LOCATING DISCRETIONARY SERVICE FACILITIES, II: MAXIMIZING MARKET SIZE, MINIMIZING INCONVENIENCE

ODED BERMAN

University of Toronto, Toronto, Ontario, Canada

DIMITRIS BERTSIMAS and RICHARD C. LARSON

Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received November 1992; revision received June 1994; accepted October 1994)

Discretionary service facilities are providers of products and/or services that are purchased by customers who are traveling on otherwise preplanned trips such as the daily commute. Optimum location of such facilities requires them to be at or near points in the transportation network having sizable flows of different potential customers. N. Fouska (1988) and O. Berman, R. Larson and N. Fouska (BLF 1992) formulate a first version of this problem, assuming that customers would make no deviations, no matter how small, from the preplanned route to visit a discretionary service facility. Here the model is generalized in a number of directions, all sharing the property that the customer may deviate from the preplanned route to visit a discretionary service facility. Three different generalizations are offered, two of which can be solved approximately by greedy heuristics and the third by any approximate or exact method used to solve the p-median problem. We show for those formulations yielding to a greedy heuristic approximate solution, including the formulation in BLF, that the problems are examples of optimizing submodular functions for which the G. Nemhauser, L. Wolsey and M. Fisher (1978) bound on the performance of a greedy algorithm holds. In particular, the greedy solution is always within 37% of optimal, and for one of the formulations we prove that the bound is tight.

Fouska (1988) and Berman, Larson and Fouska (BLF 1992) considered a problem formulation in location theory which was called "optimal location of discretionary service facilities." Independently, and at approximately the same time, Hodgson (1990) formulated the same problem and called it a "flow-capturing model." The motivation for this class of problems is a perceived behavioral change on the part of customers. Instead of undertaking a one-stop tour from home or workplace to a facility to purchase a service or product, it was argued that many customers now carry out such purchases as part of routine preplanned trips, say on the daily commute to and from home and workplace. Examples include stopping at gasoline service stations, automatic teller machines, and "convenience stores." Traditional "Hakimi type" location models focus on minimizing some measure of travel distance or travel time from home (or workplace) to the facility. The optimal location of discretionary service facilities, on the other hand, requires convenience with regard to the customer's preplanned trip.

The focus in BLF was on locating the m-discretionary service facilities to maximize the flow of potential customers who passed at least one discretionary service facility along their preselected travel paths from origin to destination. A path containing a facility was "covered;" a path not containing a facility, even if there existed a facility δ travel units from the path ($\delta > 0$), was not covered. BLF proved that an optimal set of facility locations exists on the nodes of the network, and both exact and heuristic algorithms were developed to solve the problem.

In this paper, we relax the assumption that, to be useful to a potential customer, a facility must be located at some point precisely on her preplanned travel path. Facilities located "near" the preplanned path may also be utilized by the customer.

In the first generalization, which we call *delta coverage* or problem $\mathbf{P}_1$, we assume that, as in BLF, the customer passes through each node of her preplanned trip path. If there are no facilities on the path, she is willing to detour a maximum distance Δ from any one of the path nodes to travel to a discretionary service facility. After purchasing the service or product at the facility, she returns to the same path node, implying that a total detour travel distance of up to 2Δ is incurred. This model depicts a situation in which a detour of up to 2Δ travel units, starting and ending at

Subject classifications: Financial institutions, banks: locating automatic teller machines. Networks/graphs, heuristics: locating n points to intercept max flow. Transportation, models, location: maximizing flow of potential customers.
Area of review: SERVICES.

0030-364X/95/4304-0623 $01.25
© 1995 INFORMS

one of the preplanned path nodes, is associated with a zero disutility on the part of the customer. Any detour requiring more than 2Δ travel units has, in effect, infinite disutility, so the detour will not be executed and the associated service (product) will not be purchased. We show that this problem can be reduced to the problem solved in BLF.

In the second generalization, which we call *maximize market size* or $\mathbf{P}_2$, we assume that customers are increasingly likely to balk at traveling to a service facility as the deviation distance to it increases. More precisely, we assume that the probability that a customer is willing to travel an extra d units of distance to a facility that is located off the preplanned path, assuming that this facility is the least inconvenient to the original path, is a convex decreasing function of d. The objective is to locate the m facilities to maximize the expected number of potential customers who become actual customers at the facilities. We develop both exact and approximate algorithms to solve this problem.

Problems $\mathbf{P}_1$ and $\mathbf{P}_2$ can be considered within one unified context, namely that of minimizing total customer disutility. Consider the disutility of the customers as a function of the (extra) distance d they must travel for service. Denote this disutility function by $U(d)$. One can interpret $U(d)$ as the probability of customer balking (not being willing to travel an extra d units of distance to a facility). In $\mathbf{P}_1$ the disutility function is a step function ($U(d) = 0$ for $d \leq \Delta$, $U(d) = 1$ for $d > \Delta$). In $\mathbf{P}_2$, $U(d)$ is a concave increasing function of d.

In the third generalization, which we call *minimize expected inconvenience* or $\mathbf{P}_3$, we assume that all potential customers traveling on the network must purchase the service at a service facility, regardless of the extra distance that must be traveled to get to the facility. In this sense the facilities are no longer discretionary. For some (lucky) customers, there will be a facility on their preplanned travel paths, and no inconvenience is incurred. For others, the customers must deviate from the preplanned trip to travel to the service facility causing the least digression from the originally selected path. We assume that customers select their deviation paths to be the shortest ones possible. The objective of this problem is to locate the service facilities to minimize the total deviation distance traveled per unit of time, or equivalently, to minimize the expected deviation distance traveled by a random customer. Hodgson (1981) was the first to identify and study this problem. He showed that this problem is essentially an m median problem, and any m median algorithm can be used to locate the m facilities. (The literature usually refers to the multifacility median problem as the p median problem, but p has another meaning herein; m is the number of facilities we consider.)

We present a generic worst-case analysis of all the (greedy) heuristics developed both in this paper and in BLF. We show that each of our models, with the exception of the median model ($\mathbf{P}_3$), belongs to a family of problems in which there is a e^{-1} worst-case bound associated with the greedy heuristic and the bound is tight. This result improves upon the worst-case bound published in BLF for the original form of the discretionary services location problem.

The paper includes specific algorithms, a greedy heuristic, and a branch-and-bound algorithm to solve problem $\mathbf{P}_2$, numerical results, and a section on conclusions and future research.

