A Dual-Based Procedure for Uncapacitated Facility Location
Author(s): Donald Erlenkotter
Reviewed work(s):
Source: *Operations Research*, Vol. 26, No. 6 (Nov. - Dec., 1978), pp. 992-1009
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/170260
Accessed: 11/04/2012 17:48

OPERATIONS RESEARCH
Vol. 26, No. 6, November-December 1978

0030-364X/78/2606-0992 $01.25

A Dual-Based Procedure for Uncapacitated Facility Location

DONALD ERLENKOTTER

University of California, Los Angeles, California

(Received October 1976; accepted January 1978)

We develop and test a method for the uncapacitated facility location problem that is based on a linear programming dual formation. A simple ascent and adjustment procedure frequently produces optimal dual solutions, which in turn often correspond directly to optimal integer primal solutions. If not, a branch-and-bound procedure completes the solution process. This approach has obtained and verified optimal solutions to all the Kuehn-Hamburger location problems in well under 0.1 seconds each on an IBM 360/91 computer, with no branching required. Computational tests on problems with as many as 100 potential facility locations provide evidence that this approach is superior to several other methods.

IN THE uncapacitated, or simple, facility location problem, facilities of unrestricted size are placed among m possible sites with the objective of minimizing the total cost for satisfying fixed demands specified at n locations. Costs include a fixed charge for opening each facility and a constant amount for each unit of location j's demand supplied from facility i. Although this problem description excludes many of the aspects of complex distribution systems outlined by Geoffrion [10], through simple reformulations it is directly applicable to more general problems with concave, piecewise-linear facility costs [6] and price-elastic demands [7].

Of the many solution approaches that have been proposed for the uncapacitated location problem, the branch-and-bound methods of Efroymson and Ray [6] and Khumawala [12] and the implicit enumeration method of Spielberg [22] are particularly well-known. Recently Schrage [20] developed a specialized linear programming algorithm for variable upper bound constraints and applied it to several problems, including a "tight" linear programming formulation of the location problem different from the one used in the earlier branch-and-bound methods. Schrage observed that this formulation frequently yields natural integer solutions, as noted also by ReVelle and Swain [18] for a related location formulation. Previously Spielberg [22] used this tight formulation to develop tests in his method, and Davis and Ray [4] successfully employed a similar formulation in a decomposition approach for the more general capacitated location problem. Diehr [5] and Cornuejols, Fisher, and Nemhauser [3] have approached the tight linear programming formulation through a Lagrangian dual in order to construct bounds for checking primal solutions obtained by heuristic methods. Cornuejols, Fisher, and Nemhauser (CFN) applied a subgradient optimization method to improve the dual solution and generate new primal solutions, and successfully obtained and verified optimal solutions to most of the problems they attempted.

We develop a dual-based solution procedure for the uncapacitated facility location problem that also begins with the tight linear programming formulation but differs in several ways from these earlier approaches. Solving a condensed linear programming dual through simple ascent and adjustment methods produces monotone improvement of the dual objective, a property not guaranteed by the CFN approach. Integer primal solutions are generated entirely as by-products of the dual solutions. These simple methods often yield an optimal solution directly. If not, they provide bounds for a branch-and-bound procedure that completes the search for an optimum. We present comparative computational results for several problems, including those of Kuehn and Hamburger [16]. Although such results cannot resolve the issues of computational complexity raised by CFN [3] for this location problem, this procedure is superior to other approaches in finding solutions for these particular problems.

After this paper was submitted for publication, I learned of closely related work by Bilde and Krarup [2]. Their paper, although published recently, is essentially a translation into English of a report originally prepared in Danish in 1967. From a different perspective, Bilde and Krarup develop a procedure essentially the same as the ascent procedure given here and incorporate it into a branch-and-bound procedure. They do not, however, consider the adjustment approach and other improvements developed here or provide explicit computational comparisons with other approaches.

