What You Should Know About Location Modeling

Mark S. Daskin

Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60208

Received 29 January 2008; accepted 29 January 2008
DOI 10.1002/nav.20284
Published online 28 March 2008 in Wiley InterScience (www.interscience.wiley.com).

Abstract: Facility location models have been applied to problems in the public and private sectors for years. In this article, the author first presents a taxonomy of location problems based on the underlying space in which the problem is embedded. The article illustrates problems from each part of the taxonomy with an emphasis on discrete location problems. Selected recent research in the area is also discussed. © 2008 Wiley Periodicals, Inc. Naval Research Logistics 55: 283–294, 2008

Keywords: location modeling; location problems; facility location

1. INTRODUCTION

Facility location problems have proven to be a fertile ground for operations researchers interested in modeling, algorithm development, and complexity theory. Applications of location modeling include locating emergency medical service (EMS) bases, fire stations, schools, hospitals, reserves for endangered species, airline hubs, waste disposal sites, and warehouses to list only a small subset of the numerous areas in which location models have been applied. Location models have also found applications in nontraditional areas, including medical diagnosis, vehicle routing, alignment of candidates and parties along a political spectrum, and the analysis of archeological sites [7].

Location theory and modeling has its roots in the pioneering work of Weber [65] who considered the problem of locating a single facility to minimize the total travel distance between the site and a set of customers. Later, Hotelling [29] studied the location of two facilities on a line. In his simple model, customers patronize the closer of the two facilities and the vendors locate to maximize their market share. With customers uniformly distributed along the line, the optimal location for both vendors is in the middle of the line with each vendor capturing exactly half of the market. Isard [30–32] is viewed as the founder of regional science, a merger of economics and location theory.

In this article, I begin with a taxonomy of modern location models and provide examples of three of the four major areas. The remainder of the article is devoted to the fourth area, discrete location modeling. I review the formulation of five foundational models, briefly discuss solution algorithms for these models and then summarize areas of ongoing and future research.

Correspondence to: M.S. Daskin (m-daskin@northwestern. edu)

2. A TAXONOMY OF LOCATION MODELS

While there are numerous ways of subdividing the broad spectrum of location models, Fig. 1 illustrates a breakdown based on the space in which the problems are modeled.

Analytic models are the simplest of location models. Such models typically assume that demand is distributed in some way (e.g., uniformly) over a service area and that facilities can be located anywhere within the area. Analytic models are typically solved using calculus or other simple techniques. While the strong assumptions required to develop such models limit their applicability in particular instances, the insights derived from such models tend to be applicable in a range of contexts.

To illustrate an analytic location model, assume that demands are uniformly distributed in a square area, a. Travel occurs along roads oriented at 45 degrees to the sides of the square. Under these assumptions, it is easy to show that the expected distance between a facility at the center of the service area and a randomly selected demand is $\frac{2}{3}\sqrt{\frac{a}{2}}$. Now consider the problem of locating N facilities to minimize the sum of the facility location cost and the expected transport cost. Let f be the fixed cost of locating a facility, ρ the demand density, and c the transport cost per demand per unit distance. With N facilities, each facility will serve an area of (approximately) a/N and the expected distance to the nearest of the N facilities from a randomly selected demand point

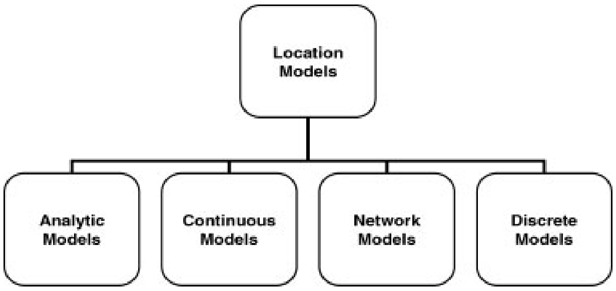

Figure 1. Taxonomy of location models.

will be $\frac{2}{3}\sqrt{\frac{a}{2N}}$. The expected facility and transport cost as a function of the number of facilities located is

$$TC(N) = fN + \left(\frac{2}{3}\sqrt{\frac{a}{2N}}\right)c\rho a. \tag{1}$$

Ignoring for the moment that N must be integer, we take the derivative of this cost function with respect to N, the number of facilities, and find

$$N^* = a\left(\frac{c\rho\sqrt{2}}{6f}\right)^{2/3}. \tag{2}$$

Thus, the optimal number of facilities grows as the $2/3$ power of the demand density and the cost per item per unit distance, and decreases with the $2/3$ power of the facility costs. As expected the number of facilities increases linearly with the area served. Substituting the optimal number (2) into the total cost function (1), we obtain

$$TC(N^*) = af^{1/3}(c\rho)^{2/3}\left[\left(\frac{\sqrt{2}}{3}\right)^{2/3} + 2\left(\frac{\sqrt{2}}{3}\right)^{2/3}\right]$$
$$\cong 1.1447af^{1/3}(c\rho)^{2/3} \tag{3}$$

where the first term in the braces of (3) comes from the fixed facility costs and the second comes from the transport costs. Thus, at the optimal solution, the facility costs are roughly 50% of the transport costs. The total cost increases linearly with the area served, with the $1/3$ power of the facility costs and the $2/3$ power of the demand density and unit transport cost.

The optimal number of facilities as given by (2) may be fractional since we ignored the condition that N must be integral when we took the derivative of (1). Thus, we should be interested in the error introduced by rounding the fractional value of N to the nearest integer. Also, for a variety of reasons exogenous to this simple model, the number of sites used may differ significantly from the optimal number. If, instead of using N^* facilities, we use $N = \beta N^*$ facilities, the ratio of the suboptimal cost to the optimal cost is given by

$$\frac{TC(N)}{TC(N^*)} = \frac{\beta + 2/\sqrt{\beta}}{3} \tag{4}$$

For $0.63 \le \beta \le 1.53$, the cost is within 5% of the optimal cost and for $0.52 \le \beta \le 1.82$ the cost is within 10% of the optimal value. In short, in this simple model, the total cost is relatively insensitive to changes in the number of facilities deployed. Daganzo [12] uses models similar to this to analyze more complex distribution systems.

While analytical models assume that demands are distributed continuously across a service region and that facilities can be located anywhere within the region, *continuous models* typically assume that demands arise only at discrete points (see Plastria [46] for an introduction to continuous location modeling). The classical Weber [65] problem is typical of this class. Demands occur at each of n discrete points. The location of demand point i is given by (x_i, y_i) for $i = 1, 2, \ldots, n$ and the intensity of demand at this location is given by h_i. The Weber problem seeks the location (X, Y) of a single facility to minimize the demand-weighted total distance between the facility and the demand points. In other words, we want to

$$\underset{X,Y}{\text{Minimize}} \quad \sum_{i=1}^{n} h_i\sqrt{(x_i - X)^2 + (y_i - Y)^2} \tag{5}$$

This model is solved using iterative numerical procedures such as the Weiszfeld [66] algorithm. Drezner et al. [18] review this model, the Weiszfeld algorithm, and a variety of enhancements to the algorithm to ensure and expedite its convergence. They also outline extensions to this basic model.

