10-2006

Solution Methods for the *p*-Median Problem: An Annotated Bibliography

Josh Reese
Trinity University, jreese@trinity.edu

Repository Citation
Reese, J. (2006). Solution methods for the *p*-median problem: An annotated bibliography. *Networks, 48*(3), 125-142. doi:10.1002/net.20128

Methods for Solving the p-Median Problem: An Annotated Bibliography

J. Reese*

August 11, 2005

Abstract

The p-median problem is a graph theory problem that was originally designed for, and has been extensively applied to, facility location. In this bibliography, we summarize the literature on solution methods for the uncapacitated and capacitated p-median problem on a graph or network.

1 Introduction

There are four primary problems in the field of *discrete location theory*: the *p-median problem*, the *p-center problem*, the *uncapacitated facility location problem (UFLP)* and the *quadratic assignment problem (QAP)* [82]. These problems decide the location of facilities and allocate demand points to one or multiple facilities. For this reason, they are often called *location-allocation* problems. The p-median problem is well studied in the literature. A few surveys on some of the solution methods have been published, the last being a chapter in [31] that appeared in 1995. Over the past ten years, however, there has been a dramatic increase in the amount of literature on solution methods (see Table 2), and to our knowledge a survey in the form of an annotated bibliography has never been produced. The goal of this paper is to present an up-to-date, exhaustive annotated bibliography.

The p-median problem is one of a larger class of problems known as *minisum location-allocation problems.* These problems find *medians* among existing points, which is not the same as finding *centers* among points, a characteristic of *minimax location-allocation problems* (the p-center problem is an example, where the goal is to minimize the maximum distance between points and center(s)). Minisum problems originated in the 17th century when Fermat posed the following question: Given a triangle (three points in a plane), find a median point in the plane such that the sum of the distances from each of the points to the median point is minimized. In the early 20th century, Alfred Weber presented the same problem with the addition of weights on each of the three points to simulate customer demand. Finding the median point corresponded to finding the best location for a facility to satisfy the demands at the points. This problem is usually acknowledged as the first location-allocation problem. It was later generalized to find the median of $n \geq 3$ points in a plane, and to the multifacility Weber problem, which generalizes to the case of $p > 1$ medians among a number of points in the plane.

The Weber problem locates medians (facilities) at continuous locations in the Euclidean plane. In the early 1960s, Hakimi developed similar problems for finding medians on a network or graph [53, 54], and his absolute median problem is similar to Weber's weighted problem. Hakimi defined the absolute median as the point on a graph that minimizes the sum of the weighted distances between that point and the vertices of the graph. Hakimi allowed this point to lie anywhere along the graph's edges, but proved that an optimal absolute median is always located at a vertex of the graph, thus providing a discrete representation of a continuous problem. In [54] Hakimi generalized

*Department of Mathematics, Trinity University, jreese@trinity.edu. Supported by the Cancer Therapy Research Center, San Antonio, TX. Special thanks to Dr. Allen Holder, Department of Mathematics, Trinity University, for his suggestion to create this bibliography and his consistent support and guidance.

the absolute median to find p medians on a graph in order to minimize the sum of the weighted distances. Again, these points were allowed to be located anywhere along the edges of the graph. Although not all optimal solutions to this problem are located at the vertices, Hakimi showed that there is always a collection of p *vertices* that minimizes the objective. Thus, Hakimi again provided a discrete representation of a continuous problem by restricting the search to the vertices. Solutions consisting of p vertices are called *p-medians* of the graph. Thus, the p-median problem differs from the Weber problem because it is discrete, a consequence of only being allowed to select medians from the candidate set V. It is also defined on a graph or a network, not a plane.

Hakimi developed the absolute median and p-median to find the optimal location of switching center(s) in a communication network. Since his work, the p-median problem has been inseparable from location theory, becoming one of the most common facility location models. Another common model is the UFLP, often referred to as the simple plant location problem or the warehouse location problem. The UFLP is similar to the p-median problem, and the methods used to solve one are often adapted to solve the other. Both problems have similar goals. As in the p-median problem, the UFLP involves locating facilities to minimize demand-weighted distance. The problems differ in the following ways. First, UFLP involves a fixed cost for locating a facility at a given vertex or node, and the p-median problem does not. Second, unlike the p-median problem, UFLP does not have a constraint on the maximum number of facilities. Lastly, typical UFLP formulations separate the set of possible facilities from the set of demand points. In the p-median problem these sets are identical. The QAP is also used to model location-allocation, but it is theoretically harder to solve than the p-median problem. Besides the fact that the objectives differ, the QAP uses flow and cost information not used in the p-median problem. The p-median problem is driven by distance alone.

2 The p-Median Problem

The p-median problem is simply stated as: Given a graph or a network $G = (V, E)$, find $V_p \subseteq V$ such that $|V_p| = p$, where p may either be variable or fixed (see Section 2.3), and that the sum of the shortest distances from the vertices in $\{V \backslash V_p\}$ to their nearest vertex in V_p is minimized. In this section we provide an extended problem definition and a unified notation scheme.

2.1 Basic Problem Definition

Let $G = (V, E)$ be a complete, weighted and undirected graph, where V is the set of vertices and E is the set of edges. Associate with each edge a weight $d(v_i, v_j)$, which is the shortest distance between vertices v_i and v_j according to the metric d. The $n \times n$ symmetric matrix $d_{ij} = [d(v_i, v_j)]$ is the *shortest distance matrix*. Each vertex v_i is assigned a weight w_i, and the *weighted distance matrix* is

$$W_{ij} = w_i d_{ij}.$$

This matrix is not generally symmetric, unless $w_i = w_j$ for each i and j. The *metric p-median problem* is a variation that restricts the weighted distance to be a metric. The p-median problem is often defined on a network and in this case a graph may be created by connecting nodes with an edge whose weight is the shortest distance between the nodes on the network.

Hakimi [53] defined a point m as an *absolute median* of G if, for every point v_j on G,

$$\sum_{i=1}^{n} w_i d(v_i, m) \leq \sum_{i=1}^{n} w_i d(v_i, v_j).$$

He later generalized this concept to the well-known p-median definition of today. Let V_p be a set of p points on G and let $d(v_i, V_p)$ and $d(v_i, V_p^*)$ be the shortest distances from vertex v_i to its nearest element in V_p and V_p^*, respectively. The definition from [54] is: A set of points V_p^* is a *p-median* of G if, for every V_p on G,

$$\sum_{i=1}^{n} w_i d(v_i, V_p^*) \leq \sum_{i=1}^{n} w_i d(v_i, V_p).$$

Under this generalization, the absolute median is the 1-median of the graph. The term *p-median* refers to the set of vertices V_p. The vertices in V_p are called *p-median vertices*.

A p-median naturally partitions a graph because for each p-median vertex m_j, there is a set of vertices that are nearer to that vertex than any other p-median vertex. The nearest neighbor partition cells P_j, for $1 \leq j \leq p$, are:

$$P_j = \{v_i : d(v_i, m_j) \leq d(v_i, m_k), i = 1, 2 \ldots, 1 \leq k \leq p\}.$$

If $d(v_i, m_j) = d(v_i, m_k)$, the vertex v_i is usually assigned to the p-median vertex with the smaller index. The total weighted distance is

$$D = \sum_{i=1}^{n} \sum_{v_j \in P_j} w_j d(v_i, v_j).$$

2.2 Integer Programming Formulation

The p-median problem is typically formulated as an integer program (IP) [94]. Let ξ_{ij} be an *allocation variable* such that

$$\xi_{ij} = \begin{cases} 1 & \text{if vertex } x_j \text{ is allocated to vertex } x_i \\ 0 & \text{otherwise.} \end{cases}$$

Then, the IP is

$$\begin{aligned} \min \quad & Z = \sum_{ij} W_{ij} \xi_{ij} \\ \text{subject to} \quad & \sum_i \xi_{ij} = 1, \text{ for } j = 1, \ldots, n, && (1) \\ & \sum_i \xi_{ii} = p, && (2) \\ & \xi_{ij} \leq \xi_{ii}, \quad \forall i, j = 1, \ldots, n, && (3) \\ & \xi_{ij} \in \{0, 1\}. && (4) \end{aligned}$$

Constraint (1) ensures that each vertex is allocated to one and only one element in the p-element subset. Constraint (2) guarantees that there are p vertices allocated to themselves, which forces the cardinality of the p-median subset to be p. Constraint (3) states that vertices cannot be allocated to non p-median vertices. The p-median is $\{v_i \mid \xi_{ii} = 1\}$. The most common relaxation is to replace constraint (4) with

$$\xi_{ij} \geq 0.$$

2.3 Complexity

While reading the literature, it was noticed that the statement, "the p-median problem is NP-hard," was often misunderstood. The problem is NP-hard on general graphs and networks for an arbitrary p (where p is a variable). Polynomial time algorithms exist for arbitrary p when the network is a tree [50, 65]. If p is fixed, the p-median problem on a general network is solvable in polynomial time [50]. This does not mean that the fixed p problem is computationally easy. Several heuristic techniques have been developed for the problem on general networks with arbitrary p, and these heuristics are often used for the fixed p problem to reduce computation time or problem size.

