

A modified p-median model to optimize the location of facilities considering the source plant

Un modelo p-mediana modificado para optimizar la ubicación de instalaciones considerando la planta matriz

SOLÍS-JIMÉNEZ, Miguel Ángel†*, CALDERÓN-PALOMARES, Luis Antonio, PIÑA-MARTINEZ, Ana Laura and GONZÁLEZ-SOBAL, Martín

Tecnológico Nacional de México-Instituto Tecnológico Superior de Huatusco / Av. 25 Poniente No. 100 Colonia Reserva Territorial, Huatusco. Ver. CP. 94100

Tecnológico de Monterrey, ITESM-Escuela de Ingeniería y Ciencias, Atizapán, Estado de México, México

ID 1st Author: *Miguel Ángel, Solís-Jimenez* / ORC ID: 0000-0002-8125-0989, Researcher ID Thomson: N-6243-2018, CVU CONACYT ID: 94216

ID 1st Coauthor: *Luis Antonio, Calderón-Palomares* / ORC ID: 0000-0001-9846-5567, Researcher ID Thomson: N-6259-2018, CVU CONACYT ID: 238274

ID 2nd Coauthor: *Ana Laura, Piña-Martinez* / ORC ID: 0000-0003-4137-6306, CVU CONACYT ID: 478063

ID 3rd Coauthor: *Martín, Gonzalez-Sobal* / ORC ID: 0000-0003-0038-8319, Researcher ID Thomson: N-6243-2018, CVU CONACYT ID: 463431

DOI: 10.35429/EJDR.2020.10.6.14.21

Received March 28, 2020; Accepted June 20, 2020

Abstract

The present general model of the distribution of the logistics value chain considers the production processes, as well as the economic factors and logistics costs operating in the different stages of a purified water company. The model has a general application for companies that distribute goods and services, and serves as support in the choice of policies and operating scenarios. Operating conditions typical of the industry in Mexico are contemplated, specifically of bottled water purification companies. It is intended to evaluate the different operating conditions that optimize delivery points, distribution centers and production plants, considering a holistic vision to meet the objectives of a distribution system, considering their random variations and their impact on the global system to take the preventive and corrective actions to improve the logistics of operation. The above is exemplified with some adjustments that are made on the P-median model where it is shown how this variation would be applied, not considering the source or parent plant as a distribution center option. The improved model offers fundamental advantages over the rest of the techniques reviewed in the literature, such as introducing the information obtained into the model by making the distribution network more optimal, thus collecting not only objective data but also the preferences and forecasts of the decision maker himself.

P-median model, Facilities location, Optimization

Resumen

El presente modelo considera los procesos de producción, así como también los factores económicos y costos logísticos operando en las diferentes etapas de una compañía de agua purificada. Tiene una aplicación general para empresas repartidoras de bienes y servicios, y sirve como apoyo en la elección de políticas y escenarios de operación. Se contemplan condiciones operativas propias de la industria en México, específicamente de las empresas purificadoras de agua embotellada. Se pretende evaluar las diferentes condiciones operativas que optimicen los puntos de entrega, centros de distribución y plantas de producción considerando una visión holística para cumplir con los objetivos de un sistema de distribución, considerando sus variaciones aleatorias y su impacto en el sistema global para tomar las acciones preventivas y correctivas que permitan mejorar la logística de operación. Lo anterior se ejemplifica con unas adecuaciones que se realizan sobre el modelo de la P-mediana donde se muestra cómo se aplicaría esta variación no considerando la fuente o planta matriz como una opción centro de distribución. El modelo ofrece ventajas fundamentales respecto al resto de técnicas revisadas en la literatura, como introducir la información obtenida haciendo más óptima la red de distribución, recogiendo de este modo no sólo los datos objetivos sino también las preferencias y previsiones del propio tomador de decisiones.

