

A Solution Proposal for the Capacitated P-Median Problem with Tabu Search

Mauricio Romero Montoya¹, María Beatríz Bernábe Loranca²,
Rogelio González Velázquez², José Luis Martínez Flores¹, Horacio Bautista Santos³,
Abraham Sánchez Flores², Francisco Macias Santiesteban²

¹ Universidad Popular Autónoma del Estado de Puebla,
Mexico

² Benemérita Universidad Autónoma de Puebla (BUAP),
Facultad de Ciencias de la Computación,
Puebla, Mexico

³ Instituto Tecnológico Superior de Tantoyuca,
Mexico

Abstract. This work presents the Capacitated P-Median problem, which seeks to solve the optimal location of p distributors. In understanding the computational complexity of this problem, partitioning principles are applied in the process of associating a distribution center with customers to be serviced. These are treated as clusters containing each distribution center and its customers. The optimization of the implicit cost function within the partitioning (distance minimization) is performed using Tabu Search with a diversified search. A series of computational experiences is performed aided by the OR-Library test instances, where in most cases an optimal solution is reached in reasonable computing time. As Tabu Search does not obtain all expected optima, OR-Library test instances are solved with the help of Lingo 16.0 and the obtained results are compared with those resulting from Tabu Search.

Keywords: Tabu Search, capacited P-median, NP-hard.

1 Introduction

The location of facilities deals mainly with choosing the physical location for a set of facilities in such a way that, given a set of restrictions, the demands from a set of customers are optimally met for some target function, from some set of candidate places. Facilities are to be understood in a broader sense, as they could be entities such as hospitals, factories, schools, ports, etc.

In this area, the P-Median and Capacitated P-Median (CPMP) problems are noteworthy due to their popularity in several areas. Here, the Capacitated P-Median problem is dealt with, as it was treated in previous work without reaching all optima.

Here, Tabu Search (TS) is introduced as an approximation tool, thus reaching the optima established in OR-Library.

This work is organized as follows: this introduction is section 1. Section 2 shows general aspects of related works. Section 3 shows the modeling leading to Section 4, which presents the approximation method and associated algorithm, section 5 shows the results of the computational experiences. Lastly, in section 7 results and conclusions are presented.

2 Related Works: P-Median and Capacitated Facility Location Problems

Location models are applied for specific cases, thus their structure (objectives, restrictions and variables) are case-dependent, therefore there are no generic models appropriate for all problems, and models with different objectives may provide different solutions for the same case study.

Facility location, also known as location analysis or location science, is concerned with the siting of facilities on a plane or within a network. These locations are dependent upon various attributes such as customers' demand on a facility, the cost of supplying those customers and the costs of opening facilities at potential locations.

A general description of location problems may be found in [1]. The problem dates back to Euclid and Pythagoras, due to their proposals of geometric models for distances; however literature suggests that it was emperor Constantine who pioneered its formal expression, as the location in a network was solved by means of discrete coordinates that placed the roman legions. Later, Fermat and Torricelli formulated the three node problem, which led to the Weber problem we know today (the infinite solution space Euclidian distance minimum single site location problem). In the 19th century Sylvester posed the problem of minimizing the maximum Euclidean distance in the infinite solution space.

In the 20th century Steiner's problem is generalized and formalized. Other contributions appear in the field of logistics (Hoover and Pallander [2]), but it was Wersan, Quon and Charnes [3] who placed the problem within the framework of mathematical optimization and programming, leading to the work of Daskin [4]. Location decisions may well be the most difficult for the efficient design of a supply chain, due to their strategic nature, for it is a long-term decision, which means it should consider facility cost, which may be quite steep as happens in manufacturing facilities. Therefore, it must be concluded that it is an inflexible decision. For a deeper discussion on the relevance of location decisions the reader is advised to refer to Daskin [5].

The quality of the location decision may make the difference in the efficiency of a supply chain, affecting customer satisfaction. Let us not forget that location decisions investigates the siting of a set of facilities in order to best satisfy the needs of a set of customers (Hale and Moberg, 2003) [6] thus, a poor location decision would hinder this purpose, contradicting the definition of supply chain as proposed by Chopra and Meindl (2001) [7]. Brandeau and Chiu (1989) [8] reviewed 50 location problems that had been studied, indicating their formulations, classification and relationships, thus showing a

series of optimization objectives that models may follow. Discrete models for the location in networks may be classified according to distance (Current et al. 2001) [9], distinguishing models based on maximum distance, and models based on total or average distance.