3 Criteria

Given the above problem definitions and distinctions, we developed a set of criteria to decide the research contained herein. The most important criteria is that the paper must focus directly on solving the p-median problem—i.e. on solution methodologies, problem formulations or complexity. We *include* articles that:

- Focus directly on solving the p-median problem

- Deal with minisum problems
- Define the p-median problem on a network or graph
- Restrict possible median locations to the set of vertices
- Do not have fixed costs
- Ensure that the set of possible locations is identical to the entire set of vertices or demand points
- Involve capacitated or uncapacitated variations of the p-median problem
- Deal with the metric p-median problem, where the weighted distance is a metric

We *exclude* articles that:

- Deal with minimax or p-center problems
- Locate medians at continuous points in the plane or deal with a continuous variation of the p-median problem
- Solve problems with stochastic travel costs (edge weights) and/or demands
- Solve problems with multiple services and/or commodities
- Deal with multiple objectives
- Do not have constraints on the number of medians or facilities
- Solve a p-median variation with maximum distance constraints
- Place medians among previously existing medians

We make exception to these criteria when a paper contains a solution method that was developed for the UFLP or other related problems but has been historically been applied to solving the p-median problem. We did not include citations for papers that we could not obtain, so there are a few papers (primarily technical reports) that may fit the above criteria that are not present. A complete list of papers that were considered for this work, including those that do not fit our criteria, is located at (`http://lagrange.math.trinity.edu/tumath/research/studtechreport.shtml`).

4 Solution Methods

This paper is not intended to be a review of mathematical programming methods, so for the sake of brevity we assume that the reader is familiar with standard techniques. In this section we simply give an overview of the techniques that have been applied to the p-median problem. Table 1 lists references by solution methodology or by subject. Table 2 gives the number of papers for each 5 year period and shows a dramatic increase in interest, particularly over the last 5 years. A complete timeline is found in Appendix A.

Enumeration and heuristics were the earliest techniques proposed. There were three primary early heuristics: Greedy [69], Alternate [77] and Vertex Substitution [114]. These prototypical heuristics have been combined with each other and numerous other techniques to form new solution methods. Of the heuristics, vertex substitution is the most common, and to this day is a typical solution method. Another heuristic of note is the branch-and-bound heuristic [64] (not to be confused with the technique for solving LP relaxations).

Several methods have been used for solving LP relaxations of the IP formulation, including branch-and-bound, dual ascent and subgradient optimization. Surrogate relaxation techniques have also been explored.

Metaheuristics and approximation algorithms are the predominant techniques explored in the literature over the last few years. The most common metaheuristics are Genetic Algorithms, Variable Neighborhood Search, Tabu Search, Heuristic Concentration, Simulated Annealing and Neural Networks. Several approximation algorithms have been produced, providing different bounds on the attainable solution quality.

Besides testing random instances, many of the papers present computational research on data from two primary test banks. These are the OR-Library (`http://people.brunel.ac.uk/~mastjjb/jeb/info.html`) and TSPLIB (`http://elib.zib.de/pub/Packages/mp-testdata/tsp/tsplib/tsplib.html`).

Heuristics	
Vertex Substitution	[4] [33] [34] [35] [39] [59] [90] [91] [92] [98] [99] [105] [114] [118] [119]
Other Heuristics	[2] [9] [15] [27] [40] [61] [64] [66] [69] [77] [89] [106] [112] [117]
Metaheuristics	
Variable Neighborhood Search	[29] [30] [48] [57] [58]
Heuristic Concentration	[101] [102] [100] [104]
Genetic Algorithms	[1] [14] [22] [28] [37] [42] [60] [71] [73] [76] [84]
GRASP Metaheuristic	[93]
Scatter Search	[49]
Tabu Search	[52] [96] [102] [107] [116]
Simulated Annealing	[23] [70] [95]
Neural Network	[79] [80]
Approximation Algorithms	[3] [18] [19] [20] [21] [62] [63] [68] [72] [81] [115] [120]
LP Relaxation	[6] [7] [11] [12] [13] [16] [17] [25] [27] [32] [36] [38] [41] [43] [45] [47] [51] [56] [67] [83] [86] [87]
Surrogate Relaxation	[74] [75] [108] [109] [110] [111]
Surveys	[31] [55] [82]
IP Formulations and Reductions	[5] [8] [10] [26] [38] [94] [103]
Complexity	[50] [65]
Graph Theoretic	[24] [44] [46] [65] [85] [113]
Enumeration	[54] [78]
Other	[53] [88]

Table 1: Papers by Type of Solution Method or Subject

References

[1] O. Alp, E. Erkut, and Z. Drezner. An efficient genetic algorithm for the p-median problem. *Annals of Operations Research*, 122:21–42, 2003.

Describes a simple and fast genetic algorithm that models the indices of vertices in the solution as genes of a chromosome. The fitness function is the objective function. Whereas traditional genetic methods use a crossover approach, this method creates a union of the parent's chromosomes, creating an infeasible solution with $m > p$ genes. A greedy deletion heuristic is applied to decrease the number of genes until p genes are left. No mutation operator is used.

[2] A. Ardalan. A comparison of heuristic methods for service facility locations. *International Journal of Operations and Production Management*, 8(2):52–58, 1988.

Introduces a heuristic for a service facility location problem. In [90], this heuristic is adapted to the p-median problem and compared to vertex substitution [114].

[3] V. Arya, N. Garg, R. Khandekar, V. Pandit, A. Meyerson, and K. Munagala. Local search heuristics for k-median and facility location problems. *SIAM Journal on Computing*, 33(3):544–562, 2004.

Defines locality gap as the maximum ratio of a locally optimal solution to the global optimum. The fact that local search, constrained by swapping only one vertex at a

<1970	7
1970–1974	9
1975–1979	12
1980–1984	9
1985–1989	6
1990–1994	11
1995–1999	24
2000–2005	42

Table 2: Number of Papers on Solution Methods Every 5 Years

time, has a locality of gap of 5 is established. When the search is allowed to swap up to p vertices at a time, the locality gap improves to $3 + 2/p$.

[4] J. Ashayeri, R. Heuts, and B. Tammel. A modified simple heuristic for the p-median problem, with facilities design applications. *Robotics and Computer-Integrated Manufacturing*, 21(4):451–464, 2005.

A simple modification to vertex substitution [114] is presented. The authors show that in some cases, this modification leads to improved solutions. The modified method, called *adjusted vertex substitution (AVS)*, is implemented with multiple starting points to avoid finding local optima. Computational results show that these methods are generally able to improve the solutions found with the original vertex substitution method.

[5] P. Avella and A. Sassano. On the p-median problem polytope. *Mathematical Programming*, 89(3):395–411, 2001.

Defines two classes of facet-defining inequalities and uses them in a cutting-plane algorithm. The authors report using this method to solve several test problems to optimality.

[6] P. Avella, A. Sassano, and I. Vasil'ev. Computational study of large-scale p-median problems. Technical Report 08-03, Università di Roma "La Sapienza", 2003.

Presents a branch-and-price-and-cut algorithm to solve large-scale instances of the p-median problem. This involves a column-and-row generation strategy to solve a relaxed LP and cutting planes to strengthen the formulation. The authors report computational results on large problems involving at least 900 nodes.

[7] P. Avella, A. Sassano, and I. Vasil'ev. A heuristic for large-scale p-median instances. *Electronic Notes in Discrete Mathematics*, 13:14–17, 2003.

Presents a three step heuristic for large p-median problems. First, a Lagrangian relaxation is solved by using a subgradient algorithm, producing a lower bound. Second, a subset of "promising" variables is selected to create a core problem. Third, branch-and-bound is run on the core problem to find an upper bound. Computational results comparing this method with the GRASP metaheuristic [93] are presented.

[8] P. Avella and A. Sforza. Logical reduction tests for the p-median problem. *Annals of Operations Research*, 86:105–115, 1999.

Proposes reduction tests for the p-median problem and shows their impact on problem size. These tests are based on characteristics of p-median solutions. These tests are implemented in Lagrangian heuristics and used to solve randomly generated instances and problems from the OR-Library.

[9] K. R. Baker. A heuristic approach to locating a fixed number of facilities. *Logistics and Transportation Review*, 10(3):195–205, 1974.

Presents a heuristic based on dynamic programming concepts for the p-median problem with fixed p. This is a suboptimal method concerned only with obtaining a good solution rapidly.

[10] M. Balinski. Integer programming, methods, uses and computation. *Management Science*, 12(3):253–313, 1965.

Provides an early integer programming formulation of the plant location problem that has historically been adapted to the p-median problem.

[11] J. E. Beasley. A note on solving large p-median problems. *European Journal of Operational Research*, 21(2):270–273, 1985.

Enhances the tree search algorithm in [25]. The enhanced algorithm was combined with a Cray supercomputer to solve large problem instances (up to 900 nodes).

[12] J. E. Beasley. Lagrangean heuristics for location problems. *European Journal of Operational Research*, 65(3):383–399, 1993.

Introduces a Lagrangian heuristic that combines vertex substitution [114] with subgradient optimization and the methods in [27] and [86]. Computational results are presented comparing this method to stand alone vertex substitution.

