000 inhabitants (iv.). To calculate the objective function, the one I want to optimize, I also use the distances between cities (vi.,i.).

A first step in the algorithm used is the generation of the initial population. The IDs that will represent the genes of each chromosome will be between 1 and 52 (1 being Bucharest, the city with the largest population, and 52 is Mangalia, a city with a population over 40,000 inhabitants located in the RO22 region). According to statistical data, the population of Romania in 2016 was 22,241,718. The population of the 52 cities considered reaches 8,761,689 inhabitants, representing over 39% of the total population.

A particular attention must be paid to the calculation of the objective function. The activity of building a function, which optimized will lead to obtaining good results for the studied problem, is decisive. In this case, when calculating the value of this function I use the population of cities, weighted by a value that depends on the distance between that city and a city from a current solution. To calculate the objective function, I introduced the notion of neighborhood. Thus, if the distance between two cities is less than or equal to 200 km, then I consider them neighbors. Knowing this, we can say that the value of the objective function is calculated using the pseudocode:

```
int val = 0;
for (int i = 0; i < 9; i++) {
    for (int j = 0; j < id; j++) {
        val += pop[i].getDifPondCum(j) * loc(j).getLoc();
    }
}
return val;
,where:
-val is the function value;
-loc(j).getLoc()-return number of people from j city;
-pop[i].getDifPondCum(j):
    - return 1, if i is equal with j;
    -otherwise, return 1-1/(distance between i and j), with
    the specification that the sum for the values returned
    for the same city j within a chromosome cannot exceed the
    value 1 (if it reaches 1, the returned value will be 0).
```

After the generation of the initial population, in each subsequent generation the crossover and mutation operators will be applied. These two operators, together with an elitist approach, related to the population that will survive, will lead to obtaining satisfactory results for the studied case. To observe the gain from using this algorithm, we should study the space of possible solutions. It consists of C_{52}^{9} solutions, which actually means approx 3,6*10^9 solutions. The complexity of this space of solutions forces us to move as far as possible from the exact algorithms and to move to methods that offer us sub-optimal solutions(for example, genetic algorithms).

4. COMPUTATIONAL RESULTS OBTAINED

Based on the generation of random numbers, the method I described does not offer the same solution every run. From the multitude of variants obtained, I extracted 5, which I will compare with the solution offered by the officials. To make the comparisons, we randomly generated 10,000 inhabitants. A inhabitant who is part of a city with a large population has a higher probability of generation, this probability decreasing at the same time as the number of inhabitants decreases.

In the following table I presented 5 results obtained using the described algorithm, but also the solution offered by the authorities:

Table 1. Obtained results

Oficial solution		1st solution		2nd solution		3th solution		4th solution		5th solution	
Id	City	Id	City	Id	City	Id	City	Id	City	Id	City
1	Bucuresti	40	Roman	7	Brasov	5	Constanta	5	Constanta	4	Cluj-Napoca
2	Iasi	8	Galati	4	Cluj-Napoca	1	Bucuresti	4	Cluj-Napoca	5	Constanta
3	Timisoara	3	Timisoara	6	Craiova	40	Roman	8	Galati	7	Brasov
4	Cluj-Napoca	1	Bucuresti	3	Timisoara	7	Brasov	6	Craiova	14	Sibiu
5	Constanta	14	Sibiu	34	Deva	4	Cluj-Napoca	14	Sibiu	40	Roman
6	Craiova	37	Sfantu	40	Roman	12	Arad	40	Roman	3	Timisoara

			Gheorghe								
7	Brasov	36	Zalau	8	Galati	34	Deva	37	Sfantu Gheorghe	1	Bucuresti
8	Galati	5	Constanta	5	Constanta	6	Craiova	3	Timisoara	8	Galati
16	Tg Mures	6	Craiova	1	Bucuresti	11	Braila	1	Bucuresti	6	Craiova

Centralizing the results from table 1, figure 3 was obtained, a figure that represents the frequency of occurrence of cities in the described solutions. It is observed that certain cities appear in all 5 solutions (Roman, Bucharest, Constanta, Craiova), in other cases, where we have two cities, geographically close (Galati-Braila, Timisoara-Arad), the results indicate a large difference between frequencies of the appearance of the cities in each case, and two of the cities in the variant of the authorities (Iaşi and Târgu Mureş) do not appear at all, the geographical position and the neighborhoods having a hard word to say in such cases.