Modelo de la p-mediana, Ubicación de instalaciones, Optimización

Citation: SOLÍS-JIMÉNEZ, Miguel Ángel, CALDERÓN-PALOMARES, Luis Antonio, PIÑA-MARTINEZ, Ana Laura and GONZÁLEZ-SOBAL, Martín. A modified p-median model to optimize the location of facilities considering the source plant. ECORFAN Journal-Democratic Republic of Congo. 2020, 6-10: 14.21

† Researcher contributing first author.

Introduction

The logistical problems of location consist of locating one or several services to meet the demand of a series of known points, usually called demand points, seeking to optimize some measure of effectiveness, however, the theory of location of facilities is relatively recent, to Despite the fact that problems related to the location of different types of facilities have accompanied man throughout history (Torres, 2006), since he has had to face deciding where to locate his facilities, taking into account the relationship that they will have with other facilities or with the clients they intend to serve. In general terms, location problems try to find out the location of the facilities of an entity say, company, warehouse, facility or customer, so that costs are minimized or benefits are maximized.

The formulation of these problems may include geographical considerations, social, economic, and personal policies, among which can be mentioned: proximity of raw material sources, availability of resources, existence of competitors or proximity between facilities, among others, however, The design, development and solution of models that adequately represent reality and that serve as a tool to make the best location decision is related to the work carried out by Hakimi (Chaudhuri, 1998) (1964-1965) for the location problem of facilities, factories, warehouses, branches, stations, etc., in which the p -median problem and its extensions have been used to optimally locate these entities.

Two of these most used models in network localization are the p -median problem and the p -center problem. The first consists of minimizing the total sum of the weighted distances by calculating the location of p -service centers in a way that minimizes the total (or average) weighted distance traveled to meet all demand (efficiency), while the second tries to minimize the maximum weighted distance from a service center to its assigned users when trying to find the location of p -service centers in a way that minimizes the maximum distance between a demand point and its closest service center (fairness).

The objective of the p -median makes it efficient but not equitable, while the bound implicit in the problem of the p -center makes it equitable, but not efficient (Toro, 2011).

Most of the aforementioned applications start from deterministic models, in which all the data of the problem are perfectly known. However, the uncertainty in the weights of the vertices or edges is inherent in most of the problems, either because these data are an estimate of future behavior for which historical data are not available, or because the economic environment is not predictable through the statistical techniques in use or because other external factors not controllable by the decision maker influence them (Darós, 1998).

The median p and extents of it often model real-world situations, such as the location of industrial plants, warehouses, and public facilities (Hansen, 1997).

The p -median problem with deterministic data has been extensively studied. The median p problem is one of the basic models in discrete location theory. As with most location problems, it is classified as NP-hard (Kariv, 1979) and therefore heuristic methods are generally used to solve it (Mladenovic, 2005). A compilation of these studies can be consulted in the works of Krarup and Pruzan (Krarup, 1983) where the approaches and methods for solving both the p -median and other related problems appear. However, in many real problems the data, that is, the demands associated with the vertices and / or the distances associated with the edges, are not exactly known.

For example, if a company wants to access new markets, the demands will be estimated from some market research. If the edges represent roads and the associated weights, the time it takes to travel them, it is impossible to know them exactly, since the time will depend on the traffic circumstances, the state of the road, etc. (Darós, 1997). The work on vehicle routing for messaging problems and sending small packages is aimed at improving mainly two indicators: time of arrival, service-oriented; and operating costs, oriented towards efficiency (Arboleda, 2018).

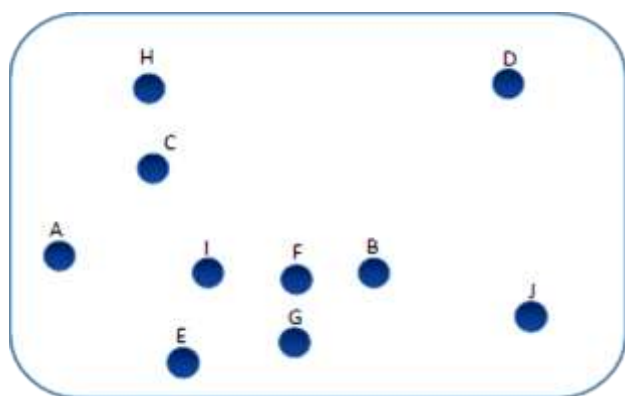
Considering the above, it should be noted that the problem of the p -median is fundamental to a large part of the discrete location theory.