Network models assume that demands arise, and facilities can be located, only on a network composed of nodes and links. Often demands occur only on the nodes, while facilities can be located anywhere on the network. The focus of much of the network location literature is on finding polynomial time algorithms, often for problems on specially structured networks such as trees.

To illustrate a network location problem, consider the following problem on a tree. Demands arise only on the nodes of the tree and the demand at node i is again given by h_i. A single facility can be located anywhere on the tree. The objective is to minimize the demand-weighted total distance between the facility and the nodes. This problem is called the *1-median problem on a tree*. Goldman [25] showed that the problem can be solved in $O(n)$ time, where n is the number of nodes in the tree: Begin with any tip node of the tree. If the demand at that node is equal to or greater than half the total demand of all nodes, the optimal location is at that node. If not, remove that node (and the link connecting it to the remainder of the tree) from the tree and add the node's demand to the demand of the node to which it had been linked. The procedure continues until the revised demand of a node is half or more of the total demand of all nodes in the tree. Since each node is examined at most once, and examination of each node requires only a comparison and an addition, the algorithm runs in $O(n)$ time. The problem of locating p facilities on a tree to minimize the demand-weighted total distance can also be solved in polynomial time [58]. Linear time algorithms also exist for the unweighted ($h_i = 1$, for all i) problem of locating one or two facilities to minimize the maximum distance between any node and the nearer facility. Numerous other network location models have been studied in the literature. A full review of this work is beyond the scope of this paper; readers should consult Tansel et al. [60] for a (somewhat dated) review of this literature.

3. DISCRETE LOCATION MODELS

The fourth and final branch of the location model taxonomy deals with *discrete models*. In such models, there may or may not be an underlying distance metric. Distances or costs between any pair of nodes may be arbitrary, although they generally do follow some rule (e.g., Euclidean, Manhattan, network, or great circle distances). Demands generally arise on the nodes and the facilities are restricted to a finite set of candidate locations.

Figure 2 further classifies discrete location models by subdividing the class into three broad areas. Covering based models assume that there is some critical coverage distance or time within which demands need to be served if they are to be counted as "covered" or "served adequately." Such models are typically used in designing emergency services as there are both practical and, in many jurisdictions, legislative guidelines for coverage. Note that coverage and service are not identical. For example, in locating fire stations, a node may not be covered (e.g., it may be more than 10 minutes from the nearest station), but demands at that location would still be served if they were within the service region. Increasingly, covering models are also being used in the private sector as coverage can be used as a proxy for high-quality service (e.g., "when it positively absolutely has to be there by. . .."). Within the class of covering models, three prototypical models are the set covering model, the maximal covering model and the p-center model.

Median-based models minimize the demand-weighted average distance between a demand node and the facility to which it is assigned. Such models are typically used in distribution planning contexts in which minimizing the total outbound or inbound transport cost is essential. Two models are shown in Fig. 2: the p-median model and the uncapacitated fixed charge location problem.

Finally, there are models which do not fall into either of these categories. For example the p-dispersion model [36] maximizes the minimum distance between any pair of facilities. This model is useful in locating franchise outlets, where minimizing the cannibalization of one outlet's market by another franchisee is desirable. The model can also be used in

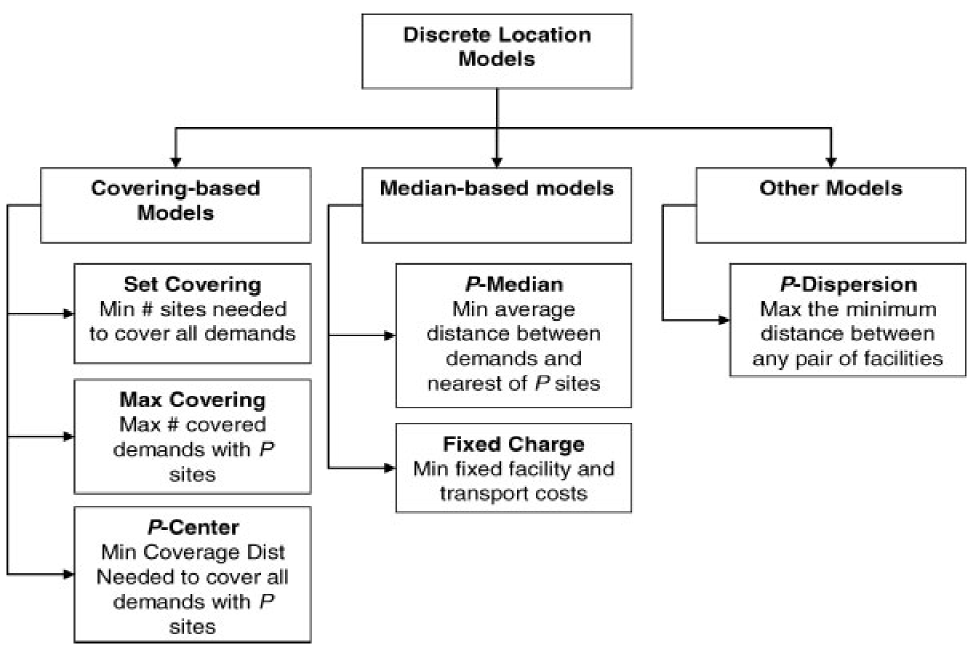

Figure 2. Breakdown of discrete location models.

locating weapon supplies (e.g., nuclear weapons) where minimizing the likelihood that the destruction of any one cache would impact other supplies is desirable. Within the category of "other" models are a large number of formulations aimed at locating undesirable facilities.

To illustrate the formulation of discrete location models, we begin with the set covering model. Let I be a set of demand nodes and J be a set of candidate locations. The distance between demand node $i \in I$ and candidate site $j \in J$ is d_{ij}. Demand node i is covered by candidate site j if $d_{ij} \leq D^c$ where D^c is the coverage distance. Define $N_i = \{j \in J : d_{ij} \leq D^c\}$. In other words, N_i is the set of all candidate sites which can cover demand node i. Finally, define a binary decision variable X_j to be 1 if we locate at candidate site j and 0 otherwise. The location set covering model can now be formulated as follows [62]:

$$\text{Min} \quad \sum_{j \in J} X_j \tag{6}$$

$$s.t. \quad \sum_{j \in N_i} X_j \geq 1 \quad \forall i \in I \tag{7}$$

$$X_j \in \{0, 1\} \quad \forall j \in J \tag{8}$$

The objective function (6) minimizes the number of facilities needed to cover all demands. Constraint (7) stipulates that each demand node must be covered. Constraints (8) are integrality constraints. While the set covering problem is NP-hard, in practice the linear programming relaxation is likely to have an all-integer solution. Even when the LP relaxation is fractional, only a small number of nodes must typically be explored in a branch and bound tree. As such, large instances of the problem can often be solved in reasonable time. Despite these observations, it is easy to construct instances that result in fractional solutions. For example, consider a problem with nodes at the four vertices of a square with sides of unit length. For any coverage distance greater than or equal to 1 and less than 2, the LP solution is to locate 1/3 of a facility at each node.

