Problem of the P-Stockings applied to the location of facilities

F Mesa¹, D Devia¹, and R Ospina²

¹ Departamento de Matemáticas, Universidad Tecnológica de Pereira, Pereira, Colombia
² Escuela de Física, Universidad Industrial de Santander, Bucaramanga, Colombia

E-mail: femesa@utp.edu.co

Abstract. One of the problems derived from the overcrowding of cities and industry is the location of optimal facilities, these types of problems suffer from the problem of dimensionality themselves. In the following article, the problem is formulated in a mathematical way and it is approached from a metaheuristic epistemology which is the only methodology for now to find local optimum, since these optimization problems are still open in the scientific literature because they are considered problems of nondeterministic polynomial time. The explored technique is that of genetic algorithms which uses the classical theory of evolution to formulate and solve optimization problems.

1. Introduction

An important line of research in the problems in research of operations or optimization has to deal with the problem of dimensionality, that is, non-polynomial problems which still do not have an analytical solution and are open to the scientific community. These issues address the uncertainty of locating a facility [1]. The facility must interact with other objects that will have fixed locations. In a location problem, you want to establish the sites where factories, warehouses, hospitals, schools, etc. are installed. with respect to a set of spatially distributed demand points [2].

The middle p problem is a classical location problem [3]. The objective is to determine the location of the facilities p in a network of nodes and to minimize the sum of the distances between each demand node [4]. The problem to be addressed involves the location of the facilities p and the assignment of clients to the facilities to minimize the sum of the distances from the clients to the facilities so that each client is served by a single facility [5].

The importance of this research lies in trying to contribute to the scientific community the line of research based on metaheuristic techniques and provide it with mathematic arguments that make them sustainable, many of the criticisms that this type of techniques receives are oriented on local solutions and not global, but using schema theory we can provide a formalism that allows us to guarantee global optimum.

2. Content

Before performing the implementation we define the following notation: Given a complete and non-oriented graph G = (V; E), where V is the set of n vertices and E is the set of edges that represent the distances between the vertices, the objective is to find a subset Vp ∈ V with cardinality p, where Vp represents the set of medians of the problem, so that the sum of the distances between the remaining vertices {V – Vp} and the nearest vertex in Vp is the smallest possible [6]. Considering that:
• \( n \): represents the number of vertices.
• \( p \): represents the number of medians to be installed.
• \( d_{ij} \) represents the distance between vertices \( i \) and \( j \);

The mathematical formulation of the \( p \)-medium problem is given by Equation (1), Equation (2), Equation (3), Equation (4), and Equation (5).

\[
\min \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij}x_{ij}, \tag{1}
\]

subject to:

\[
\sum_{j=1}^{N} x_{ij} = 1 \quad \forall \ i \in \{V - V_p\}, \tag{2}
\]

\[
x_{ij} \leq y_j \quad \forall \ i \in \{V - V_p\} \quad y_j \in V_p, \tag{3}
\]

\[
\sum_{j=1}^{N} y_j = p \quad \forall j \in V_p, \tag{4}
\]

\[
x_{ij}, y_j \in \{0,1\}, \tag{5}
\]

Where \( x_{ij} \) and \( y_j \) are defined by Equations (6) and Equation (7).

\[
x_{ij} = \begin{cases} 1, & \text{si el } i - \text{ésimo vértice es atendido por la mediana (vértice) } j \\ 0, & \text{en caso contrario} \end{cases}, \tag{6}
\]

\[
y_j = \begin{cases} 1, & \text{si el } j - \text{ésimo vértice es una de las } p \text{ - medianas} \\ 0, & \text{en caso contrario} \end{cases}. \tag{7}
\]

The restriction Equation (2) ensures that a vertex (client) is attended by a single median (facility) the restriction Equation (3) establishes that a vertex (client) is only answered by a median that is installed. Restriction Equation (4) ensures that exactly \( p \) medians are designated. Restriction Equation (5) defines that the decision variables involved are binary [7]. The functions were declared as follows [8].

2.1. Read
Function that reads the input file. The output is a matrix. Where each line contains the coordinates \((x; y)\) of a vertex except the first one contains \( n \) and \( p \) values.

2.2. Fitness
This function calculates the suitability of a candidate solution. For this, each vertex is associated with the nearest installation according to the Euclidean distance between them. Returns the total distance, which is used most during the execution of the algorithm, but also returns vectors with the order of the vertices, the median is associated with each one and the distance of each vertex is its median.

2.3. Roulette
Use the roulette strategy to define the parents that will cross to generate a new population. Individuals are assigned one by one in the segment vector in the range \([0;1]\), and the size of each segment is proportional to the probability of individual selection. Then a random number is selected, and this number is searched until it is in one of the segments of the range. This process is analogous to spinning a roulette with a single selection point. Rotate the wheel \( \lambda \) times to select \( \lambda \) individuals that will cross to generate a new population.
2.4. Cross
For the crossover, a Table 1-bit mask is applied, like the example in Table 2. The children then receive the characteristics according to the following model [9]. The first and the second line are the parents, in the third line we have the bitmask, and in the fourth and fifth line are the children. They are generated according to the mask: if 0 in the mask, the first child receives information from the second father and case 1 receives information from the first father; in the second child, case 1 in the mask masks the first child receives information from the second father and, if it is 0, receives information from the first father [10].

| Table 1. Bit mask. |
|-------------------|
| 0 1 1 0 1         |

Table 2. Individuals.

| 46 152 173 216 257 |
|-------------------|
| 58 154 199 234 275 |
| 0 1 1 0 1          |
| 58 152 173 234 257 |

2.5. Mutation
The mutation is performed by exchanging a value of the column vector, verifying if the chosen random vertex no longer exists in this candidate solution [11]. The evolutionary strategy implemented to solve the p-medium problem the proposal follows the following Algorithm 1 [12].