[13] C. Beltran, C. Tadonki, and J.-Ph. Vial. Solving the p-median problem with a semi-lagrangian relaxation. Technical report, Logilab, HEC, University of Geneva, 2004.

Describes a semi-Lagrangian relaxation that, under certain conditions, closes the integrality gap for any linear combinatorial problem with equality constraints. This process inserts Constraint (1) (see Section 2.2) into the inner minimization problem of the maxmin problem and relaxes it to a "less than or equal to" constraint. For large enough multipliers, this relaxation closes the integrality gap. The p-median problem is easier to solve, however, when the multipliers are small. The insight here is to increase the Lagrangian multipliers but keep them small enough to keep the inner problem easy until an integer optimal solution is found. Computational results are presented, showing that this method solved to optimality several "easy" problems, and improved the best known dual bounds on several unsolved "difficult" problems. The method was able to solve to optimality one of the previously unsolved "difficult" problems.

[14] B. Bozkaya, J. Zhang, and E. Erkut. An efficient genetic algorithm for the p-median problem. In Z. Drezner and H. Hamacher, editors, *Facility Location: Applications and Theory*, pages 179–205. Berlin: Springer, 2002.

A genetic algorithm that models solutions with chromosomes is developed. Each gene of the chromosome is an index of a p-median vertex. Three crossover operators are presented and tested. These variations are compared with the genetic algorithm in [60] and favorable results are reported. The genetic algorithm here improves solutions slowly but steadily, and the authors claim that it is less likely than vertex substitution [114] to be trapped in local optima.

[15] M. E. Captivo. Fast primal and dual heuristics for the p-median location problem. *European Journal of Operational Research*, 52(1):65–74, 1991.

Combines the greedy algorithms of [27] and [69] with the alternate heuristic in [77], creating a new heuristic called *GreedyG*. The author develops a dual-based procedure based upon the method in [41] and a heuristic that finds a primal solution based upon a good dual solution and complementary slackness conditions. These techniques are used to develop bounds to compare the results of several heuristics. GreedyG is compared to the alternate heuristic [77], vertex substitution [114], and hybrid techniques formed by merging a greedy heuristic with the alternate and vertex substitution heuristics.

[16] A. Ceselli. Two exact algorithms for the capacitated p-median problem. *4OR: Quarterly Journal of the Belgian, French and Italian Operations Research Society*, 1:319–340, 2003.

Introduces branch-and-bound and branch-and-price methods for solving the capacitated variation of the p-median problem. The branch-and-bound technique uses Lagrangian relaxation and subgradient optimization, and the branch-and-price algorithm uses column generation. The authors note that the ratio p/n strongly affects the behavior of these algorithms. The performance of the branch-and-bound technique worsened quickly as the ratio increased. The branch-and-price algorithm was found to be more stable.

[17] A. Ceselli and G. Righini. A branch-and-price algorithm for the capacitated p-median problem. *Networks*, 45(3):125–142, 2005.

Presents two IP formulations of the capacitated p-median problem and compares the lower bounds on the two corresponding LP relaxations. The authors offer two separate branching strategies for a branch-and-price algorithm. The first is branching on binary variables, and the second is branching on semiassignment constraints. The latter involves partitioning the set of candidate medians for each vertex into two subsets of balanced cardinality and branching in these subsets. The authors found this method to be more effective than branching on binary variables. The authors also present several pricing methods and extensive experimental results.

[18] M. Charikar, C. Chekuri, A. Goel, and S. Guha. Rounding via trees: deterministic approximation algorithms for group steiner trees and k-median. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 114–123, 1998.

A deterministic approximation algorithm that matches the best known randomized approximation is presented. This algorithm has an approximation ratio of $O(\log p \log\log p)$.

[19] M. Charikar and S. Guha. Improved combinatorial algorithms for the facility location and k-median problems. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, pages 378–388, 1999.

This modifies and improves the algorithm in [63], providing a 4-approximation algorithm for the p-median problem in $O(n^2(L+n)\log n)$ time, where L is the number of bits needed to represent connecting cost.

[20] M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys. A constant-factor approximation algorithm for the k-median problem. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 1–10, 1999.

Provides a 20/3-approximation algorithm for the metric p-median problem. The authors claim that this is the first constant factor approximation algorithm presented for the p-median problem. The algorithm solves a relaxed LP and builds a collection of trees from the solution. It then solves the problem optimally on this collection.

[21] J. Cheriyan and R. Ravi. Approximation algorithms for network problems, lecture notes, 1998.

Provides, for any $\epsilon > 0$, an approximation algorithm for the p-median problem that finds a solution within $(1 + \epsilon)$ of the optimal value and using no more than $(1 + (1/\epsilon)H(n)p)$ medians, where $H(n)$ is the n^{th} harmonic number.

[22] Y. Chiou and L. W. Lan. Genetic clustering algorithms. *European Journal of Operational Research*, 135(2):413–427, 2001.

Applies cluster seed points to solve the p-median problem. A genetic algorithm is used to select the most suitable cluster seeds (medians). The remaining vertices are assigned to clusters according to their similarity to the cluster seeds or their ability to improve the objective function. Computational results are presented.

[23] F. Chiyoshi and R. D. Galvão. A statistical analysis of simulated annealing applied to the p-median problem. *Annals of Operations Research*, 96:61–74, 2000.

Presents an algorithm that combines vertex substitution [114] with simulated annealing. The algorithm uses vertex substitution to find pairs of vertices to consider for possible exchange, instead of randomly choosing pairs of vertices. The authors adopt a cooling structure that allows for temperature adjustments rather than just temperature reductions.

[24] N. Christofides. *Graph Theory: An Algorithmic Approach.* Academic Press, New York, 1975.

Contains a section on solution methods for the p-median problem, summarizing many of the methods present in 1979 and introducing a direct tree search algorithm. This algorithm allocates vertices v_1 to v_n sequentially to their nearest neighboring vertex. Several informed observations are used to limit the number of alternative possible allocations of a vertex v_i at any stage. A method for calculating a lower bound on the objective function to further limit the search is also presented.

[25] N. Christofides and J. E. Beasley. A tree search algorithm for the p-median problem. *European Journal of Operational Research*, 10(2):196–204, 1982.

Uses the Lagrangian duals of two LP relaxations and subgradient optimization to develop two lower bounds. These bounds are combined with upper bounds found with a heuristic to develop penalty tests that often reduce the size of the problem. The bounds and penalty tests are incorporated within a tree search algorithm. Computational results are presented, comparing this method to the one in [41].

[26] R. L. Church. COBRA: A new formulation of the classic p-median location problem. *Annals of Operations Research*, 122:103–120, 2003.

Presents an alternate p-median formulation called *COndensed Balinski constraints with the Reduction of Assignment variables (COBRA).* As the name suggests, this formulation is usually smaller than the classic p-median formulation of [10], reducing the number of variables by up to 60%. COBRA is also smaller than the formulations in [94] and [103].

[27] G. Cornuéjols, M. L. Fisher, and G. L. Nemhauser. Location of bank accounts to optimize float: an analytical study of exact and approximate algorithms. *Management Science*, 23:789–810, 1977.

Presents a variety of solution techniques, including relaxations and heuristics. Also presents upper bounds for the worst case performance of the greedy interchange heuristic and the relaxations. The main result is that the relative error of the dual bound and the greedy heuristic is bounded above by $1/e$.

[28] E. S. Correa, M. T. A. Steiner, A. A. Freitas, and C. Carnieri. A genetic algorithm for solving a capacitated p-median problem. *Numerical Algorithms*, 35:373–388, 2004.

Develops a genetic algorithm for the capacitated p-median problem. The process assigns vertices to the nearest median that is not already full. A new genetic operator called *heuristic hypermutation* is introduced. This operator improves the fitness of a certain percentage of genes. Computational results of the algorithm with and without heuristic hypermutation are compared with the results of a tabu search heuristic.

[29] T. G. Crainic, M. Gendreau, P. Hansen, and N. Mladenović. Parallel variable neighborhood search for the p-median. *Les Cahiers du GERAD*, G-2003-4, 2003.

Summarizes the methods in [48] and offers a new strategy entitled *Co-operative Neighborhood VNS (CNVNS)*. This is a master-slave procedure where the master process initiates several VNS threads, each of which randomly chooses a neighborhood to explore. Threads report improved solutions to the master process in order to find an overall solution.

[30] T. G. Crainic, M. Gendreau, P. Hansen, and N. Mladenović. Cooperative parallel variable neighborhood search for the p-median. *Journal of Heuristics*, 10(3):293–314, 2004.

Describes the Cooperative Neighborhood VNS (CNVNS) strategy found in [29].

[31] M. S. Daskin. *Network and Discrete Location: models, algorithms, and applications.* John Wiley & Sons, Inc., New York, 1995.

Contains a chapter entitled "Median Problems" which describes in detail three classes of heuristics: myopic (greedy), exchange (vertex substitution), and neighborhood search. Lagrangian relaxation is also discussed. Computational results of these approaches are given.