Restricting the set of candidate sites to the nodes only is likely to be subobtimal compared to allowing locations on the nodes and links of a network. However, Church and Meadows [8] present a method to augment the nodes on a network with a finite set of additional points on the links such that the optimal solution to formulation (6)–(8) using the augmented candidate set, J, will result in a solution value equal to that obtainable by locating anywhere on the links or nodes.

The location set covering problem has a number of weaknesses. First, it is often prohibitively expensive to locate the number of facilities needed to cover all demands. In those cases, it may be necessary to either increase the coverage distance or relax the requirement of total coverage. Second, there are often a large number of alternate optima to the set covering model. In the simple 4-node problem outlined above, locating at any two of the nodes would result in total coverage. Third, the model does not distinguish between large demand nodes and small demand nodes.

The maximal covering model [9] locates p facilities to maximize the number of covered demands. This model differentiates between big and small demands and allows some nodes to be uncovered if the number of sites needed to cover all nodes exceeds p. In addition to the notation defined above, define a new decision variable, Z_i, which equals 1 if demand node i is covered and 0 otherwise. With this additional notation, the maximal covering model can be formulated as follows:

$$\text{Max} \quad \sum_{i \in I} h_i Z_i \tag{9}$$

$$s.t. \quad \sum_{j \in J} X_j = p \tag{10}$$

$$Z_i - \sum_{j \in N_i} X_j \leq 0 \quad \forall i \in I \tag{11}$$

$$X_j \in \{0, 1\} \quad \forall j \in J \tag{12}$$

$$Z_i \in \{0, 1\} \quad \forall i \in I \tag{13}$$

The objective function maximizes the number of covered demands. Constraint (10) states that p facilities are to be located. Constraints (11) link the location and coverage variables, while constraints (12) and (13) are integrality constraints. Although the maximal covering model is also NP-hard, it can be solved using a variety of heuristics including the greedy adding heuristic and the greedy adding and substitution algorithm. As in the case of the set covering model, the maximal covering model can often be solved using conventional mixed integer programming packages as the linear programming relaxation is often integer. Also, relaxing constraint (11) and embedding Lagrangian relaxation [21,22] in branch and bound works quite effectively [15,24].

Figure 3 shows the percent of all demand covered as a function of the number of facilities located for the 500 largest counties in the contiguous United States. County populations are the demand values. While these 500 counties represent less than 1/6 of the 3109 counties in the contiguous states, they account for nearly 75% of the total population. A total of 14 facilities are needed to cover all 500 demand nodes with a coverage distance of 300 miles. Figure 4 shows this solution. However, with only 7 facilities, 88.3% of the demand is covered. Figure 5 shows this solution. This result is typical of most real-world datasets. *Using only half the number of sites needed to cover all demands results in 80% to 90% of the total demand being covered.* Two other observations are worth making. First, Figs. 4 and 5 clearly illustrate that solutions with fewer facilities are not necessarily subsets of solutions with more sites. Thus, greedy algorithms will clearly be suboptimal in general. Second, using a greedy algorithm until all

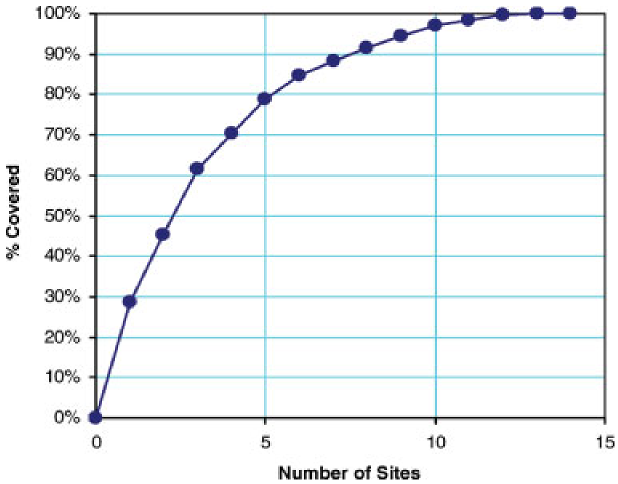

Figure 3. Percent of demand covered versus number of sites for 500 US counties. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

demands are covered results in a solution requiring 19 sites to cover all demands, 36% more sites than are truly needed. The greedy adding and substitution algorithm does somewhat better, requiring 16 sites to cover all demands or 14% more sites than are truly needed.

The p-center model finds the smallest possible coverage distance such that every node is covered. On a network, the *absolute* p-center model allows facilities to be located on the nodes and the links, while the *vertex* p-center model restricts sites to the nodes. Tansel et al. [59] provide an excellent survey of early work on the median and center problems.

Covering models generally treat distances as binary: either a node is covered or it is not. Median-based models account for the actual distances. The p-median model [26,27] locates p facilities to minimize the demand-weighted total (or average) distance between demands and the nearest facility. For this model, define an assignment variable, Y_{ij}, which equals 1 if demand node i is assigned to a facility at candidate site j, and 0 otherwise. With this notation, the p-median model can be formulated as follows:

$$\text{Min} \quad \sum_{j \in J} \sum_{i \in I} h_i d_{ij} Y_{ij} \tag{14}$$

$$s.t. \quad \sum_{j \in J} Y_{ij} = 1 \quad \forall i \in I \tag{15}$$

$$Y_{ij} - X_j \leq 0 \quad \forall i \in I; \ \forall j \in J \tag{16}$$

$$\sum_{j \in J} X_j = p \tag{17}$$

$$X_j \in \{0, 1\} \quad \forall j \in J \tag{18}$$

$$Y_{ij} \in \{0, 1\} \quad \forall i \in I; \ \forall j \in J \tag{19}$$

The objective function (14) minimizes the demand-weighted total distance. Constraints (15) stipulate that each node is assigned, while constraints (16) limit assignments to open or selected sites. Constraint (17) states that p facilities are to be located. Finally, constraints (18) and (19) are integrality constraints. Constraints (19) can be relaxed to (20) since each