1. MODEL FORMULATION AND SOLUTION PROPERTIES

To formulate a model for the uncapacitated facility location problem, we define the following notation: x_{ij} is the fraction of location $j \in J$'s demand supplied from facility $i \in I$; y_i is 1 if facility i is established and 0 otherwise; c_{ij} is the total of variable capacity, production, and distribution costs for supplying all of location j's demand from facility i; and $f_i \geqq 0$ is the fixed cost for establishing facility i. The model formulation is:

$$\min z_P = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{i \in I} f_i y_i \tag{1}$$

$$\sum_{i \in I} x_{ij} = 1, \qquad j \in J \tag{2}$$

$$y_i - x_{ij} \geqq 0, \qquad i \in I, j \in J \tag{3}$$

$$x_{ij} \geqq 0, \qquad i \in I, j \in J \tag{4}$$

$$y_i \in \{0,1\} \qquad i \in I. \tag{5}$$

In an attempt to obtain a natural integer solution, we shall solve the linear programming relaxation that replaces (5) by

$$y_i \geqq 0, i \in I. \tag{6}$$

As in [3, 4, 20, and 22], we use a strong, or "tight," formulation rather than the weaker one that substitutes the constraints

$$\textstyle\sum_{j \in J} x_{ij} \leqq n y_i, i \in I \tag{7}$$

for (3). The branch-and-bound approaches of [6] and [12] adopt this weaker formulation.

The dual problem for (1)-(4) and (6) is

$$\max z_D = \textstyle\sum_{j \in J} v_j \tag{8}$$

$$\textstyle\sum_{j \in J} w_{ij} \leqq f_i, \qquad i \in I \tag{9}$$

$$v_j - w_{ij} \leqq c_{ij}, \qquad i \in I, j \in J \tag{10}$$

$$w_{ij} \geqq 0, \qquad i \in I, j \in J. \tag{11}$$

For any feasible choice of the dual variables v_j, setting each variable w_{ij} at the lowest value possible, as in [22] will maintain feasibility and leave the value of the objective unchanged. Thus we assume

$$w_{ij} = \max\{0, v_j - c_{ij}\}. \tag{12}$$

Substituting (12) into (9) replaces (8)-(11) with the following condensed dual that involves only the explicit variables v_j:

$$\max z_D = \textstyle\sum_{j \in J} v_j \tag{13}$$

$$\textstyle\sum_{j \in J} \max\{0, v_j - c_{ij}\} \leqq f_i, i \in I. \tag{14}$$

We have not included in our formulation a restriction on the number of facilities $\sum_{i \in I} y_i \leqq p$ characteristic of the "p-median" location problem [3, 5, 18, 20], and our problem therefore is less general than those addressed by CFN [3] and Schrage [20]. This restriction would complicate the solution of the dual (13), (14) since its consequence is to introduce an implicit fixed charge for each facility as a variable in the dual. However, most applications would probably require solutions for several values of p to examine cost tradeoffs and the implicit facility fixed charge corresponding to the restriction. A similar analysis can be performed with the formulation here by obtaining solutions for various levels of surcharge added to the explicit fixed charges f_i, as in Swain [23]. Of course, as Swain points out, this approach may not produce solutions for all values of p.

To derive feasible primal solutions from dual solutions, we shall use these complementary slackness relationships for the optimal linear programming solutions:

$$y_i^*[f_i - \sum_{j\in J}\max\{0, v_j^* - c_{ij}\}] = 0 \tag{15}$$

$$[y_i^* - x_{ij}^*]\max\{0, v_j^* - c_{ij}\} = 0. \tag{16}$$

Suppose we have a feasible dual solution $\{v_j^+\}$ and a corresponding set of facility locations I^+ such that $\sum_{j\in J}\max\{0, v_j^+ - c_{ij}\} = f_i$ for each $i\in I^+$ and, for each $j, v_j^+ \geqq c_{ij}$ for some $i\in I^+$. We develop a procedure in the next section that yields such solutions. For each demand location j, define a minimum-cost source facility $i^+(j)\in I^+$ with supply cost

$$c_j^+ = c_{i^+(j),j} = \min_{i\in I^+} c_{ij}, j\in J. \tag{17}$$

The following integer primal solution satisfies the relationships (2)-(5) and (15), although it may violate the complementary slackness condition (16):

$$\begin{aligned} y_i^+ &= \begin{cases} 1, i\in I^+ \\ 0, i\notin I^+ \end{cases} \\ x_{ij}^+ &= \begin{cases} 1, i = i^+(j), j\in J \\ 0 \text{ otherwise.} \end{cases} \end{aligned} \tag{18}$$

A violation of (16) occurs whenever, for some j, more than one $i\in I^+$ has $c_{ij} < v_j$ since $x_{ij}^+ = y_i^+ = 1$ for the lowest value of c_{ij} only.