1. BACKGROUND AND NOTATION

Let $G(N, A)$ be a bidirectional urban transportation network, where N is the set of nodes with cardinality n and A is the set of arcs. We denote by $\underline{P}$ the set of non-zero flow paths through the network nodes and let f_p indicate the number of units of travel flow along any path $p \in \underline{P}$, per unit of time. The flow quantity f_p is not a decision variable, but rather is the known apriori number of units of flow along path p. Let m be the number of facilities to be located on the network. All facilities are assumed to provide identical service and thus no customer needs to stop at more than one of them on any given trip.

1.1. The Case of No Allowed Deviation

BLF examined the problem **P-BLF** of finding a set of m facilities on the network to maximize the total flow of different customers intercepted by the facilities under a specific assumption regarding the behavior of customers. It was assumed that a customer may receive service only from facilities located on his (her) preplanned trip path. In other words, customers cannot deviate from their preplanned paths. The problem can be formulated as follows.

Problem P-BLF

$$\operatorname*{Max}_{\bar{x}\in G} \sum_{p\in \underline{P}} f_p I(\bar{x}, p),$$

where $I(\bar{x}, p)$ is an indicator variable,

$$I(\bar{x}, p) = \begin{cases} 1 & \text{at least one } x \in \bar{x} \text{ is on path } p \\ 0 & \text{otherwise} \end{cases}$$

and $\bar{x}$ is a vector of m points in G. BLF show that G can be replaced with N in **P-BLF** because it is proved that an optimal set of locations exists in N.

1.2. Deviation Distances

In this paper, we relax the assumption that customers do not deviate from their preplanned trips when service is required. We define the deviation distance as the *extra* distance incurred when a customer deviates from her preplanned trip path. We denote by $d(a, b)$ the shortest distance (travel time) between a and b, $a, b \in G$.

In each analysis that follows, we allow one or more of three alternative deviation travel behaviors by the customer. The specific behavior allowed will be indicated in each respective section of the paper.

Consider an arbitrary customer who, without any travel deviation to a service facility, would travel on a preplanned path $p \in \underline{P}$, defined by a node visitation sequence $p = \{1, 2, \ldots, l\}$ (where 1 is the path p origin node and l is the destination node). To compute deviation distances we distinguish among three cases.

Case i. (*$p \in \underline{P}$ is a shortest path and shortest path deviations are allowed.*) The customer travels from point a to point b. If she chooses not to visit a service facility, she will travel on the s.p. (shortest path) from a to b. If she chooses to visit a facility at point x, she will take the s.p. from a to x, and the s.p. from x to b. The decision whether she visits a facility or not will depend on the minimum deviation. Thus, when p is a shortest path, setting the path p origin $a = 1$ and destination $b = l$, the deviation distance from the path p to the nearest of the m facilities, $D(p, \bar{x})$, is given by

$$D(p, \bar{x}) \equiv [\min_{x \in \bar{x}} D(p, x) \equiv \min_{x \in \bar{x}}\{d(1, x) + d(x, \ell) - d(1, l)\}].$$

Case ii. (*$p \in \underline{P}$ may or may not be a shortest path and all nodes in p must be visited in proper sequence.*) The customer will travel from a to b on a preplanned path. Even if she chooses to visit a facility, she must traverse the nodes on her preplanned path in the proper sequence; this could arise, for instance, because of ordered deliveries and/or pickups that must be made at the nodes visited. Setting the path p origin $a = 1$ and destination $b = l$, the deviation distance $D(p, \bar{x})$ is now given by

$$D(p, \bar{x}) \equiv \min_{j \in p, j \neq l}[\min_{x \in \bar{x}}\{d(j, x) + d(x, j + 1) - l(j, j + 1)\}],$$

where $l(j, j + 1)$ is the length of link $(j, j + 1)$.

Case iii. (*$p \in \underline{P}$ may or may not be a shortest path, all nodes in p must be visited in proper sequence, and only a simple one-tour deviation is allowed.*) The customer will travel from a to b on a preplanned path. Even if she chooses to visit a facility, she must traverse the nodes on her preplanned path in the proper sequence. In addition, any deviation to visit a facility must start and end at the same node on the preplanned path. This could arise, for instance, because of ordered deliveries and/or pickups that must be made along the arcs visited in p (as in the postal delivery or refuse collection). The deviation distance $D(p, \bar{x})$ is given by

$$D(p, \bar{x}) \equiv \min_{j \in p}[\min_{x \in \bar{x}}\{2d(j, x)\}].$$

As an example for the calculations of $D(p, \bar{x})$ we refer to Figure 1 that depicts a simple network with seven nodes. Suppose that for case i, $p = (2, 1, 4, 5)$ and a single facility is located at node 3, then

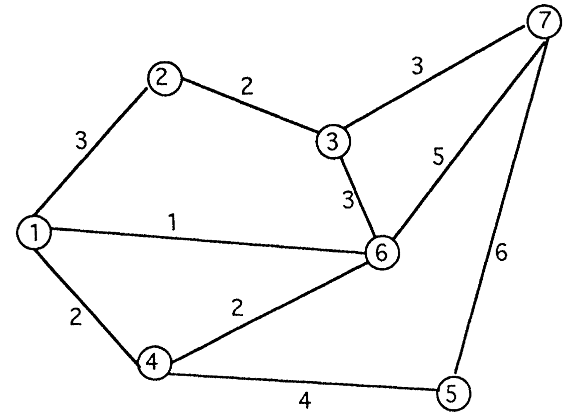

Figure 1. A 7-node network.

$D(p, \bar{x}) = 2 + 9 - 9 = 2$.

For case ii, let us assume that $p = (1, 2, 3, 6)$ and $\bar{x} = 7$; then $D(p, \bar{x}) = \min\{6 + 5 - 3, 5 + 3 - 2, 3 + 5 - 3\} = 5$. For case iii with the same path and facility as with case ii,

$D(p, \bar{x}) = (2)\min\{(6), (5), (3), (5)\} = 6$.

2. THE PROBLEMS

2.1. P_1: Delta Coverage

Delta coverage depicts a situation in which a round-trip detour of up to 2Δ travel units, starting *and* ending at the same preplanned path node, is allowed for the customer to visit a facility "nearest" to her preplanned route. It is assumed that the detour route is restricted to a tour comprising a minimum distance path from the detour-originating path node to a nearest facility and, due to network bidirectionality, a repeat of that path in reverse direction. Thus, we are assuming that the travel deviation behavior follows case iii.