Daskin (2008) and ReVelle et al. (2008) [10, 11] reviewed and categorized the models for location problems into four classes: analytical, continuous, network and discrete. Analytical models refer to simple contrived problems. They have analytical value and are solved using classical mathematical optimization techniques but are of limited use in practice, due to their simplicity. While customer demand is most often restricted to fixed points or nodes within a region or a plane, continuous problems allow for the location of facilities anywhere within the region. These problems are known as Weber-type problems, and are often solved by means of analytical techniques. The applications and algorithmic techniques for these problems were reviewed by Drezner et al. (2001) [12, 13].

2.1 P-Median

The P-median problem, credited to Hakimi (1964) is an adaption of the classical Weber problem. It is particularly difficult to solve when considering problems of a more practical size where the number of variables may run in the hundreds or even thousands. Hence heuristic solution methods are often employed, as the complete enumeration time component increases exponentially with the number of binary variables to be considered (facilities and customers) [13]. The combinatorial nature of the p-median problem caused advances in solution techniques to be often made using combinations of mathematical programming and heuristic search techniques such as branch and bound and Lagrangean relaxation. This significant work, along with its results, is still often cited today by many authors. Reese (2006) [14], provides an annotated bibliography for the p-median problem which indicates that since the late 1980s there has been a significant shift from using traditional mathematical programming relaxation techniques, such as Lagrangean and surrogate relaxation, towards the use of more modern metaheuristics techniques (Bernabé 2012) [15].

2.2 Capacitated Facility Location Problem

The p-median and uncapacitated location problems both consider facilities with more than enough capacity or supply to provide for all their assigned customers. However, as facilities have only a finite supply or capacity, in practice this may not be the case. While most customers may be assigned to their nearest facility, when supply runs out, some customers may need to be directed –or at least partially directed – to the nearest available facility. Thus, there are two cases: first, the assignment of customers to facilities is determined via a binary constraint, to ensure a customer is served from a single facility, and second, the demand constrained may be fractional, to allow for the demand to be satisfied from one or more facilities. As a direct consequence, the capacitated p-median and the capacitated facility problems have been the interest of studies. The capacity constraints result in greater theoretical difficulty for the solution

of these problems, and are generally NP-hard. The contributions to solving the capacitated p-median problem are primarily metaheuristic or hybrid based.

The capacitated facility location problem may be regarded as either deriving from the the p-median problem by including fixed opening costs for facilities or, even better, as extending the uncapacitated facility location problem by considering each facility to have a limited supply or capacity constraint. It is referred to be NP-hard, and Kariv and Hakimi (1979) [16, 17] and Garey and Johnson (1979) [18] are often cited, even though neither considered this problem. However, the problem is modeled on a network, and is therefore likely to be at least as hard as the p-median problem. Heuristic algorithms are often presented as solution techniques. ReVelle and Eislet (2005) [19] pointed out that the introduction of capacity constraints effectively eliminates the property that all demand from a given customer is to be assigned to a single facility, and thus makes the problem much more difficult to solve.

3 Metaheuristics for Capacitated P-Median CPMP

Many attempts have been made to solve both capacitated and uncapacitated P-median. Two articles are of interest regarding Tabu Search, as this approximation to the capacitated p-median problem is solved by means of TS.

A Tabu Search procedure for the CPMP was developed by Sörensen (2008) [20], which managed to produce optimal solutions for all the small to medium sized OR-Library test problems, save two. Even though their paper does not identify said instances, relative errors were reported to be less than 0.1%, which represents an improvement over any previous work for CPMP. This work both backs up the conclusions by Arostegui et al. (2006) [21] and suggests that Tabu Search may well be a dominant methodology to use when solving the CPMP.

3.1 Capacitated P-Median Model

The CPMP, also known as the Capacitated Clustering Problem, the Capacitated Warehouse Location Problem, the Sum-of-Stars Clustering Problem and some others, is a classical location problem that finds applications in many practical situations. It may be described as: given a set of n points (customers), each of them with a known demand, it is required to find p medians (centers) and assign each point to exactly one median such that the total distance of assigned points to their corresponding medians is minimized, and the capacity limit on the medians may not be exceeded. Several heuristics and metaheuristics have been proposed for these problems, known to be NP-hard [18]. Osman and Christofides [22] propose a simulated annealing and tabu search method. Maniezzo [7] present a bionomic algorithm for the solution of this problem. Lorena and Senne [6] explore local search heuristics based on location-allocation procedures and Lorena and Senne [25] use column generation to CPMP.