[32] I. R. de Farias, Jr. A family of facets for the uncapacitated p-median polytope. *Operations Research Letters*, 28(4):161–167, 2001.

Presents a non-trivial family of facet-defining inequalities for the convex hull of the feasible set of the p-median problem. The author develops a separation heuristic for this family of inequalities and then uses the inequalities as cuts in a branch-and-cut algorithm. For many problem instances, this method dramatically reduced the number of nodes. For small instances, however, it was inefficient.

[33] P. J. Densham and G. Rushton. Designing and implementing strategies for solving large location-allocation problems with heuristic methods. Technical Report 91-10, National Center for Geographic Information and Analysis, Buffalo, NY, 1991.

Discusses how to implement vertex substitution [114], specifically detailing speedup strategies. These include minimizing the volume of data and access times to that data, and exploiting the spatial structure of the problem to reduce the number of vertices to check after each substitution. The authors also discuss using an *allocation table* as an alternative method of keeping track of the changes in the objective function as vertices are substituted.

[34] P. J. Densham and G. Rushton. A more efficient heuristic for solving large p-median problems. *Papers in Regional Science*, 71(3):307–329, 1992.

Uses necessary conditions of optimal p-median solutions to develop a more efficient variant of vertex substitution [114]. This information is used to create an informed spatial search procedure that is more efficient and more effective than the original naive spatial search procedure. This procedure is implemented in a new method called Global/Regional Interchange Algorithm (GRIA).

[35] P. J. Densham and G. Rushton. Strategies for solving large location-allocation problems by heuristic methods. *Environment and Planning A*, 24:289–304, 1992.

Provides a method for solving large location-allocation problems, including the p-median problem. The method is based on vertex substitution [114] and works by exploiting the spatial structure of location-allocation problems.

[36] O. du Merle and J.-Ph. Vial. Proximal ACCPM, a cutting plane method for column generation and Lagrangian relaxation: application to the p-median problem. Technical Report 2002.23, HEC Genève, Université de Genève, 2002.

Presents a variant of the analytic center cutting plane method (ACCPM) that incorporates two features from the Bundle method. A proximal term is added to the logarithmic barrier function and a step to reduce the number of columns in the localization set is implemented. Extensive computational results are presented.

[37] J. Dvorett. Compatibility-based genetic algorithm: A new approach to the p-median problem. Technical report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1999.

Presents a compatibility measure that helps genetic algorithms to make better guesses when selecting solutions for the reproduction and crossover stages. The compatibility measure considers candidate parents together rather than choosing the two parents independently. Computational results are presented, comparing this method with greedy heuristics [27] [69], vertex substitution [114], and the alternate heuristic [77]. The authors show that in some instances, the genetic algorithm outperforms other heuristics.

[38] M. Efroymson and T. Ray. A branch-bound algorithm for plant location. *Operations Research*, 14(3):361–368, 1966.

Provides an early integer programming formulation of the plant location problem that has been historically adapted to the p-median problem. This paper presents a branch-and-bound technique that is used to solve the relaxed linear program.

[39] S. Eilon and R. D. Galvão. Single and double vertex substitution in heuristic procedures for the p-median problem. *Management Science*, 24(16):1763–1766, 1978.

Defines a set V_p of p vertices to be λ-optimal if by substituting any λ vertices in V_p with λ vertices in V no reduction in the objective value can be obtained. This is an extension of vertex substitution [114]. The calculation effort increases rapidly with λ and the authors warn against using algorithms with $\lambda > 2$. They also combine this method with the vertex addition method in [64]. This method iteratively calculates the $1, 2, \ldots, p$-medians, applying the λ-optimal heuristic on each iteration. Extensive computational results are presented.

[40] A. M. El-Shaieb. A new algorithm for locating sources among destinations. *Management Science*, 20(2):221–231, 1973.

Presents two different ways of calculating lower bounds on total weighted distance to be used in branch-and-bound. Both methods begin with an allocation set $\{S : D\}$ consisting of a set of sources S and a set of destinations D. These sets are empty at first, and on each iteration a location is added to either set, depending on which offers the least lower bound. The corresponding allocation set and its complement are added to the active set. This process continues until the number of sources equals p or the number of destinations equals $(n - p)$. Once either of these occur, the algorithm calculates the total weighted distance and purges any allocation sets. The process continues until there is only one allocation set in the active set, which is the final solution.

[41] D. Erlenkotter. A dual-based procedure for uncapacitated facility location. *Operations Research*, 26(6):992–1009, 1978.

Presents a dual ascent algorithm commonly known as DUALOC for the UFLP. This method is often applied to the p-median problem. It balances the fixed costs of the median vertices with the variable costs of allocation when finding the optimal solution. To apply this algorithm to the p-median problem one must first find a fixed cost value that results in the desired number of p median vertices. Unfortunately, such a fixed cost may not exist, so solutions for all values of p cannot be guaranteed by this procedure.

[42] V. Estivill-Castro and R. Torres-Velázquez. Hybrid genetic algorithm for solving the p-median problem. In *SEAL '98: Selected Papers from the Second Asia-Pacific Conference on Simulated Evolution and Learning*, volume 1585 of *Lecture Notes in Computer Science*, pages 19–25, 1999.

Combines genetic algorithms and vertex substitution [114]. This hybrid technique is useful for avoiding local optima that vertex substitution is prone to finding. The authors claim that this method outperforms both ordinary genetic algorithms and stand alone vertex substitution.

[43] R. D. Galvão. A dual-bounded algorithm for the p-median problem. *Operations Research*, 28(5):1112–1121, 1980.

Provides an LP relaxation of the IP formulation, relaxing the integrality constraint $\xi_{ij} \in \{0,1\}$ to $\xi_{ij} \geq 0, i,j = 1,\ldots,n$. A heuristic to solve the dual of this LP relaxation is presented. This method is similar to the method in [41] but is specifically for the p-median problem. The optimal dual solution is used in a branch-and-bound algorithm.

[44] R. D. Galvão. A graph theoretical bound for the p-median problem. *European Journal of Operational Research*, 6(2):162–165, 1981.

Finds lower bounds for a branch-and-bound algorithm. For a non-directed network the length of the shortest spanning tree minus the shortest spanning tree's $(p-1)$ longest links is shown to be a lower bound on the solution. This bound is generalized for directed networks, where the author develops a stronger bound based upon the spanning arborescences of the network.

[45] R. D. Galvão. A note on Garfinkel, Neebe and Rao's LP decomposition for the p-median problem. *Transportation Science*, 15(3):175–182, 1981.

Presents extensive computational results of the LP decomposition method in [51]. Due to the method's degenerate nature, serious convergence problems commonly occurred. Convergence surprisingly occurred more often with randomly generated initial solutions rather than "good" initial solutions found with heuristics.

[46] R. D. Galvão. Rejoinder to "A note on Galvão's a graph theoretical bound for the p-median problem". *European Journal of Operational Research*, 17(1):128, 1984.

The author acknowledges the error in [44] that was pointed out in [85], but emphasizes that the error does not affect his results.

[47] R. D. Galvão and L. A. Raggi. A method for solving to optimality uncapacitated location problems. *Annals of Operations Research*, 18:225–244, 1989.

Presents a 3-stage procedure for solving the p-median problem. The stages are 1) a primal-dual algorithm, 2) subgradient optimization to solve a Lagrangian dual, and 3) a branch-and-bound algorithm. The method is hierarchical, meaning that stages are only activated if the optimal solution was not found in the previous stage.

[48] F. García-López, B. Melián-Batista, J. A. Moreno-Pérez, and J. M. Moreno-Vega. The parallel variable neighborhood search for the p-median problem. *Journal of Heuristics*, 8(3):375–388, 2002.

Offers three parallel computing strategies for speeding up the computation time of a Variable Neighborhood Search (VNS) metaheuristic for the p-median problem. VNS finds an initial local minimum and then systematically or randomly explores increasingly distant neighborhoods. If the procedure finds a better solution, it jumps to it and continues the search until a stopping criteria is met. The first strategy presented here involves parallelizing the local search phase of the algorithm. The second approach involves running an independent VNS procedure on each processor and using the best solution at the end. The third method involves a synchronous master-slave approach.

[49] F. García-López, B. Melián-Batista, J. A. Moreno-Pérez, and J. M. Moreno-Vega. Parallelization of the scatter search for the p-median problem. *Parallel Computing*, 29(5):575–589, 2003.

Describes a population-based metaheuristic technique known as scatter search. This method begins by creating a reference set from a population of solutions and generating subsets of this reference set that are good solutions over the reference set. These solutions are combined to form a new current solution. This solution is then run through a solution improvement procedure. A reference set updating procedure then considers the improved solutions for inclusion in the reference set. Stopping procedures determine when to generate a new reference set or a new population and when to terminate the algorithm.

[50] M. R. Garey and D. S. Johnson. *Computers and Intractibility: A guide to the theory of NP-completeness*. W. H. Freeman and Co., San Francisco, 1979.

Shows that the p-median problem is NP-hard. If p is fixed then it is solvable in polynomial time. It is also solvable in polynomial time for arbitrary p if the graph is a tree.