The formulation of problem **P-BLF** is easy to extend to include problem $\mathbf{P_1}$. In $\mathbf{P_1}$ a customer is said to be intercepted by a facility if at least one facility is at a distance of, at most, Δ from a node on the customer's trip path p, that is:

Problem P_1

$$\operatorname*{Max}_{x \in G} \sum_{p \in \underline{P}} f_p I'(\bar{x}, p),$$

where we define

$$I'(\bar{X}, p) = \begin{cases} 1 & \exists\ j \in p \text{ such that } d(j, \bar{x}) \leq \Delta \\ 0 & \text{otherwise,} \end{cases}$$

where $d(j, \bar{x}) \equiv$ the shortest distance between $j \in N$ and a nearest facility located at $x \in \bar{x}$.

Let us define N' as the union of the node set N and the set of all points G_Δ in G that are exactly Δ units of distance away from a node, i.e., $N' = N \cup G_\Delta$, where $G_\Delta = \{y \in G | d(j, y) = \Delta, j \in N\}$. (Note that $\mathbf{P_1}$ reduces to problem **BLF** when $\Delta = 0$.)

Theorem 1. *An optimal set of locations for problem* $\mathbf{P}_1$ *exists in* N'.

Proof. The proof is straightforward and is omitted.

The algorithms of **BLF**, both exact and heuristic, can be applied directly to solve problem $\mathbf{P}_1$, with the set of possible facility locations now extended to N'.

2.2. $\mathbf{P}_2$: Maximize Market Size

The objective of problem $\mathbf{P}_2$ is to locate the m facilities to maximize the expected number of potential customers who become actual customers at the facilities. Here we allow customers to deviate from their preplanned route in any of the three manners described in subsection 1.2; in particular, detours are not restricted to be tours. We assume that as the deviation distance grows larger, customers become less and less likely to select the detour to visit a nearest facility. Thus, we again have a flow intercept maximization problem.

We assume that the flow of *path-p* customers to the "nearest" of m facilities in the location set $\bar{x}$ is a convex decreasing function of the deviation distance $D(p, \bar{x})$, denoted by $f_p g(D(p, \bar{x}))$, where $g(0) \equiv 1$. Here $g(D(p, \bar{x}))$ can be interpreted to be the fraction of path-p customers who would deviate to use a facility in $\bar{x}$ closest to path p, or equivalently, the probability that a random path-p customer will deviate to use that facility. Therefore, problem $\mathbf{P}_2$ is as follows.

Problem $\mathbf{P}_2$

$$\max_{\bar{x} \in G} \sum_{p \in \underline{P}} f_p g(D(p, \bar{x})).$$

Theorem 2. *An optimal set of locations to* $\mathbf{P}_2$ *exists in* N.

Proof. Case i (all customer paths are minimum distance paths):

Path/facility assignment fixed: For any $x \in \bar{x}$, let $\underline{P}_x$ be the set of all paths in $\underline{P}$ that route all or a fraction of its customers to the facility at x. If x lies on path $p \in \underline{P}_x$, there is no deviation distance and the corresponding component of the objective function is f_p. Consider a path $p \in \underline{P}_x$ not containing x, where, as usual, 1 is the origin node and l is the destination node of p. Suppose that $x \in \bar{x}$ is an interior point on link (a, b), having length $d(a, b)$. Assume that the distance from node a to x is $Qd(a, b)$, where $0 \leq Q \leq 1$. Then the deviation distance from p to x, defined as $D(p, x)$, is found by computing the minimum of four possible detour routes to and from x:

$$\begin{aligned} D(p, x) &= d(1, x) + d(x, l) - d(1, l) \\ &= \min\{d(1, a) + 2Qd(a, b) + d(a, l) \\ &\qquad - d(1, l);\ d(1, a) + d(a, b) \\ &\qquad + d(b, l) - d(1, l);\ d(1, b) \\ &\qquad + d(a, b) + d(a, l) \\ &\qquad - d(1, l);\ d(1, b) + 2(1 - Q)d(a, b) \\ &\qquad + d(b, l) - d(1, l)\}. \end{aligned}$$

Since $D(p, x)$ is the lower envelope of linear functions of Q, it is a piecewise linear concave function of Q. Since $g(y)$ is a decreasing convex function of y and $D(p, x)$ is a piecewise linear concave function of Q, $g(D(p, x))$ is a convex function of Q. Since a sum of convex functions is convex, we know that the partial sum

$$\sum_{p \in \underline{P}_x} f_p g(D(p, x))$$

is convex. Since Q is defined on the closed interval $0 \leq Q \leq 1$, a maximum of the partial sum must exist at an endpoint corresponding to $Q = 0$ or $Q = 1$.

Path/facility reassignment: Any change in location of the examined facility from an interior point on (a, b) to one of the arc's nodes may, in turn, cause a reassignment of customers on one or more paths in $\underline{P}_x$ to other facilities and/or may cause reassignment of customers on paths not in $\underline{P}_x$ to the examined facility. But such path/facility reassignment cannot decrease the objective function.

Cases ii and iii (the two instances requiring nodes to be visited in proper sequence, subsection 1.2): Both are proved in a similar fashion.

2.3. $\mathbf{P}_3$: Minimize Expected Inconvenience

In problem $\mathbf{P}_3$ *all* customers must travel to a service facility "closest" to their preplanned paths to purchase or consume the service provided there. "Closeness" of a facility to a path is measured in terms of the minimum deviation distance (subsection 1.2). The objective of $\mathbf{P}_3$ is to locate the service facilities to minimize the total deviation distance traveled per unit time, or equivalently, to minimize the expected deviation distance traveled by a random customer, that is:

Problem $\mathbf{P}_3$

$$\min_{\bar{x} \in G} \sum_{p \in \underline{P}} f_p D(p, \bar{x}),$$

where the deviation distance can be calculated according to any of the cases i, ii, and iii.

As noted by Hodgson (1981), who first posed this problem, the objective function of $\mathbf{P}_3$ is identical in form to that of the well known "m-median" type problem (see, for example, Mirchandani and Francis 1990). For the m-median problem, Hakimi (1964, 1965) proved that an optimal set of facility location exists on the nodes of the network. Thus, in $\mathbf{P}_3$ the search for optimal locations in G in the objective function can be replaced with a search limited to the node set N. Any of the algorithms, heuristic or exact, developed and used to solve the (NP-hard) m-median problem can be used for $\mathbf{P}_3$. However, as an instance of the m-median, $\mathbf{P}_3$ cannot be approximated within a constant factor unless $P = NP$ (see Nemhauser and Wolsey 1988). Moreover, it is well known that the greedy heuristic applied to the m-median problem provides solutions with values arbitrarily bad from the optimal. In the rest of the paper we focus on problems $\mathbf{P_1}$ and $\mathbf{P_2}$.