It is motivated by a series of real situations, such as having to install several plants in some points of a transport system to minimize production and shipping costs or having to install different hospitals in a city so that most of its population be covered in case you need one. In short, the p-median problem tries to locate p facilities within a network in such a way as to minimize the weighted distance between the demand points and the closest open facility (Jiménez, 2018).

In markets that demand great competitiveness, business management faces decisions and problems that must be solved strategically for the successful maintenance of companies. Among those decisions, it is of great relevance to satisfy the needs of external customers in terms of product quality and service, with delivery compliance being an element of great impact (Davis-Sramek, Mentzer and Stank, 2008), cited in (Pantoja, 2017).

Derived from the information that precedes a feasibility study for the installation of a water purification company, the following scenario arises where only the distances between the parent plant and its possible distribution centers are considered as main data. The company seeks to determine which of its customers can serve as distribution centers. For this, the logistics and delivery manager determined the distances in kilometers that exist between the different clients and the parent plant.

The distances and the plan of the customer locations are presented below in figure 1 and table 1, respectively:



Location of different clients to the parent plant.

Distances from the main plant to different clients (km).

De/ A	A	B	C	D	E	F	G	H	I	J
A	0.0	63.2	35.5	107.6	43.8	47.5	54.7	59.8	30.4	97.1
B	63.2	0.0	56.2	69.5	49.0	16.1	28.0	76.6	33.0	35.3
C	35.5	56.2	0.0	76.7	66.3	46.4	64.4	27.0	36.7	91.0
D	107.6	69.5	76.7	0.0	115.1	78.8	97.0	72.0	87.7	79.2
E	43.8	49.0	66.3	115.1	0.0	36.4	23.4	93.3	31.4	71.8
F	47.5	16.1	46.4	78.8	36.4	0.0	21.0	70.3	17.1	49.7
G	54.7	28.0	64.4	97.0	23.4	21.0	0.0	89.8	28.6	48.7
H	59.8	76.6	27.0	72.0	93.3	70.3	89.8	0.0	63.2	108.9
I	30.4	33.0	36.7	87.7	31.4	17.1	28.6	63.2	0.0	66.7
J	97.1	35.3	91.0	79.2	71.8	49.7	48.7	108.9	66.7	0.0

The parent plant will be considered by the variable X1 (A). Suppose that the company can condition up to 3 distribution centers, and it may turn out that the parent plant is also selected as the distribution center, this implies that in case the variable X1 is selected as CEDI, then it will have its own clients that attend, in addition to the other centers that are selected.

Methodology

The p-Median model originally formulated in Revelle and Swain (ReVelle, 1970), in its most general form, is defined in a network of nodes and areas where it is assumed that each node represents a local area of demand or can also represent a potential location for a plant installation. The objective is to locate exactly “p-plants”, to minimize the total transport distance to meet all the demand. It is assumed that the plants are not limited in service by maximum capacity (they have infinite capacity) so that each demand can be served by the nearest plant. We can formulate this model with the following notations.

$i, j =$ indices used to refer to a node or point numbered as 1,2,3, 4,... n.

$d_{ij} =$ minimum distance from node i to node j.

$a_i =$ demand at node i.

$X_{ij} = 1,$ if the demand at i is assigned to plant j; 0, otherwise.

$X_{jj} = 1,$ if a plant is located at site j and week j, it is also assigned; 0, otherwise.

$p =$ number of plants to be placed.

Considering the above, the logistic distribution model for the water distribution company can be formulated as follows.