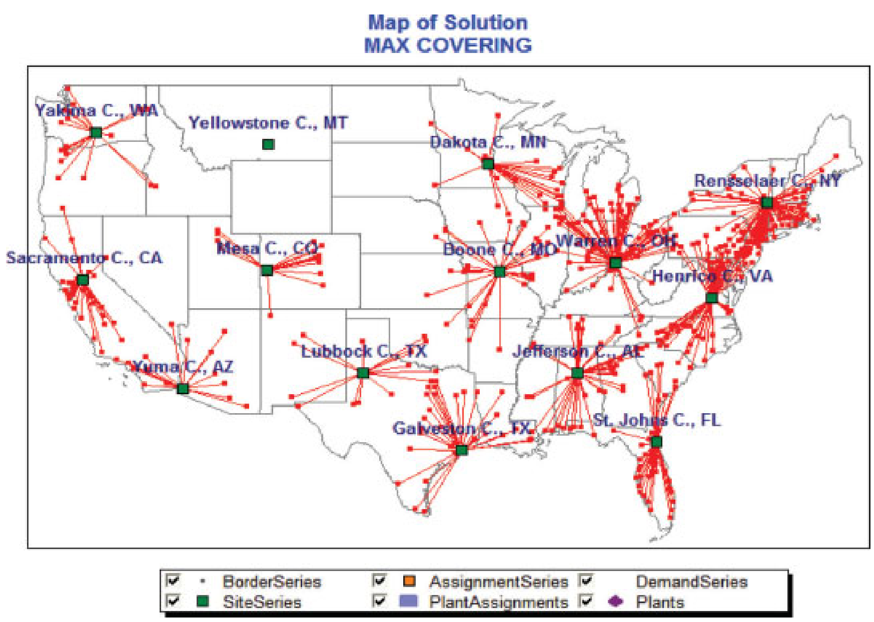

Figure 4. Fourteen sites covering all demands in 300 miles. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

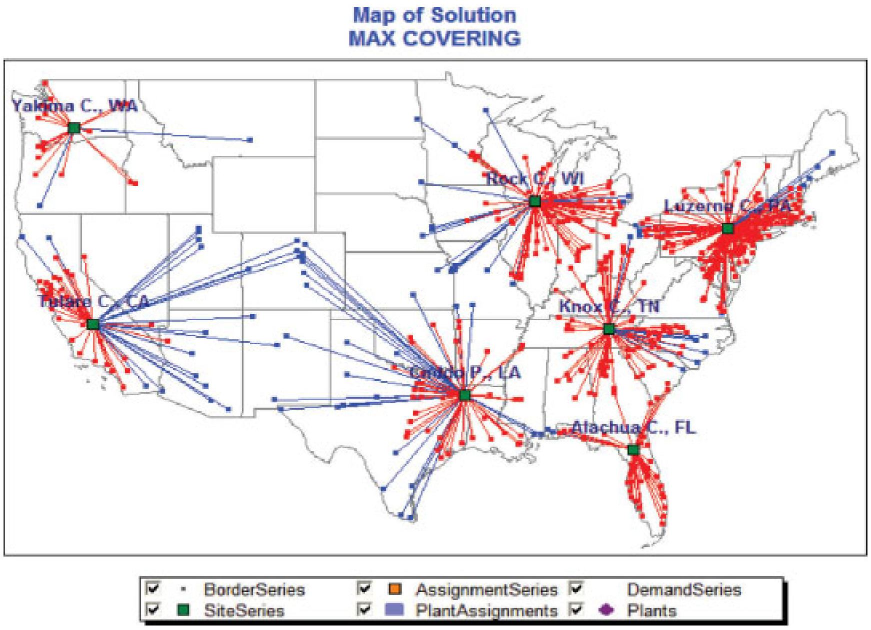

Figure 5. Seven sites covering over 88 percent of the total demand. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

demand node will automatically be assigned to the closest open site in any feasible solution.

$$0 \leq Y_{ij} \leq 1 \quad \forall i \in I; \;\; \forall j \in J \tag{20}$$

Hakimi [27] showed that at least one optimal solution to the p-median problem consists of locating only on a subset of the demand nodes. This result has been extended to problems in which d_{ij} is replaced by a concave function of the distance.

Kariv and Hakimi [35] showed that the p-median problem is NP-hard. A number of algorithms have been developed to solve the p-median model both heuristically and optimally. The greedy adding algorithm adds facilities one at

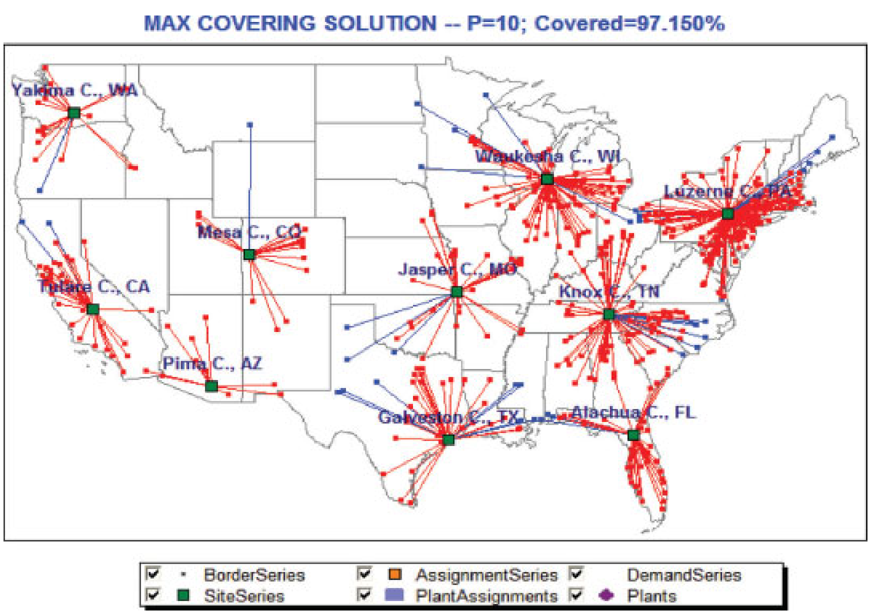

Figure 6. Maximal covering solution with 10 facilities, % covered = 97.15%. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

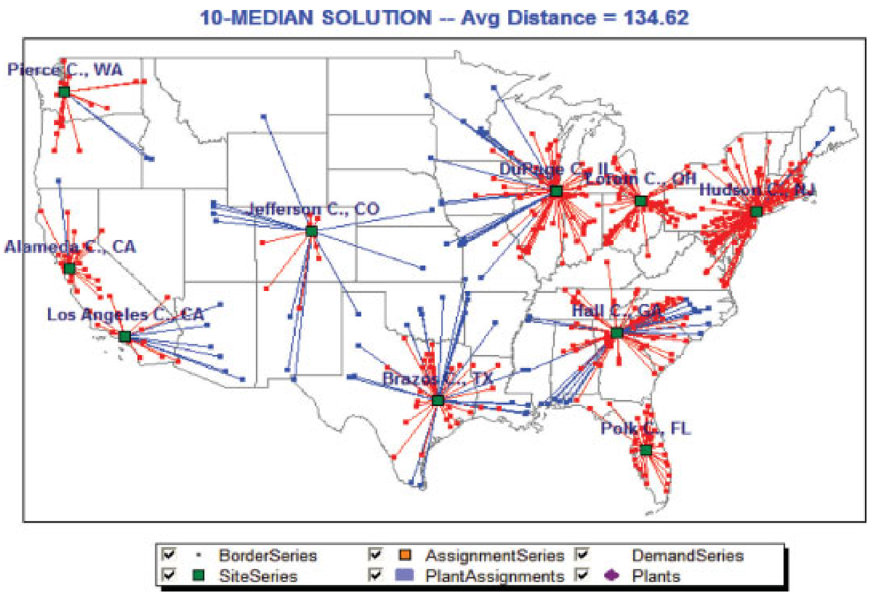

Figure 7. Ten-median solution, average distance = 134.62. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