2.4. A Simple Numerical Example

We illustrate the three alternative formulations above with a simple 3-node example network, as depicted in Figure 2. As usual, the numbers adjacent to the links are their respective lengths. Suppose that there are three paths having positive customer flow rates: $\underline{P} = \{1\text{–}2, 2\text{–}3, 3\text{–}1\}$, with the flows given as $f_{1\text{-}2} = 110, f_{2\text{-}3} = 70$, and $f_{3\text{-}1} = 80$. Consider the problem of locating only one facility, and let X be the nodal facility location (with one exception, for P_1 with $\Delta > 0$ we let $X \in N'$).

Problem $\mathbf{P_1}$. Consider problem $\mathbf{P_1}$ with $\Delta = 0$. The total flow through node i is 190, 180, and 150 for $i = 1, 2$, and 3, respectively. Hence, node 1 is the optimal location.

Suppose that for this problem we had set $\Delta = 1.5$. Inspection of the network shows that the optimal location for the facility is half-way between nodes 2 and 3, representing a situation in which all customers are "covered."

Problem $\mathbf{P_3}$. Now consider problem $\mathbf{P_3}$ for deviation distance according to cases i and ii (the calculations for case iii are similar). In each case, one set of customers, representing one of the three flow patterns, must deviate to visit the facility. The $\mathbf{P_3}$ objective function that we wish to minimize is the product of the flow of the customers who must deviate and their deviation distance. For instance, placing the facility at node 1 requires 70 customers per unit time to deviate from their 2–3 path (of length 3) to a 2-1-3 path of length 11.5; the corresponding objective function value is 70 (11.5 − 3) = 595. Similar values for a facility at nodes 2 and 3 are 80 (3.5) = 280 and 110 (2.5) = 275, respectively. Hence node 3 is optimal.

Problem $\mathbf{P_2}$. Finally, consider problem $\mathbf{P_2}$ for deviation distance according to cases i and ii with $g(D(p, x)) = \exp(-bD(p, x))$. The values of the objective function for facility placement at nodes 1, 2 and 3, respectively, are $190 + 70\exp(-8.5b)$, $180 + 80\exp(-3.5b)$ and $150 + 110\exp(-2.5b)$. For $0 < b < 0.0367$, node 3 is optimal. For $b > 0.58$, node 1 is optimal. For $0.0367 < b < 0.58$, node 2 is optimal.

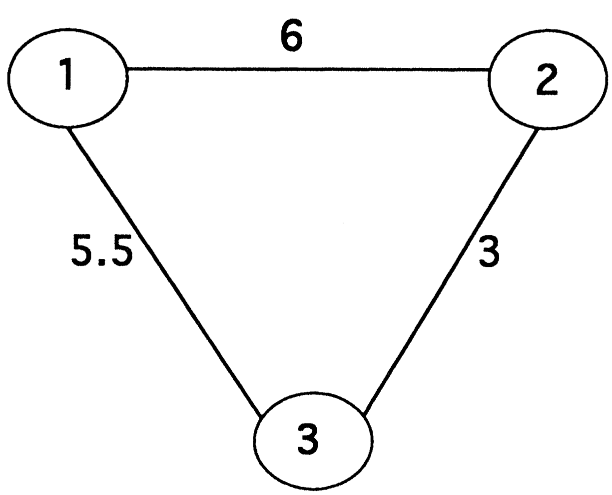

Figure 2. A 3-node network for a simple example.

3. WORST CASE ANALYSIS OF GREEDY ALGORITHMS

In this section, we analyze a generic greedy algorithm for problems $\mathbf{P_1}$ and $\mathbf{P_2}$. We show that the solutions given by the greedy algorithm are always within 37% of the optimal solution and this bound is tight. In Section 4 we develop a specific greedy algorithm for $\mathbf{P_2}$, as well as a branch-and-bound algorithm, and we give numerical results.

Problems $\mathbf{P_1}$ and $\mathbf{P_2}$ can be formulated generically as follows: Let h_i: $2^N \to R+$ $(i = 1, 2)$ be set functions defined on subsets of the set N. Then, problem $\mathbf{P_i}$, $(i = 1, 2)$ can be formulated as follows.

Problem $\mathbf{P_i}$

$$Z_i = \max_{S \subseteq N_i, |S| = m} h_i(S),$$

where the functions $h_i(S)$ are defined in Section 2 and $N_1 = N' = N \cup G_\Delta$, $N_2 = N$.

A generic greedy algorithm for problems $\mathbf{P_1}$, $\mathbf{P_2}$, proposed in the literature (see, for example, Wolsey 1983), is as follows.

Greedy Algorithm

[Input: $h_i(S)$, m, N_i]

[Output: R_G, Z_G]

1. **(Initialization)**

$R^0 \leftarrow \varnothing$, $t \leftarrow 1$.

2. **(Main Loop)**

For $t = 1, \ldots, m$

$j_t \leftarrow \operatorname*{argmax}_{j \in N_i \setminus R^{t-1}} h_i(R^{t-1} \cup \{j\})$

$R^t \leftarrow R^{t-1} \cup \{j_t\}$

3. **(Output)**

$R_G = R^m$

$Z_G = h(R^m)$.

Given the function $h(S)$, the number of facilities m and the set of potential locations N_i, the algorithm outputs a set R_G of m facilities with value Z_G.

Nemhauser, Wolsey and Fisher (1978), and Nemhauser and Wolsey (1978) studied the problem of max)($h(S)$), where $h(S)$ is a submodular and nondecreasing function and where $S \subseteq N$, $|S| \leq m$. We suppressed the subscript on the node set N_i and h_i. A set function is called submodular if for all $S, T \subseteq N$,

$$h(S \cap T) + h(S \cup T) \leq h(S) + h(T)$$

and is nondecreasing if for all S, T, $S \subseteq T$, we have $h(S) \leq h(T)$.

Obtaining the exact solution to the problem of maximizing a submodular set function is NP-hard. The greedy algorithm, however, provides very good solutions, yielding results "close" to the optimum. More precisely, Nemhauser, Wolsey and Fisher prove the next theorem.