This model aims, in addition to minimizing travel distances, both from the plant to the distribution centers and from the distribution centers to customers.

The model is formulated with the following notations:

i, j = Indices used to refer to a node or point numbered as 1,2,3,4, n.

d_{ij} = Minimum distance from node i to node j .

$X_{ij} = 1$, if demand at i is assigned to distribution center j ; 0, otherwise.

$X_{jj} = 1$, if a distribution center is located at site j and demand is assigned to j , it is also assigned; 0, otherwise.

p = Maximum number of distribution centers that can be assigned.

X_{0101} = Total distance from the distribution centers to the plant

Formulation of the model using the p-median

The next step is to formulate using the P-median method. For this purpose the following considerations will be used:

1. The objective function must minimize the distances from each of the possible distribution centers to the rest of them, including the parent plant.
2. The first set of restrictions ensures that each customer is assigned to a distribution center, provided it is activated as such.
3. The second set of restrictions indicates how many customers can be served by a CEDI, for the example it is considered that each distribution center will serve three customers, including itself.
4. The last restriction determines the maximum number of CEDIS that can be enabled.
5. All the decision variables must be binary, that is, they indicate if the location is selected as CEDI and if the location is served by CEDI.

Taking the previous information, the model is as follows:

$$Min Z = X_{0101} + \sum_{j=1}^{10} \sum_{i=1}^{10} d_{ij}X_{ij} \tag{1}$$

$$\sum_{j=1}^{10} X_{ij} = 1 \text{ for each } i = 1,2,3, \dots n \tag{2}$$

$$\sum_{i=1}^{10} X_{ij} = 1 \text{ for each } j = 1,2,3, \dots n \tag{3}$$

$$X_{0101} = \sum_{j=1}^{10} d_{ij}X_{jj} \text{ for each } j = 2,3,4, \dots n \tag{4}$$

$$X_{ij} \leq X_{jj} \text{ for each } i = 1, 2, 3, \dots n; j = 1, 2, 3, \dots n, \text{ where } i \neq j \tag{5}$$

$$X_{ij} \in \{0,1\} \text{ for each } i = 1, 2, \dots n; j = 1, 2, \dots n \tag{6}$$

Results

Solution and interpretation of the model

Once the situation has been modeled, the next point is to solve the model, for this purpose, the Excel® Solver is used. It can be mentioned that the locations corresponding to variable (B), as well as locations (C) and (G) are those selected as distribution centers. The distribution center (B) must serve the markets of (D), (F) and (J), on the other hand, the distribution center (C) must serve the markets of (A) and (H); finally, the distribution center (G) must serve the markets of (I) and (E). The above can be seen in the following Table 2:

Selected Distribution Center	customers
B	D F J
C	A H
G	E I

Table 1 Selected distribution center with their respective customers

The distances from each Distribution Center to their respective customers are:

Center of distribution	Customer	Distance to the distribution center (Km)
B	D	69.5
	F	16.1
	J	35.3
Total		120.9

Table 2 Distance from CEDI (B) to its clients

Center of distribution	Customer	Distance to the distribution center (Km)
C	A	35.3
	H	27.4
Total		62.7

Table 3 distance from CEDI (C) to its clients

$$\sum_2^{10} X_{ij} \leq 3 \text{ for each } i = 2, \dots, 10, j = 2, \dots, 10 \text{ where } i = j \tag{8}$$

$$X_{0101} = \sum_{j=2}^{10} \sum_{i=2}^{10} d_{ij} X_{ij} \text{ for each } i = 2, 3, \dots, 10; j = 1, 2, \dots, 10 \text{ where } i \neq j \tag{9}$$

$$X_{ij} \leq X_{jj} \text{ for each } i = 2, 3, \dots, 10; j = 2, 3, \dots, 10 \text{ where } i \neq j \tag{10}$$

$$\sum_2^{10} X_{ij} = 1 \text{ for each } j = 2, 3, \dots, 10 \tag{11}$$

$$X_{ij} \in \{0,1\} \text{ for each } i = 1, 2, \dots, n; j = 1, 2, \dots, n \tag{12}$$

The result of the improved algorithm indicates that the nodes that will act as distribution centers are (C), (I) and (F) and the clients that will serve each of these can be seen in the following figure:

Selected Distribution Center	customers
C	D H
F	B G J
I	E

Table 6 Selected distribution centers with their respective customers

Graphically the distribution can be seen in Fig. 3.