The CPMP considered in this paper is modeled as the following binary integer programming problem:

$$z = \min \sum_{i \in N} \sum_{j \in N} d_{ij} x_{ij} \quad (1)$$

$$\text{subject to } \sum_{j \in N} x_{ij} = 1, \forall i \in N, \quad (2)$$

$$\sum_{j \in N} x_{jj} = p, \quad (3)$$

$$\sum_{i \in N} q_i x_{ij} \leq Q x_{jj}, \forall j \in N, \quad (4)$$

$$x_{ij} \in \{0,1\}, \forall i \in N, \forall j \in N, \quad (5)$$

where: $N = \{1, \dots, n\}$ is the index set of points to allocate and also of possible medians, where p medians will be located; q_i is the demand of each point and Q the capacity of each possible median; d_{ij} is a distance matrix; x_{ij} is the allocation matrix, with $x_{ij} = 1$ if point i is allocated to median j , and $x_{ij} = 0$, otherwise; $x_{jj} = 1$ if median j is selected and $x_{jj} = 0$, otherwise.

Target function (1) minimizes the sum of distances between all demand nodes. Constraint (2) ensures each node is assigned to exactly one facility. Constraint (3) sets the number of medians to be located. Constraint (4) imposes that a total median capacity must be respected, and (5) provides the integer conditions.

Clustering Search Partitioning (CSP) was used to assign clients to suppliers. This uses clustering to detect promising areas in the search space. It is of particular interest to find out these areas as soon as possible, in order to change the search strategy over them. In CSP, a clustering process is executed simultaneously with the Tabu Search algorithm, identifying groups of individuals that call for special interest.

The CS tries to locate promising search areas by framing them by clusters. Clusters may be defined as $G = \{c; r; s\}$ where c is the center, r the radius of the area, and s is a search strategy associated with the cluster.

The center c is a solution that represents the cluster, identifying the location of the cluster inside of the search space. The radius r establishes the maximum or minimum distance, starting from the center, that a solution can be associated to the cluster. The search strategy s is a systematic search intensification, in which solutions of a cluster interact among themselves along the clustering process, generating new solutions.

4 Tabu Search for Capacitated P-Median

This approximation relies on traditional Tabu Search, which generates a random initial solution. The parameters here have been defined as: 1) numberOfIterations (global iterations to be performed by the algorithm) 2) perturbationIterations, (determines the number of times a solution may be accepted according to the solution of the previous iteration. When the counter perturbationCounter crosses this limit, a new random solution is generated, and search is restarted).

The module parameter determines the frequency with which in Vicinity 1 (Algorithm 2) is to be used instead of Vicinity 2 (Algorithm 3), as can be seen in the first conditional within the main while loop. When the cost given by the vicinity functions improves on the best known cost (bestCost), this will be replaced, along with the solution configuration (bestSolution). The last step updates the tabu lists and global iteration counter.

Algorithm 1. Tabu Search (PSEUDOCODE):

```
READ numberOfIterations
READ perturbationIterations
READ module
SET iterationCounter and perturbationCounter to 0
SET initialSolution and bestSolution to currentSolution

WHILE iterationCounter < numberOfIterations
    SET currentCost to cost of currentSolution
    IF iterationCounter % module = 0
        CALL Neighbor1 with currentSolution RETURNING newCost
    ELSE
        CALL Neighbor2 with currentSolution RETURNING newCost
    END IF
    IF newCost > currentCost
        INCREMENT perturbationCounter
    END IF
    IF newCost < bestCost
        SET bestSolution to newSolution
        SET bestCost to newCost
    END IF
    SET currentSolution to newSolution
    SET currentCost to newCost
    IF perturbationCounter > perturbationIterations
        CALL RestartSolution RETURNING newSolution
        SET currentSolution to newSolution
        SET perturbationCounter to 0
    END IF
    CALL UpdateTabuLists RETURNING updatedTabuLists
    INCREMENT iterationCounter
END WHILE
RETURN bestSolution
```

CPMP is a problem in combinatorial optimization, classified as NP-Hard, and the tabu search metaheuristic has proven efficient in searching for solutions for a wide variety of problems of this kind, such as the quadratic assignment problem, the traveling salesman problem, or p-median search, among others. The tabu search metaheuristic uses different intelligent search schemes, like the short and long term memory use, keeps a list of tabu solutions that should not be visited because they lead to trajectories

that have proven not to be good solutions; these are tools in the search for artificial intelligence.