Theorem 3. (Nemhauser, Wolsey and Fisher) *The value Z_G returned by the greedy algorithm when applied to the problem:*

$$Z_* = \max_{S \subseteq N, |S| \leq m} h(S) \tag{1}$$

for $h(S)$ being submodular and nondecreasing satisfies

$$\frac{Z_G}{Z_*} \geq 1 - \left(1 - \frac{1}{m}\right)^m \geq 1 - \frac{1}{e} \cong 0.63.$$

In other words, the greedy algorithm returns a solution that is optimal for $m = 1$ and is within 37% from the optimal solution value for any value of m. Moreover, the bound is tight, i.e., there are instances in which $Z_G = (Z_*)[1 - (1 - 1/m)^m]$.

Furthermore, Nemhauser and Wolsey (1981) have shown that within a large class of algorithms the greedy algorithm is the best possible for problem **1**.

We plan to show in the remainder of this section that problems $\mathbf{P_1}$ and $\mathbf{P_2}$ are instances of problem **1**, i.e., the functions $h_i(s)$, $(i = 1, 2)$ are submodular and nondecreasing. A function f is called supermodular if $-f$ is submodular.

3.1. Delta Coverage

With the usual bidirectional network $G = (N, A)$, let $d(x, i)$ be the length of the shortest path from node x to node $i)(x, i \in N)$ and let $N(x, \Delta) = \{i \in N: d(x, i) \leq \Delta\}$ be the set of nodal delta coverage points associated with node x. Let $\underline{P}$ be the given set of paths, assuming that a path $p \in \underline{P}$ is specified as a set of nodes in N, i.e., $p = \{1, \ldots, l\}$. A customer whose travel path includes at least one node n_j in $N(x, \Delta)$ could travel from that node to a service facility located at x and back, incurring a detour travel distance not exceeding 2Δ; in that way, a facility located at x covers any node $n_j \in N(x, \Delta)$.

Let

$$N(S, \Delta) = \bigcup_{x \in S} N(x, \Delta)$$

be the set of nodal delta coverage points corresponding to any subset S of potential facility locations, $S \subseteq N \cup G_\Delta$ (see Figure 3). A customer whose travel path includes at least one node in $N(S, \Delta)$ could travel from that path node to and from a facility, with a detour travel distance not exceeding 2Δ, if at least one facility is located in S.

Using Theorem 1, problem $\mathbf{P_1}$ is formulated as:

$$\operatorname*{Max}_{\substack{S \subseteq N \cup G_\Delta \\ |S| \leq m}} h_1(S) = \sum_{\substack{p \in \underline{P} \\ p \cap N(S, \Delta) \neq \varnothing}} f_p .$$

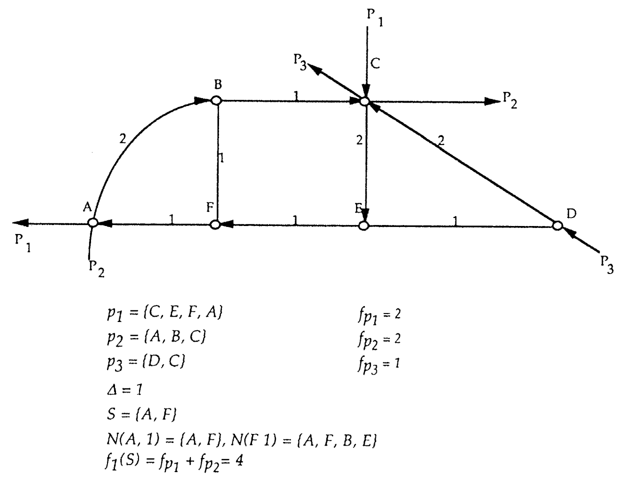

Figure 3. Example for delta coverage.

Note that for $\Delta = 0$, $N(x, \Delta) = \{x\}$, $N(S, \Delta) = S$ and the problem reduces to the problem studied in Berman, Larson and Fouska.

Proposition 1. *If $f_p \geq 0$ for all $p \in \underline{P}$, then for any $\Delta \geq 0$, $h_1(S)$ is submodular and nondecreasing.*

Proof

- If $S \subseteq T$, then clearly $N(S, \Delta) \subseteq N(T, \Delta)$ which implies that $h_1(S) \leq h_1(T)$ if $f_p \geq 0$, i.e., $h_1(S)$ is nondecreasing.
- To show that a set function is submodular it suffices to show that for *all* $S \subseteq T$ and $k \notin T$, $h_1(T \cup \{k\}) - h_1(T) \leq h_1(S \cup \{k\}) - h_1(S)$.

For any S and T, $S \subseteq T$, let $\underline{P}_T = \{p \in \underline{P}: p \cap N(T, \Delta) = \varnothing\}$, i.e., $\underline{P}_T$ is the set of all paths not covered by locating facilities in T. Since $S \subseteq T$, then $N(S, \Delta) \subseteq N(T, \Delta)$ which implies that $\underline{P}_T \subseteq \underline{P}_S$. Then,

$$\begin{aligned} h_1(T \cup \{k\}) - h_1(T) &= \sum_{\substack{p \in \underline{P} \\ p \cap N(T \cup \{k\}, \Delta) \neq \varnothing}} f_p - \sum_{\substack{p \in \underline{P} \\ p \cap N(T, \Delta) \neq \varnothing}} f_p \\ &= \sum_{\substack{p \in \underline{P} \\ p \cap N(T, \Delta) = \varnothing \\ p \cap N(k, \Delta) \neq \varnothing}} f_p = \sum_{\substack{p \in \underline{P}_T \\ p \cap N(k, \Delta) \neq \varnothing}} f_p \\ &\leq \sum_{\substack{p \in \underline{P}_S \\ p \cap N(k, \Delta) \neq \varnothing}} f_p \\ &= h_1(S \cup \{k\}) - h_1(S), \end{aligned}$$

i.e., $h_1(S)$ is submodular.

3.2. Maximize Market Size

In subsection 1.2 path deviation distances were defined for two cases: all customer paths are minimum distance paths; and not all customer paths are minimum distance paths. For the former case, for each $p \in \underline{P}$ and $S \subseteq N$, we can write the deviation distance of path p from set S as:

$$D(p, S) = \underset{\substack{i,j \in p \\ x \in S}}{\text{Min}}[d(i, x) + d(x, j) - d(i, j)\}.$$

Let $g(y): R^+ \to R^+$ be a nondecreasing function, not necessarily convex. From Theorem 2, $\mathbf{P_2}$ can be formulated as:

$$\underset{\substack{S \subseteq N \\ |S| \leq m}}{\text{Max}}\ h_2(S) = \sum_{p \in \underline{P}} f_p g(D(p, S)).$$

In the following proof, we assume that all customer paths are minimum distance paths. A directly analogous proof applies to the other case.