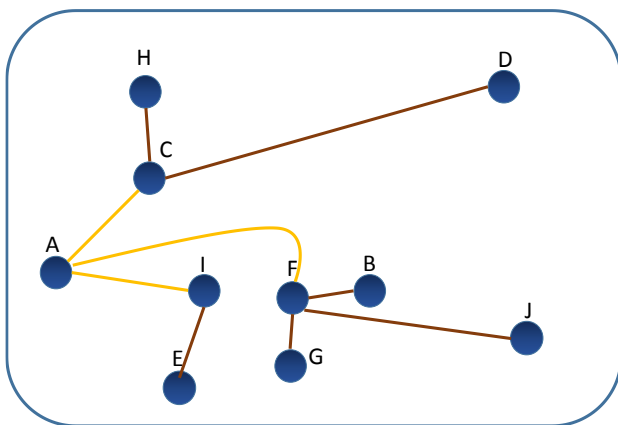


Figure 3 Result of the improved algorithm

As can be seen, node (C) will serve nodes (H) and (D), node (F) will serve nodes (B), (G) and (J) and finally node (I) will serve to node (E). The distances of each CEDI to its clients are observed in Tables 8, 9 and 10.

Center of distribution	Customer	Distance to the distribution center (Km)
C	D	76.7
	H	27
Total		103.7

Table 7 distance from CEDI (C) to its clients

Center of distribution	Customer	Distance to the distribution center (Km)
F	B	16.1
	G	21
	J	49.7
Total		86.8

Table 8 Distance from CEDI (F) to its clients

Center of distribution	Customer	Distance to the distribution center (Km)
I	E	31.4
Total		31.4

Table 9 distance from the CEDI (I) to its client

Note that the total distance traveled from the CEDIS to their clients adds up to 221.9 km, without considering the distance from the parent plant to them, which is reflected in table 10.

Center of distribution	Customer	Distance to the distribution center (Km)
A	C	35.5
	F	47.5
	I	30.4
Total		113.4

Table 10 Distance from the matrix to the CEDIS

Adding the 221.9 km from the CEDIS to the clients plus the 113.4 from the parent plant to the CEDIS, there is a total of 335.3 km. This distance is less than that presented in the two previous cases. In this particular case, an improvement can be seen in terms of the distance traveled in relation to the first model of 13.75% and in relation to the second model of 5.1%.

Conclusions

The p-median problem is central to much of discrete location theory. It is motivated by a number of real-world situations, such as having to install multiple plants at some points in a transportation system to minimize production and shipping costs or having to install different hospitals in a city so that most of the your population is covered in case you need one.

Ultimately, the p-median problem tries to locate p facilities within a network in a way that minimizes the weighted distance between demand points and the closest open facility.

In the formulation of the model, the total distance traveled to cover these markets is 235.47 km, not counting the distance to take the products from the source to the distribution centers, so that including these distances that are 153.42 km, we have a total of 388.88 km. In the model without using the plant as CEDIS, the total route of these distribution centers is 199.96 km.

To this quantity, the distance from the matrix (A) to the nodes selected as distribution centers must be added, which is 153.42 km, leaving a total of 353.37 km. And finally, in the improved model, the total distance in this case will be 335.41 km. This distance is less than the one presented in the two previous cases. In this particular case, an improvement can be seen in terms of the distance traveled in relation to the first model of 13.75% and in relation to the second model of 5.1%.