[51] R. S. Garfinkel, A. W. Neebe, and M. R. Rao. An algorithm for the m-median plant location problem. *Transportation Science*, 8:217–236, 1974.

Models the p-median problem as an integer program and solves the relaxed LP with a decomposition technique. The paper gives a method for resolving non-integer solutions that combines group theoretic and dynamic programming techniques. Computational results are presented, comparing this technique with branch-and-bound techniques similar to those in [67].

[52] F. Glover. Tabu search for the p-median problem, unpublished manuscript, 1990.

Proposes an implementation of tabu search for the p-median problem and details its associated interchange, strategic oscillation, candidate list, intensification and diversification strategies.

[53] S. L. Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459, 1964.

Shows that the absolute median of a graph G is always located at a vertex of a graph. Thus, to find the optimum location for a switching center in a communication network one must only search the vertices of the graph or network. The absolute median is equivalent to a 1-median.

[54] S. L. Hakimi. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, 13(3):462–475, 1965.

Generalizes the results in [53] and shows that the optimum collection of p switching centers in a communication network is the p-median of the corresponding weighted graph. Hakimi proved that there is a set of p points, consisting entirely of vertices on a graph, that minimizes the total weighted cost. Hakimi called this set the p-median of a graph and showed that to find it one may examine every p-element subset of the vertices of G and simply keep track of the least weighted distance after every evaluation. He then used this method, known as *direct enumeration*, to find the 3-median of a graph with 10 vertices.

[55] G. Y. Handler and P. B. Mirchandani. *Location on Networks: Theory and Algorithms*. MIT Press, Cambridge, MA, 1979.

Provides an excellent overview of the computational methods for the p-median problem up to 1979. The solution methods are classified into 5 different categories: 1) enumeration, 2) graph theoretic, 3) heuristic, 4) primal-based LP methods, and 5) dual-based LP methods.

[56] P. Hanjoul and D. Peeters. A comparison of two dual-based procedures for solving the p-median problem. *European Journal of Operational Research*, 20(3):387–396, 1985.

Presents and compares two dual-based p-median solution methods that rely on two different Lagrangian relaxations. Constraints (1) and (2), described in Section 2.2, are relaxed. Computational results indicate that these two procedures solve large-scale p-median problems successfully.

[57] P. Hansen and N. Mladenović. Variable neighborhood search for the p-median. *Location Science*, 5(4):207–226, 1997.

Variable neighborhood search (VNS) is a metaheuristic that involves a systematic change of neighborhood within a local search algorithm. The process involves exploring increasingly distant neighborhoods to avoid local minima. The authors provide a parameter-free VNS heuristic for the p-median and compare it to tabu search and the greedy interchange heuristic.

[58] P. Hansen, N. Mladenović, and D. Perez-Brito. Variable neighborhood decomposition search. *Journal of Heuristics*, 7(4):335–350, 2001.

Presents a variant of Variable Neighborhood Search (VNS) designed to enhance the efficiency of VNS on larger problems. The only difference between VNS and the new method, called Variable Neighborhood Decomposition Search (VNDS), is that during the local search phase, VNDS solves a subproblem instead of applying the local search to the whole solution space. This method is applied to the p-median problem and computational results are presented comparing VNDS with the fast interchange method in [119] and a Reduced VNS method.

[59] M. Horn. Analysis and computational schemes for p-median heuristics. *Environment and Planning A*, 28:1699–1708, 1996.

Provides a mathematical analysis of a computational scheme for the vertex substitution heuristic [114] and GRIA [34]. The author decided to use concise set-theory notation rather than the binary-array notation used in [34]. It is shown that vertex substitution and GRIA do not generally find solutions with equivalent local optimality.

[60] C. M. Hosage and M. F. Goodchild. Discrete space location-allocation solutions from genetic algorithms. *Annals of Operations Research*, 6:35–46, 1986.

Provides the first application of genetic algorithms to the p-median problem. The authors note strengths and weaknesses of this implementation, showing that it is likely to be trapped in a local optimum when solving location problems.

[61] M. Hribar and M. S. Daskin. A dynamic programming heuristic for the p-median problem. *European Journal of Operational Research*, 101(3):499–508, 1997.

Provides a heuristic that is a cross between a greedy heuristic [69] and a dynamic programming algorithm. This polynomial-time method finds several solutions and determines how often particular points are used in solutions, thus finding points that have a high or low likelihood of being in the optimal solution. Computational results are presented, comparing the solutions obtained with this method to known optimal solutions for several test problems.

[62] K. Jain, M. Mahdian, and A. Saberi. A new greedy approach for facility location problems. In *STOC '02 May 19-21, 2002, Montreal, Quebec, Canada*, 2002.

Presents a lower bound on the approximability of the metric p-median, showing that it may not be approximated with a factor strictly smaller than $1 + 2/\epsilon$.

[63] K. Jain and V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. *Journal of the ACM*, 48(2):274–296, 2001.

Deals only with the metric p-median problem, where the weighted cost W_{ij} is a metric (satisfies the triangle inequality). The authors provide an approximation algorithm with an approximation guarantee of 6. The algorithm runs in $O(n \log n(L + \log n))$ time, where L is the number of bits needed to represent connecting cost.

[64] P. Järvinen, J. Rajala, and H. Sinervo. A branch-and-bound algorithm for seeking the p-median. *Operations Research*, 20(1):173–178, 1972.

Applies a branch-and-bound technique to finding the p-median by attempting to find the $(n - p)$ vertices that are not in the p-median. It does so by starting from an availiable $(t - 1)$-median solution and adding to it the vertex that provides the maximum reduction in the objective value as $(t - 1) \to t$. This process begins with $(t - 1) = 1$ and ends when $t = (n - p)$. The insight is that vertices are removed that would not be a good choice for the p-median because their removal results in a reduction in the total weighted distance. Extensive computational results are presented.

[65] O. Kariv and S. L. Hakimi. An algorithmic approach to network location problems. II. The p-medians. *SIAM Journal on Applied Mathematics*, 37(3):539–560, 1979.

Offers a proof that the p-median problem is NP-hard on a general network based on the fact that the dominating set problem (which is NP-hard) is polynomial time reducible to the p-median problem. The authors provide an $O(n^2p^2)$ algorithm for the p-median problem on a tree where p is arbitrary.

[66] B. M. Khumawala, A. W. Neebe, and D. G. Dannenbring. A note on El-Shaieb's new algorithm for locating sources among destinations. *Management Science*, 21(2):230–233, 1974.

Compares the solution method in [40] with the methods in [51], [67], and [114].

[67] M. B. Khumawala. An efficient branch-and-bound algorithm for the warehouse location problem. *Management Science*, 18(12):718–731, 1972.

Presents a branch-and-bound technique for the warehouse location problem that has been applied to the p-median problem. The method relaxes the integrality constraint $\xi_{ij} \in \{0, 1\}$ and obtains an optimal solution Z^*. If all ξ_{ij} are integers, then Z^* is the final solution to the problem without relaxation. If fractional solutions are found, branch-and-bound is used to eliminate them.

[68] M. R. Korupolu, C. G. Plaxton, and R. Rajaraman. Analysis of a local search heuristic for facility location problems. *Journal of Algorithms*, 37(1):146–188, 2000.

Presents a summary of the work done in the field of approximation algorithms for the p-median problem. This paper deals only with the metric p-median problem. The authors show that the local search technique in [69] yields a polynomial time algorithm that, for any $\epsilon > 0$, computes a solution using at most $(3+(5/\epsilon))p$ medians with cost at most $(1+\epsilon)$ times the cost of an optimal solution with at most p facilities.

[69] A. A. Kuehn and M. J. Hamburger. A heuristic program for locating warehouses. *Management Science*, 9(4):643–666, 1963.

Introduces a greedy heuristic for the warehouse location problem that has historically been applied to the p-median problem. Let M be the set of potential warehouse locations and let N be the number of locations to evaluate at each iteration. The greedy heuristic initially chooses N locations that maximize the cost savings of replacing these N locations with warehouses. It then considers each of these locations individually and calculates the total distribution cost. Any location that does not reduce the total cost is eliminated from further consideration. The location that gives the least cost is assigned a warehouse, and any remaining locations go back to the list of possible locations to test. This process is repeated until all elements of the original list of potential warehouses have either been eliminated or assigned as a warehouse.

[70] T. V. Levanova and M. A. Loresh. Algorithms of ant system and simulated annealing for the p-median problem. *Automation and Remote Control*, 65(3):431–438, 2004.

Explores the implementation of an ant system and a simulated annealing algorithm. Ant system algorithms were suggested by the ability of ants to find the shortest path from an ant hill to food by using pheromones. The theory is that over time, the shortest path will have the greatest amount of pheromone, and will therefore be the most probable path. The authors use this insight and a simulated annealing algorithm to solve the p-median problem and present computational results.

[71] A. Lim and Z. Xu. A fixed-length subset genetic algorithm for the p-median problem. *Lecture Notes In Computer Science*, 2724:1596–1597, 2003.

The fixed-length subset genetic algorithm represents candidate solutions by a fixed-length subset. Actual computational results are not presented, but the authors claim that their method outperforms the traditional genetic algorithm and is able to find solutions very close to optimal for most problems.