5 Computational Experience

Table 1 gathers the attained results for two kinds of tests with the Tabu Search approach: 1) varying parameters (random) and 2) using default set parameters and, on the other hand, the results obtained by the Branch & Bound method in Lingo 16 are presented.

It can be seen that varying parameters gives better results, and the OR-Library [26] demanded optima are approached. On the other hand, the results obtained are better than those reached in previous work [27].

Table 1. Results for Capacited P-Median with TS.

I	Capacited Q=120 Random						Default			Lingo 16	
	BV	C	P	RC	T	GAP	DC	T	GAP	OPT	T OPT
1	713	50	5	717	48340	0.56	722	10040	1.26	713	7.77
2	740	50	5	740	50933	0	740	9950	0	740	1.05
3	751	50	5	752	72445	0.13	772	9810	2.8	751	2.34
4	651	50	5	653	53402	0.31	654	10030	0.46	651	1.18
5	664	50	5	664	50073	0	664	9950	0	664	2.80
6	778	50	5	778	24426	0	778	9930	0	778	1.55
7	787	50	5	796	52371	1.14	810	9820	2.92	787	63.02
8	820	50	5	826	54607	0.73	851	9690	3.78	820	236.96
9	715	50	5	718	23938	0.42	721	9840	0.84	715	9.95
10	829	50	5	844	47455	1.81	844	9810	1.81	829	129.25
11	1006	100	10	1054	3E+05	4.77	1074	11900	6.76	1006	119.65
12	966	100	10	978	2E+05	1.24	983	11900	1.76	966	219.48
13	1026	100	10	1032	1E+05	0.58	1055	12050	2.83	1026	27.22
14	982	100	10	1012	3E+05	3.05	1027	12110	4.58	982	1190.93
15	1091	100	10	1133	7E+05	3.85	1160	12010	6.32	1091	711.64
16	954	100	10	975	2E+05	2.2	985	12120	3.25	954	108.53
17	1034	100	10	1077	64331	4.16	1072	11960	3.68	1034	112.21
18	1043	100	10	1086	67329	4.12	1088	12040	4.31	1043	543.21
19	1031	100	10	1057	63633	2.52	1049	11780	1.75	1031	380.67
20	1005	100	10	1133	66159	12.7	1146	11850	14.03	1005	48439.28

In the table 1, the notation is the next: I=Instance, BV is the best known value published in the OR-Library, C are the n Customers, P are the medians, RC are the values in the experiment when we were varying parameters (random), T (sec), is the execution time for each test, The GAP is the difference between the best cost and the obtained value by the approximation method, DC are the default parameters by default with TS, OPT are the values obtained with optimization software Lingo 16, T OPT is

the run time for each instance in Lingo 16. Tests were performed on a PC with 6GB RAM and a second generation i5 processor.

6 Conclusions

Throughout this work the extensive literature, efforts and applications surrounding the Capacitated P-Median problem have been confirmed.

Our objective is to obtain all optima dictated by the OR-Library, and at this point in the computational experience, it is assumed that satisfaction of optima is possible by adjusting the parameters in Tabu Search in such a way that the design of factorial statistical experiments becomes necessary.

In this work the robustness of the tabu search performed on the 20 OR-Library instances has been proven and compared with the results from Lingo, with an acceptable GAP, as shown in table 1.

The contribution from this work is the design of a sequential program based in object oriented programming and a data structure that allows for covering trajectories that partially explore the set of feasible solutions, known to have a combinatory explosion as the size of the problem increases even in moderate cases. This Tabu Search proposal is competitive even with the implementation of Lingo 16, a professional version without restrictions that is recommended for large scale testing.

Lastly, the capacitated P-median is part of the localization-assignment problems which define the strategy of an organization, that is, they are long-term decisions, and a good decision in the distribution and location of facilities has positive implications in the efficiency of the supply chain and the cost competitiveness of products and services reaching the customer, thus in general being benefic for the population sector that demands said product or service. In this point, as a future development from this work, a series of applications with both TS and Lingo 16 is to be developed, in particular for the meat industry.