a time so that the objective is minimized given the previously located sites. Maranzana [41] proposed a neighborhood search algorithm which is based on the observation that the 1-median can be found in polynomial time by total enumeration. The algorithm defines the neighborhood of a facility as the set of demand nodes assigned to the facility; it then finds the 1-median in each neighborhood. If any facility has changed, new neighborhoods are computed and the process continues. Teitz and Bart [61] introduced a more general exchange algorithm which tends to outperform Maranzana's algorithm. Rolland et al. [48] present a tabu search algorithm for the p-median problem. Rosing and ReVelle [49] introduced the notion of heuristic concentration and applied it successfully to the p-median problem. Rosing et al. [50]

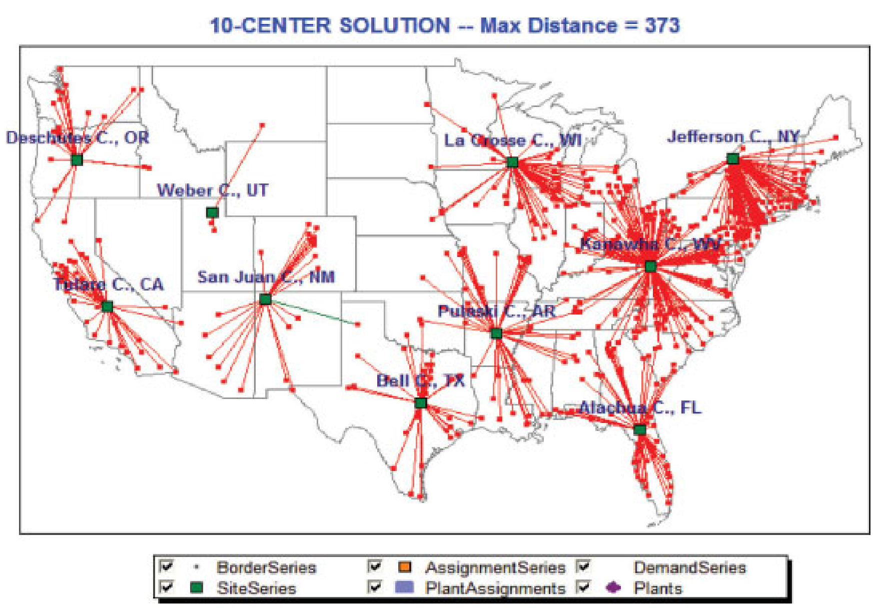

Figure 8. Ten-center solution, maximum distance = 373. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 1. Summary of covering center and median solutions.

	Solution		
	Max covering	p-center	p-median
Objective			
Max covering	**97.1**	86.8%	88.2%
p-center	507	**373**	565
p-median	176.5	219.2	**134.6**
UFLP	$2,839,650	$3,285,141	$2,403,191
Sites	52 - Pima C., AZ	159 - Tulare C., CA	1 - Los Angeles C., CA
	150 - Knox C., TN	162 - Pulaski C., AR	21 - Alameda C., CA
	159 - Tulare C., CA	242 - Bell C., TX	42 - DuPage C., IL
	163 - Waukesha C., WI	257 - Alachua C., FL	70 - Pierce C., WA
	180 - Luzerne C., PA	274 - Kanawha C., WV	87 - Hudson C., NJ
	229 - Galveston C., TX	280 - Weber C., UT	104 - Jefferson C., CO
	253 - Yakima C., WA	462 - Deschutes C., OR	118 - Polk C., FL
	257 - Alachua C., FL	469 - San Juan C., NM	196 - Lorain C., OH
	460 - Mesa C., CO	478 - Jefferson C., NY	348 - Brazos C., TX
	500 - Jasper C., MO	496 - La Crosse C., WI	389 - Hall C., GA
Average site number	240.3	327.9	137.6

compared heuristic concentration and tabu search. Bozkaya et al. [4] recently introduced a genetic algorithm that works well for the p-median problem. Relaxing constraints (15) results in an effective Lagrangian relaxation which can be embedded in branch and bound to find optimal solutions [15]. Hale and Moberg [28] review extensions to median, center and covering models.

The p-median problem ignores differences in the facility location costs at different sites. The Uncapacitated Fixed Charge Location Problem (UFLP) problem is a close cousin of the p-median problem. Letting f_j be the fixed facility cost of locating at candidate site $j \in J$, the UFLP minimizes (21) subject to all of the constraints of the p-median model with the exception of constraint (17)

$$\sum_{j \in J} f_j X_j + c \sum_{j \in J} \sum_{i \in I} h_i d_{ij} Y_{ij} \qquad (21)$$

where c is the transport cost per item per mile.

The UFLP is also NP-hard. It too can be solved effectively for fairly large instances using procedures similar to those developed for the p-median problem, including Lagrangian relaxation [23]. In addition, Erlenkotter [20] introduced a dual ascent algorithm for this problem. Van Roy and Erlenkotter [64] extend this algorithm to a dynamic location context in which new facilities can be added and existing sites can be closed.

Finally, we formulate the p-center problem for the case in which a finite set of candidate sites, J, is given. Define W as the maximum demand-weighted distance (the unweighted case is formulated identically with $h_i = 1$ for all i). The p-center problem is then simply to minimize W, subject to constraints (15) through (19) and the additional constraint (22) which defines W in terms of the assignment variables.

$$h_i \sum_{j \in J} d_{ij} Y_{ij} - W \leq 0 \quad \forall i \in I \qquad (22)$$

The p-center problem is also NP-hard [34].

To illustrate the differences between some of these models, consider again the 500 largest counties in the contiguous United States. Figures 6–8 show the solutions to the maximal covering, median and unweighted center problems for 10 sites. Again, a coverage distance of 300 miles was used. Table 1 summarizes the key differences between these solutions, lists the sites used in each of the 3 solutions, and also shows the objective function value for the UFLP with $f_j = 100,000$, $\forall j \in J$ and $\alpha = 0.00005$. Site numbers beside the county names in Table 1 are the rank of the counties in decreasing order of population (with Los Angeles, number 1, being the most populous).