[72] J.-H. Lin and J. S. Vitter. Approximation algorithms for geometric median problems. *Information Processing Letters*, 44:245–249, 1992.

Deals only with the metric p-median problem. The authors present a polynomial-time algorithm that, for any $\epsilon > 0$, finds a solution of no more than $2(1 + \epsilon)$ times the optimal cost and of at most $(1 + (1/\epsilon))p$ median vertices.

[73] L. A. N. Lorena and J. C. Furtado. Constructive genetic algorithm for clustering problems. *Evolutionary Computation*, 9(3):309–328, 2001.

The constructive genetic algorithm presented here differs from the traditional genetic algorithm in that it uses a dynamic population. Two separate fitness functions are described and the clustering problems are formulated as bi-objective optimization problems. Computational results are presented.

[74] L. A. N. Lorena and E. L. F. Senne. Local search heuristics for capacitated p-median problems. *Networks and Spatial Economics*, 3:409–419, 2003.

Combines the Lagrangian/surrogate relaxation techniques in [108] with some local search heuristic techniques to produce a new method for solving the capacitated p-median problem. The heuristic techniques are used to improve solutions that are made feasible by the Lagrangian/surrogate process and involve swapping medians within clusters and reallocating vertices. This is repeated until no further improvement is made.

[75] L. A. N. Lorena and E. L. F. Senne. A column generation approach to capacitated p-median problems. *Computers and Operations Research*, 31(6):863–876, 2004.

Applies the column generation and Lagrangian/surrogate techniques in [109] to the capacitated p-median problem.

[76] V. Maniezzo, A. Mingozzi, and R. Baldacci. A bionomic approach to the capacitated p-median problem. *Journal of Heuristics*, 4:263–280, 1998.

Presents a metaheuristic technique similar to a genetic algorithm for the capacitated p-median problem. The technique is known as a bionomic algorithm and differs from genetic algorithms in how the parent set is obtained. Extensive computational results are presented.

[77] F. E. Maranzana. On the location of supply points to minimize transport costs. *Operations Research Quarterly*, 15(3):261–270, 1964.

Presents a heuristic for finding the p-median that, after each iteration, yields a collection of p vertices that is guaranteed to either reduce or leave unchanged the weighted distance. The algorithm begins with an arbitrary p vertices of the graph and partitions the remaining vertices into corresponding nearest neighborhood cells. The algorithm then determines a "center of gravity" c_j for each partition cell P_j. If $m_j = c_j, \forall j$, then the algorithm terminates, and m_j is the p-median. Otherwise, set $m_j = c_j$ and repeat the process, re-creating partition cells based upon the new vertices. The author shows that with respect to each iteration, the total weighted distance D associated with m_j is monotonically non-increasing, but the algorithm does not always converge to an optimum.

[78] R. E. Marsten. An algorithm for finding almost all the medians of a network. Technical Report 23, Center for Math Studies in Economics and Management Science, Northwestern University, 1972.

Shows that for a network of n nodes every p-median $(1 \leq p \leq n)$ is an extreme point of a polyhedron. An algorithm that tours these extreme points is presented. Some extreme points correspond to p-median solutions with a fractional value of p. Furthermore, the tour may not hit a p-median for every value of p. Aside from these exceptions the algorithm is shown to create a complete set of medians.

[79] E. D. Merino and J. M. Pérez. An efficient neural network algorithm for the p-median problem. In *IBERAMIA 2002: Proceedings of the 8th Ibero-American Conference on AI*, volume 2527 of *Lecture Notes in Computer Science*, pages 460–469, 2002.

Presents an alternate formulation of the p-median problem and then applies a Hopfield neural network model to solve it. The problem formulation involves two types of neurons, one for location and the other for allocation. The constrained formulation is turned into an unconstrained problem by introducing a penalty function that penalizes the objective function if constraints are violated. Computational results are presented comparing this method with vertex substitution [114].

[80] E. D. Merino, J. M. Pérez, and J. J. Aragonés. Neural network algorithms for the p-median problem. In *ESANN 2003: Proceedings of 11th European Symposium on Artificial Neural Networks, Belgium, April 2003*, pages 385–391, 2003.

Offers three different variations of a neural network algorithm for the p-median problem. The objective function is modelled as an energy function, which is guaranteed to decrease or remain unchanged as the system changes according to a given dynamical rule. The different variations, *iterative*, *agglomerative*, and *stepwise*, are tested against vertex substitution [114] and the results are presented.

[81] R. Mettu. *Approximation Algorithms for NP-Hard Clustering Problems.* PhD thesis, Department of Computer Science, University of Texas at Austin, 2002.

Develops a randomized $O(1)$-approximate algorithm for the p-median problem that runs in $O(n(p+\log n)+p^2\log^2 n)$ time. For a wide range of p values, i.e $\log n \leq p \leq \frac{n}{\log^2 n}$, the complexity is shown to reduce to $O(np)$.

[82] P. B. Mirchandani and R. Francis, editors. *Discrete location theory.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1990.

Contains a chapter titled “The p-Median Problem and Generalizations.” The author describes the classical p-median problem and discusses variations such as the p-median problem on oriented networks, probabilistic costs and demands, and multidimensional networks. The author then gives a survey of solution methods, classifying them into the same five categories as [55].

[83] P. B. Mirchandani, A. Oudjit, and R. T. Wong. ‘Multidimensional’ extensions and a nested dual approach for the m-median problem. *European Journal of Operational Research*, 21(1):121–137, 1985.

Describes a “nested dual” approach for solving the p-median problem that uses the method in [41] as a subroutine. The problem is first dualized with respect to Constraint (2) (see Section 2.2). The Lagrangian dual is then solved by a simplex method, where the method in [41] is used to solve the Lagrangian subproblems.

[84] J. A. Moreno-Pérez, J. L. Roda García, and J. M. Moreno-Vega. A parallel genetic algorithm for the discrete p-median problem. *Studies in Locational Analysis*, 7:131–141, 1994.

Implements a variant of the genetic algorithm that does not involve the traditional binary string solution representation (chromosome), but instead represents solutions by a collection of indices of the demand points in V. This algorithm is known as an evolutive algorithm, and it is implemented in parallel, using software to create a parallel virtual machine on a local network. The solution is split into several colonies, each of which is sent to a different processor. This method was tested on TSP data.

[85] L. Morgenstern. A note on Galvão's “A graph theoretical bound for the p-median problem”. *European Journal of Operational Research*, 12(4):404–405, 1983.

Points out that a step of a bound improvement algorithm in [44] is flawed.

[86] S. C. Narula, U. I. Ogbu, and H. M. Samuelsson. An algorithm for the p-median problem. *Operations Research*, 16(5):955–961, 1968.

Introduces an algorithm for solving the IP formulation of the p-median problem when the constraint $\sum_i \xi_{ij} = 1$, for $j = 1, \ldots, n$ is relaxed (i.e. absorbed into the Lagrangian). The algorithm begins with a "good" initial value for the dual variable and moves in a direction where the subgradient is not zero. The algorithm terminates when all the components of the subgradient are zero or when there is no duality gap. Computational results comparing this with the methods in [40], [67], [94], and [114] are presented.

[87] A. W. Neebe and M. R. Rao. A subgradient approach to the m-median problem. Technical Report 75-12, University of North Carolina, Chapel Hill, N.C., 1975.

Presents a subgradient method for solving the dual of the relaxed LP to overcome degeneracy in decomposition techniques. This procedure often terminates with the maximum dual objective value, and in some cases it leads to integer values for all of the primal variables. If not, the bounds may be used in a branch-and-bound algorithm.

[88] H. Pirkul, R. Gupta, and E. Rolland. VisOpt: a visual interactive optimization tool for p-median problems. *Decision Support Systems*, 26(3):209–223, 1999.

Introduces a visual software tool for finding good solutions by attempting to combine human graphical processing power with computer computational power. This paper only discusses the implementation of the software for the capacitated p-median problem. The authors present a comparison of the results of users with this tool to a stand alone heuristic that performs the same task.

[89] N. D. Pizzolato. A heuristic for large-size p-median location problems with application to school location. *Annals of Operations Research*, 50:473–485, 1994.

Introduces a heuristic that begins with an initial collection of p trees generated by the technique in [77] and then reshapes them iteratively through a root interchange method. These steps are fast and provide a near-optimal solution. The author then applies some slower techniques for improving the final solution. This method is compared with vertex substitution [114].

[90] S. Rahman and D. K. Smith. A comparison of two heuristic methods for the p-median problem with and without maximum distance constraints. *International Journal of Operations and Production Management*, 11(6):76–84, 1991.

Compares vertex substitution [114] with a heuristic originally developed for a service facility location problem, found in [2]. The authors call it the Ardalan heuristic. Unlike vertex substitution, once a vertex is chosen as a median, it remains in the solution set until termination. Computational results are presented showing that vertex substitution consistently outperformed the Ardalan heuristic on test instances.

[91] M. G. C. Resende and R. F. Werneck. On the implementation of a swap-based local search procedure for the p-median problem. In R. Ladner, editor, *ALENEX '03: Proceedings of the Fifth Workshop on Algorithm Engineering and Experiments*, pages 119–127, Philadelphia, 2003. SIAM.