Proposition 2. *If $g(y)$ is a nonincreasing function, then $h_2(S)$ is a submodular, nondecreasing set function.*

Proof. If $S \subseteq T$, then $D(p, S) \geq D(p, T)$ which implies that $g(D(p, S)) \leq g(D(p, T))$, because $g(y)$ is nonincreasing. Therefore, $h_2(S) \leq h_2(T)$, i.e., $h_2(S)$ is nondecreasing. We will now show that $g(D(p, S))$ is submodular. Let $S \subseteq T$ and $k \notin T$. We first note that

$$[D(p, T \cup \{k\})]$$
$$= \min\{D(p, T), \underset{i,j \in p}{\text{Min}}[d(i, k) + d(k, j) - d(i, j)].$$

To check whether $g(D(p, S))$ is submodular it suffices to show that

$$g(D(p, T \cup \{k\})) - g(D(p, T))$$
$$\leq g(D(p, S \cup \{k\})) - g(D(p, S)),$$

or equivalently, defining

$$\alpha \equiv \underset{i,j \in p}{\text{Min}}[d(i, k) + d(k, j) - d(i, j)],$$

$$g(\text{Min}[D(p, T), \alpha]) - g(D(p, T))$$
$$\leq g(\text{Min}[D(p, S), \alpha]) - g(D(p, S)). \tag{2}$$

To check (2) we distinguish three cases:

1. ($\alpha \geq D(p, S) \geq D(p, T)$) Then (2) becomes

$$g(D(p, T)) - g(D(p, T))$$
$$\leq g(D(p, S)) - g(D,(p, S)),$$

which is obviously satisfied.

2. ($D(p, S) \geq \alpha \geq D(p, T)$) Then (2) becomes

$$0 \leq g(\alpha) - g(D(p, S)),$$

which is satisfied because $g(y)$ is nondecreasing.

3. ($D(p, S) \geq d(p, T) \geq \alpha$) Then (2) becomes

$$g(\alpha) - g(D(p, T)) \leq g(\alpha) - g(D(p, S)),$$

which is satisfied because $g(D(p, S)) \leq g(D(p, T))$. ($S \subseteq T$ and $D(p, S) \geq D(p, T)$.)

Therefore, the set function $g(D(p, S))$ is submodular. Then, if $f_p \geq 0$, $h_2(S)$ is also submodular because it is a sum of submodular set functions.

Note that the submodularity of $h_2(S)$ is independent of any convexity assumption on $g(x)$.

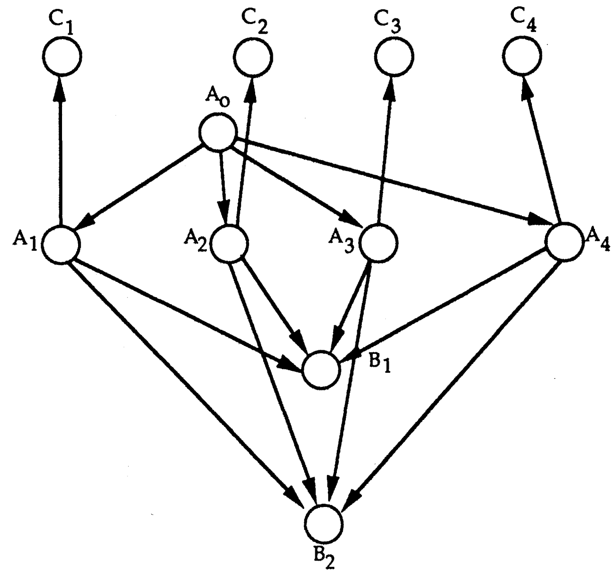

Figure 4. Example in the proof for $m = 4$.

3.3. Main Theorem

As a result of Propositions 1 and 2, Theorem 3 holds for $\mathbf{P_1}$ and $\mathbf{P_2}$, i.e., the generic greedy algorithm produces a value Z_{G_i} ($i = 1, 2$) such that

$$\frac{Z_{G_i}}{Z_{*i}} \geq 1 - \left(1 - \frac{1}{m}\right)^m. \tag{3}$$

We note that the tightness of the bound is proved in Nemhauser and Wolsey (1981) for arbitrary submodular functions. Still remaining is the question whether a better bound can be found for our problem. Therefore, in the following theorem we investigate the tightness of the bound.

Theorem 4. *For problem $\mathbf{P_1}$ with $\Delta = 0$, the bound (3) is tight.*

Proof. Consider a network with nodes $A_0, A_1, \ldots, A_m, B_1, B_2, \ldots, B_{m-2}, C_1, \ldots, C_m$. The set of edges is as follows: $E = \{(A_0, A_i), (A_i, C_i); i = 1, \ldots, m\} \cup \{(A_i, B_j); i = 1, \ldots, m, j = 1, \ldots, m-2\}$. (See Figure 4 for $m = 4$.) The set $\underline{P}$ of paths is as follows: For $i = 1, \ldots m$

m paths (A_0, A_i) each with value $f_0 = m^{m-2}$

m paths (A_i, B_1) each with value $f_1 = \dfrac{m-1}{m} m^{m-2}$

$\vdots$

m paths (A_i, B_j) each with value $f_j = \left(\dfrac{m-1}{m}\right)^j m^{m-2}$, $j = 1 \ldots m-2$

m paths (A_i, C_i) each with value $f_{m-1} = (m-1)^{m-1}$.

The optimal solution of $\mathbf{P_1}$ with $\Delta = 0$ and up to m facilities is the set $(A_1, \ldots, A_m)$ of the nodes at the first level that covers all the paths. The value of the optimal solution is

$$Z_* = mf_0 + m \sum_{j=1}^{m-2} f_j + mf_{m-1} = m^m.$$

The greedy algorithm selects node A_0 first, then node B_1, then node $B_2, \ldots,$ node B_{m-2} and, finally, one of $A_1, \ldots, A_m$.

The cost of the greedy algorithm is

$$Z_G = mf_0 + m \sum_{j=1}^{m-2} f_j + f_{m-1} = m^m - (m-1)^m.$$

The reason is that the value of A_0 is at least as large as any other's node, B_1 is at least as high as any from the remaining nodes, etc. Thus, $Z_G/Z_* = 1 - (1 - 1/m)^m$.