Linear programming theory provides a simple relationship between these complementary slackness violations and the difference between the dual objective value z_D^+ for $\{v_j^+\}$ and the primal objective value z_P^+ corresponding to the integer solution (18) (see [21, p. 99]). We state this result as a lemma and show it directly.

LEMMA. $z_P^+ - z_D^+ = \sum_{j\in J}\sum_{i\in I^+, i\neq i^+(j)}\max\{0, v_j^+ - c_{ij}\}$.

Proof.

$$\begin{aligned} z_D^+ &= z_D^+ + \sum_{i\in I^+}[f_i - \sum_{j\in J}\max\{0, v_j^+ - c_{ij}\}] \\ &= \sum_{j\in J} v_j^+ + \sum_{i\in I^+} f_i + \sum_{j\in J}(c_j^+ - v_j^+) \\ &\quad - \sum_{j\in J}\sum_{i\in I^+, i\neq i^+(j)}\max\{0, v_j^+ - c_{ij}\} \\ &= z_P^+ - \sum_{j\in J}\sum_{i\in I^+, i\neq i^+(j)}\max\{0, v_j^+ - c_{ij}\}. \end{aligned}$$

Clearly, an integer primal solution (18) that exhibits no complementary slackness violations is optimal. Our solution approach will attempt to close the gap between primal and dual solutions by reducing these violations.

A set I^+ may exclude some facility locations i in the eligible set $I^* = \{i: \sum_{j \in J} \max\ \{0, v_j^+ - c_{ij}\} = f_i\}$. We require only for each j that $v_j^+ \geqq c_{ij}$ for some $i \in I^+$. Adding an additional facility i' to a minimal set I^+ that satisfies this condition cannot improve the solution. From the lemma the inclusion of i' changes the primal objective value z_P by $\sum_{j \in J} \max\ \{0, v_j^+ - \max\ \{c_j^+, c_{i'j}\}\} \geqq 0$. Conversely, deletion from a set I^+ of a *non-essential* facility i' that is not required for such a minimal set cannot make the solution worse.

If more than one facility can be designated as non-essential, the construction of a best minimal set I^+ is a combinatorial problem in itself. Since the solution approach does not require a best minimal set, we use a rapidly obtained approximation to one. First we include in I^+ all *essential* facilities, i.e., any $i \in I^*$ for which only a single $c_{ij} \leqq v_j^+$ for $i \in I^*$ and some j. We then examine sequentially all j for which there is no $i \in I^+$ with $c_{ij} \leqq v_j^+$ and augment I^+ by the eligible facility $i \in I^*$ that has minimum c_{ij}.

2. THE DUAL SOLUTION PROCEDURE

One might conceivably devise a specialized simplex-type method for solving the condensed dual formulation (13), (14), as Schrage [20] has done for the primal problem through special treatment of the "variable upper bound" constraints (3). But two aspects of the dual suggest that a simpler approach might be possible and desirable. First, the dual problem has a very simple structure and typically has multiple solutions (primal solutions are usually quite degenerate). Second, since an exact solution to the relaxed linear problem does not always yield an optimal integer primal solution, we should not be obsessed with obtaining an exact solution. Resolution of these non-integer solutions seems best accomplished through a branch-and-bound phase. A simple approach that provides good bounds from suboptimal dual solutions will be adequate.