Develops a more efficient variant of vertex substitution [114] known as *fast interchange*. Several techniques to hasten the algorithm are developed. These include storing partial results in a matrix to speed up later steps, compressing this matrix, and preprocessing techniques. They report obtaining speedups of up to 3 orders of magnitude over the original fast interchange method.

[92] M. G. C. Resende and R. F. Werneck. A fast swap-based local search procedure for location problems. Technical Report TD-5R3KBH, AT & T Labs Research, 2004.

Provides an implementation of vertex substitution [114] that builds on the implementation in [119]. This method uses *gain*, *loss*, *extra*, and *netloss* functions to determine the value of different substitutions. Values for these functions are stored in data structures so that the best swap may easily be found.

[93] M. G. C. Resende and R. F. Werneck. A hybrid heuristic for the p-median problem. *Journal of Heuristics*, 10(1):59–88, 2004.

Introduces a randomized multistart iterative metaheuristic known as GRASP (Greedy Randomized Adaptive Search Procedure). Each iteration of this process applies a greedy randomized algorithm followed by a local search procedure. A pool of some of the best solutions of previous iterations is stored, and after each iteration, the new candidate solution is combined with one stored solution in a process called *path-relinking*. Once the algorithm terminates, the stored solutions are combined with each other.

[94] C. ReVelle and R. Swain. Central facilities location. *Geographical Analysis*, 2:30–42, 1970.

Provides the first linear programming formulation of the p-median problem. Constraint (4), described in Section 2.2, is relaxed. The authors recommend using a branch-and-bound technique when dealing with fractional solutions. This formulation was tested on several problem instances and no fractional solutions occurred.

[95] G. Righini. A double annealing algorithm for discrete location/allocation problems. *European Journal of Operational Research*, 86(3):452–468, 1995.

Presents a double annealing algorithm and tests it on instances of the p-median problem. The algorithm is a variant of mean-field annealing, which is a deterministic version of simulated annealing. This method splits the annealing process into two synchronized parallel processes. The author shows that if a *deannealing* process is used, allowing the annealing temperature to increase instead of just decrease, then the algorithm experimentally improves.

[96] E. Rolland, D. A. Schilling, and J. R. Current. An efficient tabu search procedure for the p-median problem. *European Journal of Operational Research*, 96(2):329–342, 1996.

Tabu search is a metaheuristic designed to be used in conjunction with heuristics that move nodes between sets, such as vertex substitution [114]. The tabu search method described here involves tabu restrictions, aspiration criteria, diversification and strategic oscillation. The restrictions prevent the search from moving back to previous solutions. Aspiration criteria allow the search to move a node even if it has been restricted. Diversification is used to escape from a local minimum by deterring the search from performing the same moves too often. Strategic oscillation allows the search to intentionally pass through infeasible solutions in order to avoid local optima. Computational results are presented and this method is compared to the method in [34].

[97] K. E. Rosing. An empirical investigation of the effectiveness of a vertex substitution heuristic. *Environment and Planning B*, 24:59–67, 1997.

Presents a comparison of optimal solutions to solutions found with the vertex substitution implementation in [105] on 90 test problems. The values of n and p were varied systematically, and the results showed a degradation of solution quality as either n or p increased.

[98] K. E. Rosing, E. L. Hillsman, and H. Rosing-Vogelaar. A note comparing optimal and heuristic solutions to the p-median problem. *Geographical Analysis*, 11(1):86–89, 1979.

Compares the optimal solution of six test problems to solutions found with the vertex substitution implementation in [105] and solutions found with the alternate heuristic [77]. The vertex substitution heuristic found the optimal solution more frequently than the alternate heuristic.

[99] K. E. Rosing, E. L. Hillsman, and H. Rosing-Vogelaar. The robustness of two common heuristics for the p-median problem. *Environment and Planning A*, 11:373–380, 1979.

Compares the solutions found using the alternate heuristic [77] and vertex substitution [114] with known optimal solutions for six test problems. The authors found that the quality of solutions generated by the alternate heuristic decreases rapidly as p grows. Vertex substitution was more stable.

[100] K. E. Rosing and M. J. Hodgson. Heuristic concentration for the p-median: An example demonstrating how and why it works. *Computers and Operations Research*, 29(10):1317–1330, 2002.

Uses combinatorial and map analysis to show how heuristic concentration [101] can be more effective than vertex substitution [114]. The authors show how vertex substitution may get stuck in traps of particular groups of nodes that heuristic concentration avoids. Good vertex substitution solutions only have one such trap, so a set of several of these solutions is likely to identify all of the nodes required to find the optimum. If all of the optimal nodes are captured in the concentration set, then the second stage of heuristic concentration is able to find the optimal solution.

[101] K. E. Rosing and C. S. ReVelle. Heuristic concentration: Two stage solution construction. *European Journal of Operational Research*, 97(1):75–86, 1997.

The method presented here is a two stage heuristic. In the first stage, the method analyzes several solutions found with vertex substitution [114] and creates a concentration set that contains vertices that have a high probability of being the facilities in the optimal solution. The second stage involves using an exact algorithm to solve a subproblem on this set.

[102] K. E. Rosing, C. S. ReVelle, E. Rolland, D. A. Schilling, and J. R. Current. Heuristic concentration and tabu search: A head to head comparison. *European Journal of Operational Research*, 104(1):93–99, 1998.

Presents computational results comparing tabu search [96] with heuristic concentration [101]. The authors report that heuristic concentration found the superior solution in about 95% of the test instances. When the optimal solution was previously known, it found the optimal solution in about 80% of the cases.

[103] K. E. Rosing, C. S. ReVelle, and H. Rosing-Vogelaar. The p-median and its linear programming relaxation: An approach to large problems. *Journal of the Operational Research Society*, 30(9):815–823, 1979.

Summarizes the plant location IP formulations in [10] and [38], noting the superiority of the former in terms of terminating with little or no non-integer solutions. It also cites [94] as the standard formulation of the p-median problem. Techniques that use these formulations to solve larger problems are developed. These involve removing columns or rows from the fully specified problem.

[104] K. E. Rosing, C. S. ReVelle, and D. A. Schilling. A gamma heuristic for the p-median problem. *European Journal of Operational Research*, 117(3):522–532, 1999.

Presents a heuristic that is a variant of heuristic concentration [101]. The first stage is the same as heuristic concentration, but the second stage applies a 2-opt procedure that assures no exchange of any two median nodes for any two non-median nodes results in further reduction of the objective function. This is followed by a 1-opt procedure. The entire process is a metaheuristic technique called *gamma heuristic*.

[105] G. Rushton and J. A. Kohler. ALLOC: Heuristic solutions to multi-facility location problems on a graph. In G. Rushton, M. Goodchild, and L. Ostresh, editors, *Computer Programs for Location-Allocation Problems*, pages 163–187. Monograph No. 6, Department of Geography, University of Iowa, Iowa City, 1973.

Presents a FORTRAN implementation of two common heuristics for solving the p-median problem. The alternate heuristic [77] and vertex substitution [114] are implemented. The authors note that in 75 test runs, vertex substitution almost invariably outperformed the alternate heuristic.

[106] S. Salhi. A perturbation heuristic for a class of location problems. *Journal of the Operational Research Society*, 48(12):1233–1240, 1997.

Presents an algorithm that allows the cardinality of the p-median to become infeasible, as long as $p \leq q$. This method solves two relaxed problems, looking for the $(p + q)$-median and the $(p - q)$-median. When a solution for the $(p - q)$-median is found, the cardinality of the solution is small enough to nearly guarantee that these vertices will remain in the final solution. Similarly, by dropping some of the vertices in the $(p + q)$-median, a better feasible solution is likely to be found. This process is repeated several times and has a filtering effect where the most effective vertices tend to remain in the best configuration. After a fixed amount of time or after no better solution may be found, a diversification strategy is used to guide the search to other regions that may not be reached otherwise.

[107] S. Salhi. Defining tabu list size and aspiration criterion within tabu search methods. *Computers and Operations Research*, 29(1):67–86, 2002.

Develops a functional representation of the tabu list size, allowing for a dynamic tabu list size. The author develops a softer aspiration criteria that takes into account 1) the tabu status of the attribute for that solution, 2) how much a solution differs from the best found so far, and 3) the change in the objective function. Computational results are presented.

[108] E. L. F. Senne and L. A. N. Lorena. Lagrangean/surrogate heuristics for p-median problems. In M. Laguna and J. Gonzalez-Velarde, editors, *Computing Tools for Modeling, Optimization and Simulation: Interfaces in Computer Science and Operations Research*, pages 115–130. Kluwer Academic Publishers, 2000.

Relaxes Constraint (1) (see Section 2.2) using a surrogate relaxation. This relaxation is not easily solved, so a Lagrangian relaxation is used on the surrogate formulation of the problem. Heuristic techniques are used to find a range of Lagrangian/surrogate multiplier values that improve the bounds of the usual Lagrangian relaxation technique. This approach was able to generate approximate solutions at least as good as the traditional Lagrangian relaxation technique while reducing computational effort for larger problems.