The unweighted p-center problem ignores the demand levels. As such it tends to locate in less populous, but more central, areas than do the other two models. The average county number for the p-center solution is 328. The maximum covering model accounts for population, but deals with distance in a binary (covered or uncovered) manner. As shown in Table 1, it tends to locate facilities in more populous locations than does the p-center model. The p-median model accounts for both population and the actual distances, and it tends to locate in more populous sites.

Significant differences in the objective function values can also be seen. While 97.15% of the total demand can be covered with 10 facilities within 300 miles, the coverage distance has to increase to 373 (almost a 25% increase) for all demands to be covered by only 10 sites. This solution, however, has

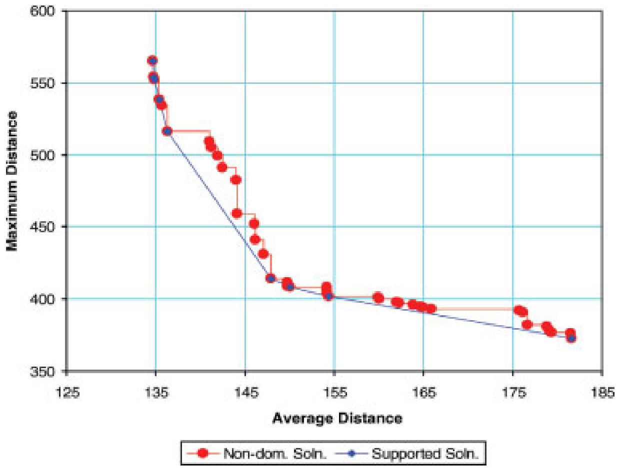

Figure 9. Tradeoff between average and maximum distance with 10 facilities serving the 500 largest counties in the contiguous United States. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

an average distance of 219 which is 63% greater than the minimum possible distance of 135 miles shown in Fig. 7.

There is a clear tradeoff between the average and maximum distance objectives. Figure 9 plots the 38 nondominated solutions on this tradeoff curve. The 9 supported solutions that define the lower convex hull of the nondominated solutions are shown with diamonds. This figure illustrates another common observation in discrete location modeling. *The optimal solution for any one objective is likely to do poorly with respect to other objectives, but there are very often many good compromise solutions*. For example, the fourth supported solution from the right in the figure (147.95, 414) reduces the maximum distance by nearly 27% (compared with the p-median solution shown at the far left of the figure) while increasing the average distance less than 10%. The next supported solution to the left (136.34, 516) reduces the maximum distance by 8.67% while increasing the average distance only 1.3% compared with the p-median solution.

Finally, we can compare the qualitative recommendations of the fixed charge analytic model [Eqs. (1)–(4)] with the results of a discrete fixed charge location model. Table 2 shows the results of 21 runs (9 with equal fixed costs per site and 12 with costs that varied linearly with the population) for the fixed charge location model applied to the 3109 counties of the contiguous United States. The cost per item per mile shipped [c in (21)] was varied as shown in the second row of the table (note that c also includes an embedded conversion factor from population to demand). The 9 problems with equal fixed costs were solved to optimality while the maximum gap between the lower and upper bounds for the 12 unequal fixed cost runs was 0.036 percent. Solution times for the equal fixed cost runs were all under 40 min while the maximum time for the unequal fixed cost runs was over 13 h. For strategic decision making problems of this size (3109 integer location variables and nearly 10 million assignment variables), these times are quite reasonable.

The last two rows of the table show the number of sites and total cost as a function of the cost per item per mile [c in (1)–(3) and (21)]. The analytic model results suggest that both the number of sites [Eq. (2)] and the total cost [Eq. (3)] should increase as the 2/3 power of c. The exponent of c in the four regression equations ranges from about 0.5 to 0.7. Thus, the regression equations—all with very high R^2 values—show that the number of sites and the total cost increase as a function of the cost per mile in a manner similar to that predicted by the analytic model, despite significant violations of the assumptions underlying the analytic model. For example, the contiguous United States is not a convex region, great circle distances were used in the discrete model as opposed to Manhattan, or right angle, distances used in the analytic model, and the demand density is anything but uniform. In fact, the population or demand density varies by 5 orders of magnitude from a high of over 9.5 million in Los Angeles County, CA to a low of 67 in Loving County, TX.

4. WHERE TO FROM HERE?

Location modeling and theory remain active areas of research. At the two most recent INFORMS Annual Meetings

Table 2. Number of sites and total cost as a function of the cost per item per mile.

	Equal fixed costs	Unequal fixed costs
Number of runs	9	12
Min and Max cost per item/mile, c	0.000001 and 0.01	0.000000025 and 0.01
Min and Max number of sites	1 and 320	1 and 560
Average time (s)	1,317	5,280
Max time (s)	2,202	48,454
Max % gap	0.000	0.036
Number of sites	No sites $= 5442.4c^{0.614}$ $R^2 = 0.9987$	No sites $= 4907.5c^{0.4945}$ $R^2 = 0.9962$
Total cost	Total cost $= 1.044 \cdot 10^9 c^{0.5871}$ $R^2 = 0.9999$	Total cost $= 2.126 \cdot 10^9 c^{0.7003}$ $R^2 = 0.9985$

in Pittsburgh and Seattle, respectively, there were 22 sessions with 88 papers dealing explicitly with location problems as well as 1 tutorial. Numerous other talks—on such topics as homeland security, healthcare, and telecommunications—also dealt with location problems. Clearly a review of all of this is beyond the scope of this paper. Rather, I will focus on a *few* emerging areas of research in discrete location modeling, with apologies to researchers in other areas.

One active area of research has involved relaxing key assumptions in the basic location models outlined earlier. For example, Berman et al. [2] consider a gradual covering model in which a demand is fully covered if the nearest facility is within l, uncovered if the nearest facility is further than u and partially covered if the nearest facility is between l and u distance units away. They show that, for convex decay functions, the maximal gradual covering problem can be restructured as a special case of the UFLP.

The incorporation of stochastic elements into location models has been of interest for years. Larson [38, 39] introduced the hypercube queueing model which addresses individual server availability in the face of Poisson demand and exponential service times. Jarvis [33] embedded this model in an iterative heuristic search algorithm to identify good facility locations. Daskin [13, 14] introduced the maximum expected covering location model which extends the maximal covering model to accommodate facilities which are busy or unavailable for part of the time. In the maximum availability location model developed by ReVelle and Hogan [47], demand nodes are covered only if the probability of a facility within the coverage distance being available exceeds a given value. Marianov and Serra [42] explicitly embed a queueing model into the maximal covering model. Berman and Krass [1] provide a review of stochastic location models with congestion effects. Berman et al. [3] merge partial coverage and queueing in set covering context.