The first component of our solution approach is a dual ascent procedure that constructs a dual solution $\{v_j^+\}$ and an associated set of facility locations I^+, with the properties described earlier. This procedure begins with any dual-feasible solution $\{v_j\}$ and repeatedly cycles through the demand locations j one by one, attempting to increase v_j to the next higher value of c_{ij}. This incremental approach to increasing the v_j appears to reduce the likelihood of complementary slackness violations by initially distributing the number of $c_{ij} \leqq v_j$ evenly among the demand locations j. If some constraint (14) blocks the increase of v_j to the next higher c_{ij}, v_j is increased to the maximal level allowed by the constraint. When all the v_j are blocked from further increases, the procedure terminates. A simple example in the next section illustrates this procedure.

We will specify the procedure formally. For notational and computational convenience, we reindex the c_{ij} for each j in non-decreasing order as c_j^k, $k=1,\cdots,m$, and include a high-cost dummy source with $c_j^{m+1}=+\infty$. To obtain an initial feasible dual solution, set $v_j=c_j^1$ for each j. For use in the subsequent adjustment procedure, we restrict changes in the v_j to a subset of demand locations $J^+\subseteq J$. If all v_j are eligible for change, as in the initial application of the procedure, we set $J^+=J$. We now give the

Dual Ascent Procedure

1. Initialize with any feasible dual solution $\{v_j\}$ such that $v_j\geqq c_j^1$ for each $j\in J$ and $s_i=f_i-\sum_{j\in J}\max\{0,v_j-c_{ij}\}\geqq 0$ for each $i\in I$. For each $j\in J$, define $k(j)=\min\{k: v_j\leqq c_j^k\}$. If $v_j=c_j^{k(j)}$, increase $k(j)$ by 1.
2. Initialize $j=1$ and $\delta=0$.
3. If $j\notin J^+$, go to Step 7.
4. Set $\Delta_j=\min_{i\in I}\{s_i: v_j-c_{ij}\geq 0\}$.
5. If $\Delta_j>c_j^{k(j)}-v_j$, set $\Delta_j=c_j^{k(j)}-v_j$ and $\delta=1$, and increase $k(j)$ by 1.
6. Decrease s_i by Δ_j for each $i\in I$ with $v_j-c_{ij}\geqq 0$; then increase v_j by Δ_j.
7. If $j\neq n$, increase j by 1 and return to Step 3.
8. If $\delta=1$, return to Step 2. Otherwise, terminate.

Since each v_j is increased until blocked by an equality in some constraint (14), the final solution from this procedure with $J^+=J$ satisfies the conditions required for a solution $\{v_j^+\}$ and facility set $I^+\subseteq I^*=\{i: s_i=0\}$. As discussed in the preceding section, non-essential facilities can be excluded from I^+ to improve the corresponding primal solution (18).

As described, the dual ascent procedure arbitrarily processes demand locations according to ascending order. We have experimented with two other processing strategies: descending order and alteration between ascending and descending order at each passage through Step 2 of the procedure. Different strategies can lead to somewhat different solutions, and none has been uniformly superior. The alternating strategy has been the most consistent and was used for all the computational results reported here.

The dual ascent procedure yields a candidate dual solution and a candidate integer primal solution through (18). If these solutions satisfy all the complementary slackness conditions (16), the solutions are optimal. If not, we try to improve the solutions through a dual adjustment procedure. To initiate this adjustment procedure, we select some j' for which (16) is violated. Suppose we decrease $v_{j'}$. This creates slack on at least two binding constraints (14), corresponding to those $i\in I^+$ with $v_{j'}>c_{ij'}$. We then attempt to increase other v_j that are limited by these constraints. If more than one v_j can be increased unit for unit as $v_{j'}$ is decreased, the dual objective value will be increased by this change. If only one v_j can be increased, the dual objective value will remain at the same level, but the slack created on some constraint (14) will alter the primal solution since I^+ will be changed. We cycle through this process for all locations j' and may repeat the procedure as long as the dual objective continues to improve. Although more complex adjustment rules can be devised, we have not implemented them since this simple procedure has been very effective.