4. SOLVING PROBLEM $\mathbf{P_2}$

Since $\mathbf{P_1}$ can be solved using the algorithms of BLF and $\mathbf{P_3}$ is essentially an m-median problem, in this section we focus on solution methods for $\mathbf{P_2}$. Building on our standard notation, let $D(p, i)$ be the deviation distance for a path p customer to detour through node i. This notation applies whether or not strictly minimum distance paths are used. Define the *potential* expected path p flow through node i as

$$C_{pi} = f_p g(D(p, i)) \equiv f_p g(\min_{j \in p} D(j, i)),$$

where

$$D(j, i) = \begin{cases} d(j, x) + d(x, l) - d(j, l) & \text{for case i} \\ d(j, x) + d(x, j+1) - d(j, j+1) & \text{for case ii} \\ 2d(j, x) & \text{for case iii.} \end{cases}$$

Since $g(.)$ is a convex decreasing function,

$$\sum_{p \in \underline{P}} f_p g(D(\bar{x}, p)) = \sum_{p \in \underline{P}} \max_{i \in \bar{x}} C_{pi}.$$

Therefore $\mathbf{P_2}$ can be formulated as:

$$\max_{\bar{x} \in N} \sum_{p} \max_{i \in \bar{x}} C_{pi}.$$

We can now identify binary decision variables and assignment variables. Let

$$X_{pi} = \begin{cases} 1 & \text{if a facility } i \text{ is assigned to path } p \\ 0 & \text{otherwise} \end{cases}$$

and

$$X_j = \begin{cases} 1 & \text{if a facility is located at node } j \\ 0 & \text{otherwise.} \end{cases}$$

Now $\mathbf{P_2}$ can be stated as a mathematical programming problem:

$$\max \sum_{p \in \underline{P}} \sum_{i=1}^{n} C_{pi} X_{pi}$$

subject to $\displaystyle\sum_{j=1}^{n} X_j = m$

$$X_j - X_{pj} \geq 0 \quad j = 1, 2, \ldots, n$$

$$\sum_{j=1}^{n} X_{pj} = 1 \quad \text{for all } p \in \underline{P}.$$

The first constraint guarantees that m facilities are located. In the second set of constraints we have made sure that path p cannot be assigned to a node that does not house a facility, whereas the last set of constraints ensures that each path is assigned to exactly one facility.

To solve $\mathbf{P_2}$, we developed a greedy heuristic (having the properties of the generic greedy heuristic of Section 3) and an exact algorithm. The greedy heuristic is a modification of the greedy heuristic to solve **BLF**.

The Greedy Heuristic to Solve $\mathbf{P_2}$

STEP 1. $l = 1$.

STEP 2. Compute $b_i = \sum_{p \in \underline{P}} f_p g(d(i, p)),\ i = 1, \ldots, n$.

STEP 3. Find $b_{i_{max}} = \max_{i \in N}\{b_i\}$; locate facility l at node i_{max} and delete i_{max} from N.

STEP 4. For all $p \in \underline{P}$, for all $j \in N$, set

$$g(D(i, p)) = [g(D(i, p)) - g(D(i_{\max}, p))]^+.$$

Delete from $\underline{P}$ the set $\underline{P}_{i_{\max}}$ (which is the set of all paths p for which $g(D(i_{\max}, p)) = 0$.)

STEP 5. If $l = m$ or if $\underline{P} = 0$, STOP. Otherwise set $l = l + 1$ and go to Step 2.

As an example, let us refer again to the network in Figure 1 and let us solve problem $\mathbf{P_2}$ with an exponential customer damping factor, i.e., $g(D(p, \bar{x})) = e^{-bD(p, \bar{x})}$. The paths and all extra distances are given in Table I. Suppose that $b = 0.05$ and $f_p = 1$ for all $p \in \underline{P}$. We start the heuristic with $l = 1$.

Since $b_1 = 9.413$, $b_2 = 8.781$, $b_3 = 8.993$, $b_4 = 9.065$, $b_5 = 8.171$, $b_6 = 9.4999$, $b_7 = 8.722$, $b_{i_{max}} = b_6$, we locate facility 1 at node 6 and we delete node 6 from N. We find new $e^{-bD(p,j)}$ for all $p, j \neq 6$ and delete from $\underline{P}$ paths 4, 5, 6, 7, 8, and 9. Since $l = 1$, $\underline{P} = \varnothing$, we let

Table I

Paths and Extra Distances for the Example

Nodes /	Path	1	2	3	4	5	6	7
1	2,145	0	0	2	0	0	1	2
2	237	4	0	0	7	10	4	0
3	145	0	6	7	0	0	1	6
4	167	0	2	1	3	6	0	0
5	361	0	1	0	3	11	0	5
6	1	2	0	0	0	0	0	
7	463	1	2	0	0	6	0	5
8	546	1	2	6	0	0	0	5
9	612	0	0	1	3	11	0	6
10	75	6	8	6	5	0	5	0

$l = 2$. Now $b_1 = 0.0975$, $b_2 = 0.2300$, $b_3 = 0.1812$, $b_4 = 0.0975$, $b_5 = 0.3187$, $b_7 = 0.4024$, $b_{i_{max}} = b_7$ and we locate the second facility at node 7, and we delete node 7 from N. We find new $e^{-bD(p,j)}$ and delete from $\underline{P}$ paths 2 and 10. Since $l = 2$, $P = \varnothing$ we set $l = 3$, and find that $b_1 = 0.0975$, $b_2 = 0.0487$, $b_3 = 0$, $b_4 = 0.0975$, $b_5 = 0.0975$, $i_{\max} = 1$ (or 4 or 5). Since $l = 3$, we are done. The total flow intercepted is 9.4999 + 0.4024 + 0.0975 = 9.9998.