In the post-9/11 era, facility reliability has become an increasingly important topic. Three broad categories of models have emerged. In the first, facilities fail randomly and the objective typically deals with expected performance. Such models are most appropriate for facility failures that are uncorrelated. For example, Snyder and Daskin [54] extend the p-median and UFLP models to account for facility reliability and explore the tradeoff between the traditional objective functions for these models and the costs when facilities fail. Lim et al. [40] extend the analytic (1) and mixed integer formulations (21) of the UFLP to account for two types of facilities: those that fail with some probability q and more costly facilities that do not fail at all. A carefully constructed Lagrangian relaxation algorithm was able to solve a problem with 3109 nodes (representing the counties of the contiguous United States) in just over an hour. The gap between the upper and lower bounds was 0.005% of the lower bound for this instance with 29 million variables and 38 million constraints.

Defender/interdictor models, in which disruptions are due to enemy attacks and the objective generally deals with minimizing the worst-case results, represent the second broad area of facility reliability modeling. Church et al. [11] extend the p-median and maximal covering models to identify the r most critical sites, the sites whose elimination results in the greatest degradation of the original objective function. Church and Scaparra [10] and Scaparra and Church [52] extend the p-median model to a case in which a defender can protect Q of the p facilities against attacks on R of the facilities by an enemy. The objective is to find the Q facilities to defend to minimize the worst-case attack. Snyder et al. [57] and Snyder and Daskin [55] review these streams of research.

Scenario planning represents a compromise between the independence assumption that is often made in dealing with random disruptions and the single worst-case assumptions made in defender/interdictor modeling. A variety of objectives ranging from expected performance to worst-case performance and minimax regret have been employed in location modeling as well as objectives that represent a compromise between expected and worst-case objectives (e.g., Daskin et al. [17]; Chen et al., [6]; Snyder and Daskin, [56]).

Facility location problems typically represent long-term strategic planning issues in supply chain design. As such these decisions impact, and are affected by, lower level tactical and operational decisions. Nozick and Turnquist [44, 45] iteratively solve location and inventory models to integrate these two decisions. Daskin et al. [16] and Shen et al. [53] present a non-linear extension of the fixed charge location model (21) that incorporates inventory decisions at a distribution center as well as shipments from a plant to the distribution centers. Vehicle routing problems have also been integrated into facility location models (see Min et al. [43] for a review of this work). It is worth noting that a capacitated fixed charge location model is at the heart of one of the more effective heuristics for the vehicle routing problem (Bramel and Simchi-Levi, [5]). Finally, researchers interested in the field of reverse logistics or closed-loop supply chains are utilizing extensions of the UFLP to identify desirable sites for forward distribution centers and reverse collection, sorting, recycling and remanufacturing facilities (Easwaran and Uster, [19]; Sahyouni et al. [51]; and Uster et al. [63]).

5. CONCLUSIONS

Facility location problems have been a fertile ground for the development of new modeling techniques, innovative solution algorithms and exciting applications. This paper presented a taxonomy of location models based on the underlying space in which the problem is embedded. After a brief review of three of the four areas, the paper outlined the formulation of five key discrete location models: the set covering model, the maximal covering model, the p-median model, the uncapacitated fixed charge location problem and the p-center problem. In contrast to the vehicle routing problem for which only relatively small instances can be solved to provably optimal solutions (Laporte, [37]), many location problems, though NP-hard, can be solved effectively for problems with thousands of demand nodes and candidate sites.

Recent extensions of the basic discrete location models were also highlighted with emphasis on models that account for stochasticity, reliability and uncertainty and the integration of other logistical decisions into location modeling frameworks. Location modeling remains an exciting area of research for operations researchers and of application for practitioners.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation Grant DMI-0457503. The opinions expressed in the article are those of the author and not of the National Science Foundation.

REFERENCES

[1] O. Berman and D. Krass, "Facility location problems with stochastic demand and congestion," Facility location: Applications and theory, Z. Drezner and H.W. Hamacher (Editors), Springer-Verlag, Berlin, 2002, pp. 329–371.
[2] O. Berman, D. Krass, and Z. Drezner, The gradual covering decay location problem on a network, Eur J Oper Res 151 (2003), 474–480.
[3] O. Berman, D. Krass, and J. Wang, Locating service facilities to reduce lost demand, IIE Trans 38 (2006), 933–946.
[4] B. Bozkaya, J. Zhang, and E. Erkut, "An efficient genetic algorithm for the p-median problem," Facility location: Applications and theory, Z. Drezner and H.W. Hamacher (Editors), Springer Verlag, Berlin, 2002, pp. 179–205.
[5] J. Bramel and D. Simchi-Levi, A location-based heuristic for general routing problems, Oper Res 43 (1995), 649–660.
[6] G. Chen, M.S. Daskin, Z-J. Shen, and S. Uryasev, The α-reliable mean-excess regret model for stochastic facility location modeling, Naval Res Logist 53 (2006), 617–626.
[7] R.L. Church and T.L. Bell, An analysis of ancient Egyptian settlement patterns using location-allocation covering models, Ann Assoc Am Geog 78 (1988), 701–714.
[8] R.L. Church and M. Meadows, Location modeling utilizing maximum service distance criteria, Geog Anal 11 (1979), 348–373.
[9] R.L. Church and C. ReVelle, The maximal covering location problem, Papers Reg Sci Assoc 32 (1974), 101–118.
[10] R.L. Church and M.P. Scaparra, Protecting critical assets: The r-interdiction median problem with fortification, Geog Anal 39 (2007), 129–146.
[11] R.L. Church, M.P. Scaparra, and R.S. Middleton, Identifying critical infrastructure: The median and covering facility interdiction problems, Ann Assoc Am Geog 94 (2004), 491–502.
[12] C.F. Daganzo, Logistics systems analysis, Springer, Berlin, 1991.
[13] M.S. Daskin, Application of an expected covering model to EMS system design, Decis Sci 13 (1982), 416–439.
[14] M.S. Daskin, A maximum expected covering location model: Formulation, properties and heuristic solution, Transport Sci 17 (1983), 48–70.
[15] M.S. Daskin, Network and discrete location: Models, algorithms and applications, Wiley, New York, 1995.
[16] M.S. Daskin, C. Coullard, and Z.-J.M. Shen, An inventory-location model: Formulation, solution algorithm and computational results, Ann Oper Res 110 (2002), 83–106.
[17] M.S. Daskin, S.M. Hesse, and C.S. ReVelle, α-Reliable P-Minimax regret: A new model for strategic facility location modeling, Location Sci 5 (1997), 227–246.
[18] Z. Drezner, K. Klamroth, A. Schöbel, and G.O. Wesolowsky, "The Weber problem," Facility location: Applications and theory, Z. Drezner and H.W. Hamacher (Editors), Springer, Heidelberg, Germany, 2001, pp. 1–36.
[19] G. Easwaran and H. Uster, Tabu search and benders decomposition approaches for a capacitated closed-loop supply chain network design problem, Working paper, Department of Industrial and Systems Engineering, Texas A&M University, 2007.
[20] D. Erlenkotter, A dual-based procedure for uncapacitated facility location, Oper Res 26 (1978), 992–1009.
[21] M.L. Fisher, The Lagrangian method for solving integer programming problems, Management Sci 27 (1981), 1–18.
[22] M.L. Fisher, A applications oriented guide to Lagrangian relaxation, Interfaces 15 (1985), 10–21.
[23] R.D. Galvão, The use of Lagrangean relaxation in the solution of uncapacitated facility location problems, Location Sci 1 (1993), 57–70.
[24] R.D. Galvão and C. ReVelle, A Lagrangean heuristic for the maximal covering location problem, Eur J Oper Res 88 (1996), 114–123.
[25] A.J. Goldman, Optimal center location on simple networks, Transport Sci 5 (1971), 212–221.
[26] S.L. Hakimi, Optimum locations of switching centers and the absolute centers and medians of a graph, Oper Res 12 (1964), 450–459.
[27] S. Hakimi, Optimum location of switching centers in a communications network and some related graph theoretic problems, Oper Res 13 (1965), 462–475.
[28] T.S. Hale and C.R. Moberg, Location science research: A review, Ann Oper Res 123 (2003), 21–35.
[29] H. Hotelling, Stability in competition, Econom J 39 (1929), 41–57.
[30] W. Isard, Location and space-economy: A general theory relating to industrial location, market areas, land use, trade and urban structure, M.I.T. Press, Cambridge, MA, 1956.
[31] W. Isard, Methods of regional analysis, M.I.T. Press, Cambridge, MA, 1960.