These changes in the dual solution are accomplished conveniently with the dual ascent procedure. We decrease $v_{j'}$ to the next lower value $c_{j'}^k$ and then apply the dual ascent procedure to those v_j identified as likely candidates for increase. A final pass of the dual ascent procedure with $J^+=J$ completes the adjusted solution.

To specify this procedure formally, we define the following additional notation: $I_j^*=\{i\in I^*: v_j\geqq c_{ij}\}$ and $I_j^+=\{i\in I^+: v_j>c_{ij}\}$ for $j\in J$, and $J_i^+=\{j: I_j^*=\{i\}\}$ for $i\in I$. Supplementing the definition of $i^+(j)$ in (17), we define a second-best source $i'(j)\in I^+$ with

$$c_{i'(j),j}=\min_{i\in I^+, i\neq i^+(j)} c_{ij} \text{ for each } j\in J \text{ with } |I_j^+|>1,$$

and $c_j^-=\max_{i\in I}\{c_{ij}: v_j>c_{ij}\}$ for $j\in J$. If $|I_j^+|>1$, we have a violation of the complementary slackness condition (16); if $|I_j^+|\leqq 1$ for each $j\in J$, the solution (18) corresponding to I^+ is optimal. If $|I_j^*|=1$, a single constraint (14) for $i^+(j)$ blocks v_j from increasing, and therefore v_j is a candidate for increase if some $v_{j'}$ contributing to the left-hand side of this constraint is decreased. We now give the

Dual Adjustment Procedure

1. Initialize $j=1$.
2. If $|I_j^+|\leqq 1$, go to Step 7.
3. If $J^+_{i^+(j)}=\phi$ and $J^+_{i'(j)}=\phi$, go to Step 7.
4. For each $i\in I$ with $v_j>c_{ij}$, increase s_i by $v_j-c_j^-$, then decrease v_j to c_j^-.
5. (a) Set $J^+=J^+_{i^+(j)}\cup J^+_{i'(j)}$, and execute the dual ascent procedure.
   (b) Augment J^+ by J and repeat the dual ascent procedure.
   (c) Set $J^+=J$ and repeat the dual ascent procedure.
6. If v_j has not resumed its original value, return to Step 2.
7. If $j\neq n$, increase j by 1 and return to Step 2. Otherwise, terminate.

Since we are making a discrete reduction in v_j, the purpose of Step 5(b) is to increase v_j to take up any of this decrease that has not been absorbed by increases in the other dual variables. Each unit of decrease for v_j in the final solution is matched by a unit increase in at least one other dual variable, and the value of the dual objective therefore cannot be decreased by this procedure. We execute the dual ascent procedure in Step 5(c) with $J^+=J$ to ensure termination with a valid solution $\{v_j^+\}$. If the dual adjustment procedure increases the value of the dual objective, repetition may give further improvement.

If the dual ascent and adjustment procedures do not yield an optimal integer solution to (1)–(5), we complete the search with a branch-and-bound phase that uses as bounds the solutions provided by these procedures. The branch-and-bound approach is the simplest possible—the bounds have been so effective that a more elaborate version seems unnecessary. Salient features of this branch-and-bound phase are:

(a) For branching, we select some facility location contributing to the violation of a complementary slackness condition (16). No attempt is made to discriminate among possible choices—at the first violation encountered, we branch on the corresponding lowest-cost source $i^+(j)$.

(b) For ease in updating the solution and restarting when backtracking, initially we always fix the branching facility closed.

(c) As described by Geoffrion [9], an elementary backtracking scheme with last-in, first-out processing of nodes minimizes computer storage requirements and simplifies updating of solutions.

(d) To fix facility i closed, we replace the fixed charge f_i by $+\infty$. The current dual solution remains feasible. Application of the dual ascent and adjustment procedures may improve the bound given by the value of the dual objective through increasing those v_j no longer restricted by the dual constraint (14) for i.

(e) To fix facility i open, we replace the fixed charge f_i by 0 and restore dual feasibility by reducing the value of each $v_j > c_{ij}$ to c_{ij}. Adding to (13) the fixed charges f_i for all facilities fixed open gives the value of the dual objective for the restricted problem. Since $\sum_{j \in J} \max \{0, v_j - c_{ij}\} = f_i$ for $i \in I^+$, reduction of those $v_j > c_{ij}$ does not change the dual objective value. Reducing these v_j is likely to add slack to other dual constraints (14), and applying the dual ascent and adjustment procedures may increase the value of the dual objective.