[109] E. L. F. Senne and L. A. N. Lorena. Stabilizing column generation using Lagrangean/surrogate relaxation: an application to p-median location problems. In *Proceedings of the EURO2001 Conference, Erasmus University, Rotterdam, July 2001*, 2001.

Describes relationships between the Lagrangian/surrogate relaxation technique in [108] and column generation. This applies a column generation approach to the p-median problem that uses Lagrangian/surrogate relaxation as an acceleration process, generating new productive sets of columns on each iteration. Computational results show some improvement over the traditional column generation techniques.

[110] E. L. F. Senne and L. A. N. Lorena. Complementary approaches for a clustering problem. In *XI CLAIO: Latin-Iberian American Congress of Operations Research, Concepcin, Chile, October 2002*, 2002.

Combines the Lagrangian/surrogate relaxation techniques in [108] with subgradient optimization and column generation to create two new heuristics. Computation results are presented, and the authors note that the heuristic using column generation is best when used on large-scale instances. The subgradient heuristic performed better on small scale instances.

[111] E. L. F. Senne, L. A. N. Lorena, and M. A. Pereira. A branch-and-price approach to p-median location problems. *Computers and Operations Research*, 32(6):1655–1664, 2005.

Combines the traditional column generation approach with Lagrangian/surrogate relaxation. The method is a tree search algorithm employing column generation at each search node. The authors claim that the Lagrangian/surrogate multiplier modifies the reduced cost criterion so that more productive columns are selected at each search node than in the traditional column generation method. The authors also claim that the algorithm is faster than the traditional approach.

[112] É. D. Taillard. Heuristic methods for large centroid clustering problems. Technical Report 96-96, Istituto Dalle Molle Di Studi Sull Intelligenza Artificiale, 1996.

Applies three new clustering methods to solving the p-median problem—*candidate list search (CLS)*, *local optimization (LOPT)* and *decomposition/recombination (DEC)*. CLS begins with the alternate heuristic [77], obtaining a locally optimal solution. This solution is perturbed by eliminating a vertex and adding another, similar to vertex substitution [114]. The new solution is chosen only if it is better than the initial one. LOPT selects a median and some nearby medians and generates and solves a corresponding subproblem. DEC uses LOPT to obtain a good solution for the overall problem.

[113] A. Tamir. An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs. *Operations Research Letters*, 19(2):59–64, 1996.

This paper shows that the total running time of the "leaves to root" dynamic programming algorithm is $O(pn^2)$.

[114] M. B. Teitz and P. Bart. Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research*, 16(5):955–961, 1968.

Introduces a *vertex substitution* or *interchange* heuristic for finding the p-median. This method begins by selecting an initial vertex subset $V_p = \{m_j : 1 \leq j \leq p\}$. For every $v_i \notin V_p$, the heuristic finds the location $m_j \in V_p$, if it exists, that would improve the solution the most if m_j was replaced with v_i. If this location exists, then the vertices are switched, and the process is repeated on the new solution. When all vertices have been checked, the algorithm terminates with a local minimum solution.

[115] M. Thorup. Quick k-median, k-center, and facility location for sparse graphs. In *ICALP '01: Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 249–260, 2001.

Presents a $12 + o(1)$ constant factor approximation algorithm for the p-median problem.

[116] S. Voß. A reverse elimination approach for the p-median problem. *Studies in Locational Analysis*, 8:49–58, 1996.

Applies a tabu search procedure known as reverse elimination to the p-median problem. This procedure implements a necessary and sufficient condition so that known solutions are not revisited. This involves storing the entire search history in a running list. When used in conjunction with a diversification strategy, this was found to yield favorable results. Computational results are presented, comparing this method with the greedy interchange method in [119].

[117] R. A. Whitaker. A tight bound drop exchange algorithm for solving the p-median problem. *Environment and Planning A*, 13:669–680, 1981.

Details a drop algorithm that starts with all n nodes in the solution set and an objective value of 0. On each iteration, k nodes are removed from the solution set and $(k - 1)$ nodes are brought back into the set, yielding a net loss of one node per iteration. On each iteration, the procedure attempts to minimize the amount that the objective value increases. The algorithm terminates when p vertices remain in the solution set. Computational results are presented comparing this method to a greedy interchange method.

[118] R. A. Whitaker. Some interchange algorithms for median location problems. *Environment and Planning B*, 9:119–129, 1982.

A variation of vertex substitution [114] that involves exchanging multiple vertices at once is presented. Several ways of producing initial solutions for this method are outlined, including greedy heuristics and drop algorithms (similar to the one presented in [117]). Extensive computational results are presented.

[119] R. A. Whitaker. A fast algorithm for the greedy interchange of large-scale clustering and median location problems. *INFOR*, 21(2):95–108, 1983.

Presents a method that initializes vertex substitution [114] with the solution obtained with a fast greedy algorithm based upon the methods in [27] and [69]. This method found solutions to test problems more than an order of magnitude faster than normal vertex substitution initialized with a normal greedy technique.

[120] N. E. Young. Greedy approximation algorithms for k-medians by randomized rounding. Technical Report PCS-TR99-344, Department of Computer Science, Dartmouth College, Hanover, NH, 1999.

Applies randomized rounding and other probabilistic methods to understanding the operation of greedy approximation algorithms. An approximation algorithm is presented that, for any $\epsilon > 0$, finds a solution using at most $\ln(n + (n/\epsilon))p$ medians and with objective value no more than $(1 + \epsilon)$ times the optimal solution. This algorithm requires $\ln(n + (n/\epsilon))p$ linear time iterations.

Appendix A

1963	Heuristic [69]
1964	Heuristic [77]
	Other [53]
1965	Direct Enumeration [54]
	IP Formulation [10]
1966	IP Formulation, LP Relaxation [38]
1968	Vertex Substitution [114]
1970	IP Formulation, LP Relaxation [94]
1972	Heuristic [64]
	LP Relaxation [67]
	Enumeration [78]
1973	Heuristic [40]
	Vertex Substitution [105]
1974	Heuristic [9] [66]
	LP Relaxation [51]
1975	Graph Theoretic [24]
	LP Relaxation [87]
1977	LP Relaxation [86]
	Heuristic, LP Relaxation [27]
1978	LP Relaxation [41]
	Vertex Substitution [39]
1979	IP Formulation [103]
	Vertex Substitution [98] [99]
	Complexity, Graph Theoretic [65]
	Survey [55]
	Complexity [50]
1980	LP Relaxation [43]
1981	LP Relaxation [45]
	Graph Theoretic [44]
	Heuristic [117]
1982	LP Relaxation [25]
	Vertex Substitution [118]
1983	Vertex Substitution [119]
	Graph Theoretic [85]
1984	Graph Theoretic [46]
1985	LP Relaxation [11] [56] [83]
1986	Genetic Algorithm [60]
1988	Heuristic [2]
1989	LP Relaxation [47]
1990	Survey [82]
	Tabu Search [52]
1991	Vertex Substitution [33] [90]
	Heuristic [15]
1992	Approximation Algorithm [72]
	Vertex Substitution [34] [35]
1993	LP Relaxation [12]
1994	Heuristic [89]
	Genetic Algorithm [84]
1995	Survey [31]
	Simulated Annealing [95]
1996	Graph Theoretic [113]
	Vertex Substitution [59]
	Tabu Search [96] [116]
	Heuristic [112]
1997	Heuristic [61] [106]
	Variable Neighborhood Search [57]
	Heuristic Concentration [101]
	Vertex Substitution [97]
1998	Genetic Algorithm [76]
	Approximation Algorithm [18] [21]
	Heuristic Concentration, Tabu Search [102]
1999	Heuristic Concentration [104]
	Other [88]
	Genetic Algorithm [37] [42]
	IP Formulation [8]
	Approximation Algorithm [19] [20] [120]
2000	Simulated Annealing [23]
	Surrogate Relaxation [108]
	Approximation Algorithm [68]
2001	Approximation Algorithm [63] [115]
	Surrogate Relaxation [109]
	Variable Neighborhood Search [58]
	IP Formulations [5]
	Genetic Algorithm [22] [73]
	LP Relaxation [32]
2002	Approximation Algorithm [62] [81]
	Heuristic Concentration [100]
	Variable Neighborhood Search [48]
	Genetic Algorithm [14]
	Surrogate Relaxation [110]
	Tabu Search [107]
	Neural Network [79]
	LP Relaxation [36]
2003	Variable Neighborhood Search [29]
	Vertex Substitution [91]
	Neural Network [80]
	Scatter Search [49]
	Genetic Algorithm [1] [71]
	Surrogate Relaxation [74]
	LP Relaxation [6] [7] [16]
	IP Formulation [26]
2004	Simulated Annealing [70]
	Genetic Algorithm [28]
	Approximation Algorithm [3]
	Surrogate Relaxation [75]
	Vertex Substitution [92]
	LP Relaxation [13]
	Variable Neighborhood Search [30]
	GRASP Metaheuristic [93]
2005	Surrogate Relaxation [111]
	LP Relaxation [17]
	Vertex Substitution [4]