The improved model offers two fundamental advantages over the rest of the techniques discussed above. The first of these is that it allows us to focus on the information obtained by making the distribution network more optimal, thus establishing an adequate policy for the conformation of the distribution network. The second is that it allows the decision-maker to intervene in the resolution of the model, thus collecting not only the objective data but also the preferences and forecasts of the decision-maker himself.

References

- Arboleda-Castillo J. J., A. D. Heredia-Giraldo, and J. P. Orejuela-Cabrera, "Two-phase method for motorcycle messenger routing problem with time windows. Método de dos fases para el problema de ruteo de mensajeros en motocicleta con ventanas de tiempo," *Entramado*, vol. 14, no. 1, pp. 268-281, 2018-06 2018, doi: 10.18041/entramado.2018v14n1.27120.
- Chaudhuri, Shiva; Garg Naveen; Ravi R.,"The p-neighbor k-center problem," *Information Processing Letters*, vol. 65, no. 3, pp. 131 - 134, 1998, doi: [https://doi.org/10.1016/S0020-0190\(97\)00224-X](https://doi.org/10.1016/S0020-0190(97)00224-X).
- Darós Canos, M. a. A. M. a. N. V. a. S. V. a. S. R. J., "Aplicaciones económicas de la optimización robusta," *Rect@*, vol. Actas_5, 1997.
- Darós M. C. and M. M. Arce, "Técnicas de optimización robusta Aplicadas al problema de la p-mediana en condiciones de incertidumbre," *Actas de VI Jornadas de la Asociación Española de Profesores Universitarios de Matemáticas para la Economía y la Empresa*, Santiago de Compostela, 1998.
- Hansen P., Mladenovic,"Variable neighborhood search for the p-median," *Location Science*, vol. 5, no. 4, pp. 207 - 226, 1997, doi: [https://doi.org/10.1016/S0966-8349\(98\)00030-8](https://doi.org/10.1016/S0966-8349(98)00030-8).
- Jiménez S. L., "Problemas de la p-mediana Aplicación," 2018.
- Kariv O. and S. L. Hakimi, "An Algorithmic Approach to Network Location Problems. I: The p-Centers," *SIAM Journal on Applied Mathematics*, vol. 37, no. 3, pp. 513-538, 1979.
- Krarup J., Pruzan, P.M.,"The simple plant location problem: Survey and synthesis," *European Journal of Operational Research*, vol. 12, no. 1, pp. 36 - 81, 1983, doi: [https://doi.org/10.1016/0377-2217\(83\)90181-9](https://doi.org/10.1016/0377-2217(83)90181-9).
- Mladenovic N., J. Brimberg, P. Hansen, and J. A. Moreno-Perez, "The p-median problem: A survey of metaheuristic approaches," *European Journal of Operational Research*, vol. 179, no. 3, pp. 927-939, Jun 16 2007, doi: 10.1016/j.ejor.2005.05.034.
- Pantoja C., J. P. Orejuela, and J. J. Bravo, "Metodología de distribución de plantas en ambientes de agrupación celular," *Estudios Gerenciales*, vol. 33, no. 143, pp. 132-140, 2017.
- ReVelle C. S. and R. W. Swain, "Central Facilities Location," *Geographical Analysis*, vol. 2, no. 1, pp. 30-42, 1970, doi: 10.1111/j.1538-4632.1970.tb00142.x.
- Torres Díaz L, "Análisis de diferentes políticas de ubicación de una instalación en ambientes competitivos," Maestría en Ciencias, Programa de Posgrado de Ingeniería en Sistemas, Universidad Autónoma de Nuevo León, México, 2006.

Toro Ocampo E., R. Gallego Rendón, and C. A. Muñoz, "Comparación del desempeño del algoritmo genético de CHU-BEASLEY y el algoritmo colonia de hormigas en el problema de p-mediana," *Scientia et technica*, vol. 1, no. 47, p. 6, 2011, doi: 10.22517/23447214.531.