To solve the problem $\mathbf{P_2}$ exactly we developed a branch-and-bound code for the problem. This branch-and-bound procedure is based on two upper bounds. The first one, which is a minor modification of the one developed in (1) for $\mathbf{P_1}$, is called UB_1. The variables $X_1, \ldots, X_n$ are the decision variables for the branch-and-bound tree. Let us define $D \subset N$ as a set of all variables that constitute a partial solution in the branch-and-bound tree (i.e., $j \in D \Rightarrow X_j = 0, 1$) and let $U = N - D$. Let $D_1 \subset D$, $|D_1| = l$, be the set of all nodes in D that house a facility ($j \in D_1 \Rightarrow X_j = 1$) and let $D_0 = D - D_1 (j \in D_0 \Rightarrow X_j = 0)$. Let

$$r_i = \max_{j \in D_1} C_{ij},\ i = 1, \ldots, |\underline{P}|,$$

i.e., r_i is the maximum amount of flow that the facilities in the partial solution intercept from path i. For each $j \in U$ we define

$$\delta_j = \sum_{i=1}^{|\underline{P}|} \max\{0,\ C_{ij} - r_j\},$$

i.e., δ_j is maximum amount of flow node j can intercept after deleting all the flow intercepted by the already-located facilities. Now UB_1 can be defined as

$$UB_1 = \sum_{i=1}^{|\underline{P}|} r_i + L$$

where L is the sum of the $(m - l)$ greatest δ_j's. The second upper bound called UB_2 can be now defined as

$$UB_2 = \sum_{p \in P} \max_{j \in D_1 \cup U} C_{pj}.$$

We note that UB_2 is useful when there are many variables in a partial solution for which $X_j = 0$, which is exactly the situation when UB_1 is not useful. Therefore, UB_1 and UB_2 complement each other in the branch and bound. Finally, the upper bound is the minimum of UB_1 and UB_2.

The computer code implementing the branch-and-bound algorithm for problem $\mathbf{P_2}$ is written in C and tests were run on a DEC 5810. To provide test results we randomly generated network sizes, their paths and corresponding flows. Table II illustrates a typical sample of our test cases for the problem with an exponential damping factor. The table provides the CPU time and the ratio of the solution value provided by the greedy heuristic and the branch and bound for networks with a number of nodes and a number of paths ranging from 20–100 and a number of facilities ranging from 2–5. We see that for this set of runs the greedy heuristic performs considerably better than its worst case bound.

Table II
CPU Times in Seconds (Rounded to Closest Integer) and the Ratio of Objective Function of the Greedy Heuristic to the Objective Function of the Branch-and-Bound Procedure for Several n, $|P|$ and m Values

n	$\|P\|$	m	CPU of the Branch-and-Bound Algorithm	Value of Greedy / Value of Branch and Bound
10	10	2	0	1
		3	0	1
		4	0	1
		5	0	1
30	30	2	0	0.998
		3	1	0.937
		4	5	0.951
		5	22	0.993
50	50	2	1	0.983
		3	11	0.935
		4	69	0.968
		5	355	0.955
100	100	2	5	1
		3	169	0.936
		4	693	0.934
		5	2,147	0.943

5. CONCLUSIONS AND FUTURE RESEARCH

In this paper, the **BLF** model is generalized to consider the possibility for customers of discretionary services to deviate from preplanned tours to visit a discretionary service facility. Three models are considered: the delta coverage problem $\mathbf{P_1}$, where a customer will deviate from the preplanned tour if the facility is at a distance of at most Δ from the tour; the maximize market size problem $\mathbf{P_2}$, where the number of customers that travel to the facility is a decreasing function of the deviation distance; and the minimize expected inconvenience problem, where the total deviation distance traveled is minimized.

The main results of the paper are:

1. Node optimality property holds for $\mathbf{P_2}$ and $\mathbf{P_3}$. For problem $\mathbf{P_1}$ an optimal set of locations exists in a set N' which is the union of N and the set G_Δ (set points that are at a distance of Δ from the nodes.)
2. Problem $\mathbf{P_1}$ can be solved using the algorithms developed for the **BLF** (applied to the set of candidate locations N').
3. Problems $\mathbf{P_1}$ and $\mathbf{P_2}$ belong to a family of problems for which the greedy heuristic gives a solution that is guaranteed to be within 37% of the optimal solution and this bound is tight.
4. Problem $\mathbf{P_2}$ can be formulated as an integer program problem and a branch-and-bound algorithm to solve it is given.

We offer two directions for future research: Study problems $\mathbf{P}_1$ and $\mathbf{P}_2$ when the facilities are allowed to compete with each other. Study the polyhedral structure of problem $\mathbf{P}_2$ to improve the branch-and-bound algorithm developed in this paper.

ACKNOWLEDGMENT

The research of Oded Berman was partially supported by the National Science and Engineering Council of Canada. The research of Dimitris Bertsimas was partially supported by the National Science Foundation under grant DDM-9014751 and by the Presidential Young Investigator award DDM-9158118 with matching funds from Draper Laboratory. The work of Richard Larson was supported by National Science Foundation grant SES-9119962. We thank two anonymous referees whose suggestions helped to improve the style of the paper.

REFERENCES

BERMAN, O., R. C. LARSON AND N. FOUSKA. 1992. Optimal Location of Discretionary Service Facilities. *Trans. Sci.* **26**(3), 201–211.

FOUSKA, N. 1988. Optimal Location of Discretionary Service Facilities. MS Thesis, Operations Research Center, MIT, Cambridge, Mass.

HAKIMI, S. L. 1964. Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. *Opns. Res.* **12,** 450–459.

HAKIMI, S. L. 1965. Optimum Distribution of Switching Centers in a Communication Network and Some Related Graph Theoretic Problems. *Opns. Res.* **13,** 462–475.

HODGSON, J. 1981. The Location of Public Facilities Intermediate to the Journey to Work. *Eur. J. Opnl. Res.* **6**(2), 199–204.

HODGSON, J. 1990. A Flow-Capturing Location-Allocation Model. *Geograph. Anal.* **22**(3), 270–279.

MIRCHANDANI, P. B., AND R. L. FRANCIS. 1990. *Discrete Location Theory*. John Wiley, New York.

NEMHAUSER, G. L., AND L. A. WOLSEY. 1978. Best Algorithms for Approximating the Maximum of a Submodular Set Function. *Math. Opns. Res.* **3,** 177–188.

NEMHAUSER, G. L., AND L. A. WOLSEY. 1981. Maximizing Submodular Set Functions: Formulations and Analysis of Algorithms. *Anns. Discr. Math.* **11,** 279–301.

NEMHAUSER, G. L., AND L. A. WOLSEY. 1988. *Integer and Combinatorial Optimization*. John Wiley, New York.

NEMHAUSER, G. L., L. A. WOLSEY AND M. L. FISHER. 1978. An Analysis of Approximations for Maximizing Submodular Set Functions, I. *Math. Prog.* **14,** 265–294.

WOLSEY, L. A. 1982. Maximizing Real-Valued Submodular Functions: Primal and Dual Heuristics for Location Problems. *Math. Opns. Res.* **7,** 410–425.

WOLSEY, L. A. 1983. Fundamental Properties of Certain Discrete Location Problems. In *Locational Analysis*, J. F. Thisse and H. G. Zoller (eds.). North-Holland, Amsterdam, 331–355.