(f) Fathoming of nodes in the branch-and-bound phase is either by bounding or by obtaining a primal integer soution (18) that also satisfies the complementary slackness conditions (16).

(g) At each node we construct primal integer solutions (18) only after the initial application of the dual ascent procedure and after each completion of the dual adjustment procedure, and not each time the ascent procedure is completed with $J^+ = J$.

We have investigated four levels of dual improvement effort. All levels first apply the dual ascent procedure at each node. The first level (*no dual improvement*) excludes the dual adjustment phase entirely. The second level (*one-pass dual improvement*) applies the dual adjustment procedure once at each node with repetition only when the primal solution improves. The fourth level (*maximum dual improvement*) repeats the dual adjustment procedure at each node as long as the dual objective value continues to increase. The third level (*maximum/one-pass dual improvement*) is a compromise between the second and fourth levels that obtains maximum dual improvement at the initial node and one-pass improvement at subsequent nodes. The rationale for this level is that the tightest possible bound for the base solution at the initial node is valuable as a starting solution, but extensive effort at subsequent nodes may yield only marginal improvement. Higher levels of dual improvement typically lead to less branching. Although further dual improvement might be possible through subgradient optimization [3], we have not explored this option since the third and fourth levels of improvement have provided optimal linear programming solutions for most test problems without branching. In presenting our computational results, we shall focus primarily on those for the "maximum/one-pass" strategy.

3. ILLUSTRATIVE EXAMPLES

We shall demonstrate the procedure by solving two versions of an example presented by Khumawala [12]. This example has five potential facility locations and eight demand locations. Table I gives the total demand costs c_{ij}.

TABLE I
DATA FOR EXAMPLE PROBLEM
Total demand cost c_{ij}

i \ j	1	2	3	4	5	6	7	8
1	120	180	100	$+\infty$	60	$+\infty$	180	$+\infty$
2	210	$+\infty$	150	240	55	210	110	165
3	180	190	110	195	50	$+\infty$	$+\infty$	195
4	210	190	150	180	65	120	160	120
5	170	150	110	150	70	195	200	$+\infty$

Khumawala's example has the facility fixed-charge vector $(f_i)=(100, 70, 60, 110, 80)$. As we shall see, the dual ascent procedure solves this problem immediately.

Initializing the dual variables v_j at the values c_j^1 provides the vector of dual variables $(v_j^0)=(120, 150, 100, 150, 50, 120, 110, 120)$, with $\sum v_j^0=920$. The corresponding initial vector of dual slack variables is $(s_i^0)=(f_i)=(100, 70, 60, 110, 80)$.

We now apply the dual ascent procedure to the demand locations in ascending order. At the end of the first cycle, we have $(v_j^1)=(170, 180, 110, 180, 55, 195, 160, 155)$ with $\sum v_j^1=1205$, and $(s_i^1)=(40, 20, 55, 0, 20)$. All dual variables v_j^1 have been increased to the values c_j^2 except $v_8^1=155$, which has been blocked by $s_4^1=0$.

At the end of the second cycle, we have $(v_j^2)=(180, 190, 110, 180, 60, 195, 160, 155)$ with $\sum v_j^2=1230$, and $(s_i^2)=(20, 15, 50, 0, 0)$. Now all dual variables v_j^2 except v_5^2 are blocked from further increase by the zero slacks s_4^2 and s_5^2. One last cycle of the dual ascent procedure completes the solution: $(v_j^3)=(180, 190, 110, 180, 65, 195, 160, 155)$ with $\sum v_j^3=1235$, and $(s_i^3)=(15, 10, 45, 0, 0)$.