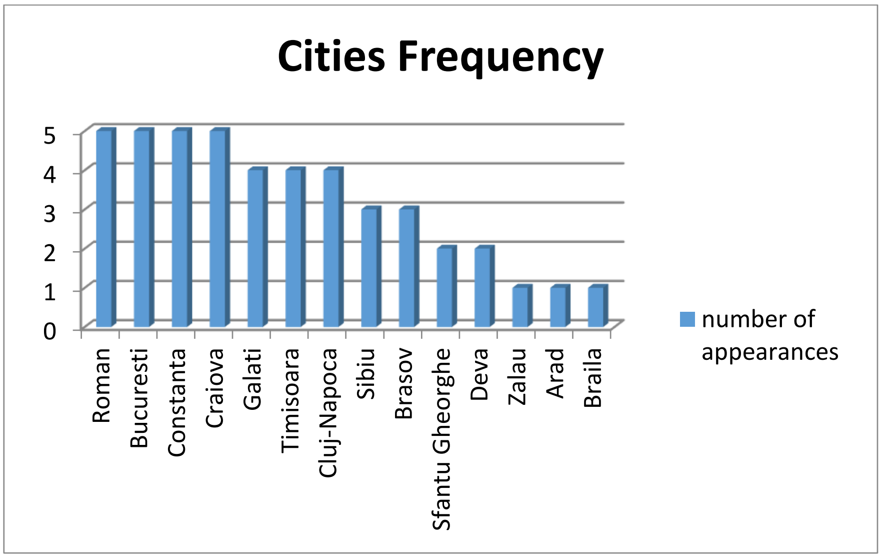

Figure 3. Frequency of occurrence of cities in solutions

A first problem that I mentioned in the first part of the paper refers to the fact that there are certain regions without any hospital in the version found by the officials. In table 2, I present a centralization, by regions, of the solutions that were presented in table 1.

Table 2. Number of hospitals/region

	Oficial solution	**1st solution**	**2nd solution**	**3th solution**	**4th solution**	**5th solution**
RO11	1	1	1	1	1	1
RO12	2	2	1	1	2	2
RO21	1	1	1	1	1	1
RO22	2	2	2	2	2	2
RO31	0	0	0	0	0	0
RO32	1	1	1	1	1	1
RO41	1	1	1	1	1	1
RO42	1	1	2	2	1	1

It is observed that, even in the variants found using the algorithm described, the RO31 region remains without a hospital, the explanation from the first chapter being able to be taken into account in the case of all the obtained solutions. Related to the second problem, the one about the populated region that has only one allocated hospital, through the offered solutions a small part of the problem was solved. If in the initial version, in that region, a locality (with 362,000 inhabitants) located at the edge of the region was proposed, it was replaced, in all the solutions that were found, by a locality that has around 70,000 inhabitants (Roman). The great advantage of the smaller locality is that it is located in the center of the region, being accessible to a large part of its population.

The last two mentioned problems are solved by the appearance, in the final solutions of some sparsely populated localities (less than 200,000 inhabitants), such as: Roman (5times/5solutions), Sibiu (3/5), Sfantu Gheorghe (2/5), Zalau (1/5), Deva (2/5), Arad (1/5)

We validated these solutions by randomly generating 10,000 inhabitants from the cities from our database, the probability that a citizen will be among the 10,000 generated depending on the number of inhabitants for the city they belong to. Thus, the larger it is, the more likely this inhabitant is to be generated. The obtained results are presented in table 2. And in this case we performed 5 runs, because behind the tests is the random number generator, which leads to different results on each run.