**Algorithm 1.** Algorithm implemented, P-median.

T $\leftarrow$ 1

Initialize population $P_t = \{n_{i,1} i = 1, 2, ..., p\}$

for $i = 1 : n$ do

Calculate fitness

while some stopping criteria is not satisfied do

Select roulette parents

for $i = 1 : n$ do

Cross

Mutation

Validation $\mu + \lambda$

Select two most suitable

Return $P$ – median

3. Results
The first test was carried out with the following parameters: population size ($n = 30$), crossover probability (crossP rob = 1) and mutation probability (crossMut = 0: 9). For this test, a spatially distributed set of points of size $n = 324$ is used and you want to establish $p = 5$ (medium) locations Figure 1.

Figure 1. Test 1 $n = 30$. 
The result after just over 300 interactions and a fitness $= 1.2252 \times 10^5$, with monotonic convergence as shown in Figure 2.

![Figure 2. Convergence test 1.](image)

The best individual never worsens as is the rest of the population, which is possible thanks to the $\mu + \lambda$ strategy implemented. In the second test, the initial parameters of the previous test are repeated, and the number of demand points is increased for $n = 818$. In this test it can be seen that the strategy could detect some groups that can be confirmed visually Figure 3.

![Figure 3. Test 2.](image)

The result after 300 generations in the algorithm, a fitness value $= 6.0603 \times 10^5$ was found and again monotonic convergence is found Figure 4.

![Figure 4. Convergence test 2.](image)
In the last test, the number of demand points is increased \( n = 3282 \), that is to say more than three times the amount of the previous test and you want to find \( p = 5 \) facilities (medium), Figure 5.

![Figure 5. Test 3.](image)

In this case, there are very clear groups again Figure 6, and it was possible to verify that the proposed evolutionary strategy was able to place them well. The result after just over 200 interactions It was a monotonic convergence best solution with fitness \( = 6.3855 \times 10^6 \), Figure 6.

![Figure 6. Convergence test 3.](image)

4. Conclusions
The results show that the strategy based on genetic algorithms has a good performance even for large inputs. Not many interactions were required for the algorithm to converge on a promising solution. A good choice of parameters. they make the algorithm converge faster and more satisfactorily, so it is essential to know the problem well to decide the value of the parameters. The decision to implement a strategy like this has proven to be an important point of implementation because convergence is guaranteed for a solution as seen in Figure 6.

As pure mathematics has not yet found an answer to this type of problem, the most appropriate thing is to resort to techniques inspired by physics or biology, which can help us to give satisfactory solutions as demonstrated in the research. For locations less than 20 units, it is possible to coincide with the analytical and numerical solutions, but as the number of variables increases, the possible combinations imply a distance between the analytical and numerical solutions.
References

[1] de Andrade R L, Rêgo L C 2019 P-Means centrality Communications in Nonlinear Science and Numerical Simulation 68 41-55
[2] Cunha C, Gaspar A, Takahashi C R, Antunes C 2012 Manual de Computação Evolutiva e Metaheurística (Coimbra: Imprensa da Universidade de Coimbra)
[3] Alp O, Erkut E, Drezner Z 2003 An efficient genetic algorithm for the p-median problem Annals of Operations Research 122(1-4) 21-42
[4] Moreno J, Osorio M 2008 A Lyapunov approach to second-order sliding mode controllers and observers 47th IEEE Conference on Decision and Control (México: IEEE) pp 2856-2861
[5] Mladenović N, Brimberg J, Hansen P, Moreno J 2007 The p-median problem: A survey of metaheuristic approaches European Journal of Operational Research 179(3) 927-939
[6] Adasme P 2018 P-Median based formulations with backbone facility locations Applied Soft Computing 67 261-275
[7] Grine F, Kamach O, Sefiani N 2018 A new efficient metaheuristic for solving the uncapacitated single allocation p-hub median problem International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA) (Morocco: IEEE) pp 69-74
[8] Wang Y, Chen W, Wang J, Wang H 2015 Active global localization based on localizability for mobile robots Robotica 33(8) 1609-1627
[9] Biajoli F, Chaves A, Lorena L 2019 A biased random-key genetic algorithm for the two-stage capacitated facility location problem Expert Systems with Applications 115 418-426
[10] Mirjalili S 2019 Evolutionary Algorithms and Neural Networks (Berlin: Springer)
[11] Led Jiménez S P 2018 Problemas de la P-Mediana (Zaragoza: Universidad de Zaragoza).
[12] Eiben A E, Smith J E 2003 Introduction to Evolutionary Computing (Berlin: Springer Verlag)