[32] W. Isard, General theory: Social political and regional, M.I.T. Press, Cambridge, MA, 1969.
[33] J.P. Jarvis, "Optimization in stochastic service systems with distinguishable servers," Report TR-19-75, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1975.
[34] O. Kariv and S.L. Hakimi, An algorithmic approach to network location problems. I: The p-centers, SIAM J Appl Math 37 (1979), 513–538.
[35] O. Kariv and S.L. Hakimi, An algorithmic approach to network location problems. II: The p-medians, SIAM J Appl Math 37 (1979), 539–560.
[36] M. Kuby, Programming models for facility dispersion: The p-dispersion and maxisum dispersion problems, Geog Anal 19 (1987), 315–329.
[37] G. Laporte, What you should know about the vehicle routing problem, Naval Res Logist 54 (2007), 811–819.
[38] R.C. Larson, A hypercube queueing model to facility location and redistricting in urban emergency services, Comput Oper Res 1 (1974), 67–95.
[39] R.C. Larson, Approximating the performance of urban emergency service systems, Oper Res 23 (1975), 845–868.
[40] M. Lim, M.S. Daskin, S. Chopra, and A. Bassamboo, Managing risks of facility disruptions, Working paper, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 2008.
[41] F.E. Maranzana, On the location of supply points to minimize transport costs, Oper Res Quart 15 (1964), 261–270.
[42] V. Marianov and D. Serra, Probabilistic, maximal covering location-allocation models for congested systems, J Reg Sci 38 (1998), 401–424.
[43] H. Min, V. Jayaraman, and R. Srivastava, Combined location-routing problems: A synthesis and future research directions, Eur J Oper Res 108 (1998), 1–15.
[44] L.K. Nozick and M.A. Turnquist, Integrating inventory impacts into a fixed-charge location model for locating distribution centers, Transport Res E: Log Transport Rev 34 (1998), 173–186.
[45] L.K. Nozick and M.A. Turnquist, A two-echelon inventory allocation and distribution center location analysis, Transport Res E 37 (2001), 425–441.
[46] F. Plastria, "Continous location problems," Facility location: A survey of applications and methods, Z. Drezner (Editor), Springer, New York, 1995.
[47] C. ReVelle and K. Hogan, The maximum availability location problem, Transport Sci 23 (1989), 192–200.
[48] E. Rolland, D.A. Schilling, and J.R. Current, An efficient tabu search procedure for the p-median problem, Eur J Oper Res 96 (1996), 329–342.
[49] K.E. Rosing and C.S. ReVelle, Heuristic concentration: Two stage solution construction, Eur J Oper Res 97 (1997), 75–86.
[50] K.E. Rosing, C.S. ReVelle, E. Rolland, D.A. Schilling, and J.R. Current, Heuristic concentration and tabu search, Euro J Oper Res 104 (1998), 93–99.
[51] K. Sahyouni, R.C. Savaskan, and M.S. Daskin, A facility location model for bidirectional flows, Transport Sci 41 (2007), 484–499.
[52] M.P. Scaparra and R.L. Church, A bilevel mixed-integerr program for critical infrastructure planning, Comput Oper Res 35 (2008), 1905–1923.
[53] Z.-J.M. Shen, C. Coullard, and M.S. Daskin, A joint location-inventory model, Transport Sci 37 (2003), pp. 40–55.
[54] L.V. Snyder and M.S. Daskin, Reliability models for facility location: The expected failure case, Transport Sci 39 (2005), 400–416.
[55] L.V. Snyder and M.S. Daskin, "Models for reliable supply chain network design," in Reliability and vulnerability in critical infrastructure: A quantitative geographic perspective, A. Murray and H.T. Grubesic (Editors), Advances in Spatial Science Series, Springer, 2006.
[56] L.V. Snyder and M.S. Daskin, Stochastic p-robust location problems, IIE Trans 38 (2006), 971–985.
[57] L.V. Snyder, M.P. Scaparra, M.S. Daskin, and R.L. Church, Planning for disruptions in supply chain networks, Tutorials Oper Res, INFORMS (2006), 234–257.
[58] A. Tamir, An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs, Oper Res Lett 19 (1996), 59–64.
[59] B.C. Tansel, R.L. Francis, and T.J. Lowe, Location on networks: A Survey. Part I: The p-center and p-median problems, Management Sci 29 (1983), 482–497.
[60] B.C. Tansel, R.L. Francis, and T.J. Lowe, Location on networks: A survey. Part II: Exploiting tree network structure, Management Sci 29 (1983), 498–511.
[61] M.B. Teitz and P. Bart, Heuristic methods for estimating the generalized vertex median of a weighted graph, Oper Res 16 (1968), 955–961.
[62] C. Toregas, R. Swain, C. ReVelle, and L. Bergman, The location of emergency service facilities, Oper Res 19 (1971), 1363–1373.
[63] H. Uster, G. Easwaran, E. Akcali, and S. Cetinkaya, Benders decomposition with alternative multiple cuts for a multi-product closed-loop supply chain network design model, Naval Res Logist 54 (2007), 890–907.
[64] T.J. Van Roy and D. Erlenkotter, A dual-based procedure for dynamic facility location, Management Sci 28 (1982), 1091–1105.
[65] A. Weber, Űber den standort der industrien, tűbingen, english translation, by C.J. Friedrich (1929), Theory of the location of industries, University of Chicago Press, 1909.
[66] E. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnes est minimum, Tohoku Math J 43 (1936), 355–386.