Table 3. Final tests

	Oficial solution	**1st solution**	**2nd solution**	**3th solution**	**4th solution**	**5th solution**
1st run						
Total number of km (solution city - inhabitant)	506063	487058	484869	488103	484963	**477718**
Number of inhabitants exceeding 100 km	3105	**2314**	2550	2356	2509	2561
Maximum distance per inhabitant	187	**163**	187	187	187	187
2nd run						
Total number of km	498500	490041	483053	486086	483267	**475077**
Number of inhabitants exceeding 100 km	3045	**2280**	2559	2339	2479	2538
Maximum distance per inhabitant	187	**163**	187	187	187	187
3th run						
Total number of km	504853	491003	487846	494576	490150	**481050**
Number of inhabitants exceeding 100 km	3094	**2230**	2547	2328	2511	2561
Maximum distance per inhabitant	187	**163**	187	187	187	187
4th run						
Total number of km	504602	493129	483785	488131	486819	479154
Number of	3112	**2323**	2555	2348	2528	2572

inhabitants exceeding 100 km						
Maximum distance per inhabitant	187	**163**	187	187	187	187
5th run						
Total number of km	494032	478944	475905	483202	476054	**469194**
Number of inhabitants exceeding 100 km	3078	**2220**	2504	2304	2470	2529
Maximum distance per inhabitant	187	**163**	187	187	187	187

From table 3, it is observed that the solution offered by the authorities is not the best considering the indicators we have treated. Thus, apart from the last indicator, where I obtained a value equal to 4 other solutions found through the proposed algorithm (= 187 km), in the case of the other 2 indicators the values obtained in the proposed by the authorities are lower than the values obtained in all 5 solution variants.

CONCLUSIONS

Genetic algorithms are methods that are successfully applied to complex problems, for which exact methods would have no solution. They apply to those combinational problems that involve a huge search for solutions. In our problem, a linear increase in input data (in our case, the number of cities), leads to an exponential increase in the number of possible solutions (those found in the solution space). For this problem, we applied this method and obtained better solutions than the solution recommended by the Romanian authorities regarding the future locations used for the construction of large hospitals, at regional level. For the present problem, the results obtained by applying the algorithm on a larger set of cities (maybe even all of them exceeding 5000 inhabitants) would become interesting. The same problem can be studied from a multicriteria perspective, but also by introducing new constraints, or even changing the objective function by introducing new features in its calculation. In this case we discussed regional hospitals, but the same approach can be used in any other issue that involves an organization at the regional level, such as: the establishment of logistics centers, issues related to tourism strategies, transport, education, military etc.

REFERENCES

1. Balan, I., (2014), *Parallel Hybrid Methods Used In Optimization Problems Solving*, The USV Annals of Economics and Public Administration 14 (2 (20)),pp.168-176.
2. Chavehpour, Y., Rashidian, A., Woldemichael, A., Takian, A., (2019), *Inequality in Geographical distribution ofhospitals and hospital beds in densely populated metropolitan cities of Iran*, BMC Health Services Research 19.
3. Lin, C.-Y., Lee, Y.-C.,(2021), *Effectiveness of hospital emergency department regionalization and categorization policy on appropriate patient emergency care use: a nationwide observational study in Taiwan*, BMC Health Services Research volume 21, Article number: 21.
4. Mitchell, M., (1996), *An Introduction to Genetic Algorithms*, Cambridge, MA: MIT Press. ISBN 9780585030944.

5. Ramos, M.C., Barreto, J.O.M., Shimizu, H.E.,Moraes, A.P.Gd., Silva, E.Nd, (2020), *Regionalization for health improvement: A systematic review*. PloS ONE 15(12): e0244078. https://doi.org/10.1371/journal.pone.0244078
6. Rashidian, A., Omidvari, A., Vali, Y., Mortaz, S., Yousefi-Nooraie, R., Jafari, M., et al, (2014), *The effectiveness of regionalization of perinatal care services: a systematic review*. Public Health; 128(10):872–85
7. Whitley, D., (1994), *A genetic algorithm tutorial*, Statistics and Computing. 4 (2): 65–85. CiteSeerX 10.1.1.184.3999.,doi:10.1007/BF00175354
8. https://distanta.ro/ [accessed 10.04.2021]
9. Eurostat, https://ec.europa.eu/eurostat/cache/digpub/regions/ [accessed 10.04.2021]
10. Eurostat, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Region [accessed 10.04.2021]
11. Populaţia României pe localităţi, Institutul Naţional de Statistică, 2016, https://insse.ro/cms/sites/default/files/field/publicatii/populatia_romaniei_pe_localitati_la_1ianuarie2016_0.pdf [accessed 10.04.2021]
12. Council Of Europe, https://www.coe.int/en/web/congress/regionalisation [accessed 10.04.2021]
13. https://www.tourismguide.ro/x/distante_rutiere/ [accessed 10.04.2021]