References

1. Revelle, C.: A perspective of location Science. *Location Science*, 5, pp. 3–13 (1997)
2. Hoover, E.: *The location of Economic Activity*. New York (1948)
3. Wersan S. J., Quon J. E., Charnes, A.: System Analysis of refuses collection and disposable practice. *Yearbook American Public Works Association*, 14, pp. 195–211 (1962)
4. Daskin, M.S.: *Network and Discrete Location: Models, Algorithms, and Applications*. New York, John Willey & Sons, Inc., pp. 201–203 (1995)
5. Shen, Z. J., Coullard, C., Daskin, M. S.: A joint location-inventory model *Transportation science*. 37(1), pp. 40–55 (2003)
6. Hale, T., Moberg, C.: Location science research: A review. *Annals of Operations Research*, 123(1), pp. 21–35 (2003)
7. Chopra, S., Meindl, P.: *Supply Chain Management: Strategy, Planning, and Operation*. 2 Edition, New Jersey: Pearson Prentice Hall (2001)
8. Brandeau, M., Chiu, S.: An Overview of Representative Problems in Location Research. *Management Science*, 35(6), pp. 645–674 (1989)

9. Kaufman, L., Current, J., Daskin, M., Schilling, D.: Discrete Network Location Models. In: Facility Location Theory: Applications and Methods, Berlin: Z. Drezner and H. Hamacher Eds., pp. 85–112 (2001)
10. Daskin, M.: What you should know about location modelling. *Naval Res Logis*, 55, pp. 283–294 (2008)
11. Reville, C., Eislet, H., Daskin, M.: A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*, 184, pp. 817–848 (2008)
12. Drezner, Z., Klamroth, K., Schobel, A., Wesolowsky, G.: *The Weber Problem*. Berlin, Germany: Springer-Verlag, pp. 1–36 (2001)
13. Venables, H.: Ant Colony Optimization: A Proposed Solution Framework for the Capacitated Facility Location Problem. <http://sure.sunderland.ac.uk/policies.html> (2011)
14. Reese, J.: Methods for solving the p-median problem: An annotated bibliography. *Networks*, 48(3), pp. 125–142 (2006)
15. Diaz, J. A., Bernabe, M. B., Luna, D., Olivares, E., Martínez, J. L.: Lagrangean relaxation for the geographical partitioning problem. *Rev. Mate. Teor. Aplic.*, 19(2), pp. 169–181 (2012)
16. Hakimi, S.: An algorithmic approach to network location problems. part i: The p-centres, *SIAM Journal of Applied Mathematics*, 37, pp. 513–538 (1979)
17. Kariv, O., Hakimi, S.: An algorithmic approach to network location problems. part ii: The p-median, *SIAM Journal of Applied Mathematics*, 37, pp. 539–560 (1979)
18. Garey, M. R., Johnson, D. S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman (1979)
19. Reville, C., Eislet, H.: Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165, pp. 1–19 (2005)
20. Sörensen, K.: Investigation of practical, robust and flexible decisions for facility location problems using tabu search and simulation. *J Operations Research Society*, 59(5), pp. 624–636 (2008)
21. Arostegui, M. J., Kadipasaogul, S., Khumawala, B.: An empirical comparison of tabu search, simulated annealing and genetic algorithms for facilities location problems. *International Journal of Production Economics*, 103, pp. 742–754 (2006)
22. Osman, I. H., Christofides, N.: Capacitated clustering problems by hybrid simulated annealing and tabu search. *Intern. Trans. in Operational Research*, Vol.1 (3), pp. 317–336 (1994)
23. Maniezzo, V., Mingozzi, A., Baldacci, R.: A bionomic approach to the capacitated p-median problem. *J Heuristics*, (4), pp. 263–280 (1998)
24. Lorena, L. A., Senne, E. L.: Local search heuristics for capacitated p-median problems. *Networks and Spatial Economics*, 3(4), pp. 407–419 (2003)
25. Lorena, L. A., Senne, E. L.: A column generation approach to capacitated p-median problems. *Computers and Operational Research*, Vol.31, pp. 863–876 (1994)
26. Bernábe, B., González, R., Estrada, M.: Adaptación de Particionamiento sobre Medoides a PMediana Capacitado. En: 5° Congreso Internacional de computación, México – Colombia, XV Jornada Académica en Inteligencia Artificial Cicom, pp. 2462-9588 (2015)
27. OR-library: <http://people.brunel.ac.uk/~mastjib/jeb/orlib/pmedinfo.html>, Retrieved (2014)