The primal solution (18) for $I^+=\{4, 5\}$ yields $y_4=y_5=1$ and $x_{51}=x_{52}=x_{53}=x_{54}=x_{45}=x_{46}=x_{47}=x_{48}=1$. There are no complementary slackness violations, and the primal objective value is 1235, the same as for the dual. Therefore, the primal solution is optimal. We note that the dual solution is not unique: there are many alternate optima.

For a more challenging example, change the fixed-charge vector to $(f_i)=(200, 200, 200, 400, 300)$. We shall sketch the solution process without giving complete details. At the termination of the dual ascent procedure, we have the dual solution $(v_j^1)=(210, 190, 150, 240, 65, 285, 195, 195)$, with $\sum v_j^1=1530$, and associated slack vector $(s_i^1)=(30, 0, 70, 65, 0)$. For $I^+=\{2, 5\}$, the primal objective value is 1605. Since $v_6^1>c_{26}>c_{56}$, the primal solution violates a complementary slackness condition (16).

We initiate the dual adjustment procedure and find $|I_6^+|=2$. Since $J_2^+=\{5, 7, 8\}$ and $J_5^+=\{2\}$, we increase s_2^1, s_4^1, and s_5^1 by 75 and reduce v_6^1 to 210. The dual adjustment procedure yields the improved solution $(v_j^2)=(210, 220, 150, 240, 65, 245, 195, 235)$, with $\sum v_j^2=1560$, and slack vector $(s_i^2)=(0, 0, 0, 35, 10)$ and then terminates. Facility $i=2$ is essential, and we add $i=1$ to obtain $I^+=\{1, 2\}$ with a primal objective value of 1580.

Applying the dual adjustment procedure again, we find $v_5^2>c_{15}>c_{25}, |I_5^+|=2$, and $J_2^+=\{6\}$. We increase s_1^2, s_2^2, and s_3^2 by 5 and reduce v_5^2 to 60. The procedure then provides the solution $(v_j^3)=(210, 225, 150, 240, 60, 250, 195, 235)$, with $\sum v_j^3=1565$, and $(s_i^3)=(0, 0, 0, 25, 0)$. This is an optimal dual solution, and no further improvement is possible.

As the reader may verify, the optimal relaxed primal solution for this problem is fractional, with $y_2=⅔$ and $y_1=y_3=y_5=⅓$. Except for $x_{24}=⅓$, all nonzero primal variables x_{ij} equal the corresponding y_i values as indicated by the complementary slackness conditions (16).

To complete the solution, we initiate the branch-and-bound phase with $I^+=\{1, 2\}$, deleting non-essential facilities to obtain the best integer primal solution found thus far. Since $v_7^3>c_{17}>c_{27}$, initially we fix y_2 closed by increasing s_2^3 to $+\infty$. The dual ascent phase does not change the dual solution (v_j^3), but now $I^+=\{1, 3, 5\}$. Applying the dual adjustment procedure, we have $v_1^3>c_{31}>c_{51}>c_{11}, |I_1^+|=3$, and $J_1^+=\{7\}, J_5^+=\{6\}$. After increasing s_1^3, s_3^3, and s_5^3 by 30 and reducing v_1^3 to 180, the procedure gives the dual solution $(v_j^4)=(185, 225, 150, 240, 65, 275, 195, 235)$, with $\sum v_j^4=1570$, and $(s_i^4)=(25, +\infty, 20, 0, 0)$. Next, with $I^+=\{4, 5\}$, we have $v_2^4>c_{42}>c_{52}$, and $J_4^+=\{5, 7, 8\}$, $J_5^+=\{1\}$. Increasing s_1^4, s_3^4, s_4^4, and s_5^4 by 35 and reducing v_2^4 to 190, we obtain $(v_j^5)=(210, 190, 150, 240, 70, 275, 200, 260)$, with $\sum v_j^5=1595$. Since this exceeds the value of 1580 for the best-known integer primal solution, the node is fathomed.

We now fix y_2 open by reducing all $v_j>c_{2j}$ to c_{2j} and setting $s_2=0$, yielding $(v_j^6)=(210, 190, 150, 240, 55, 210, 110, 165)$, and $(s_i^6)=(50, 0, 80, 205, 75)$.

Applying the dual ascent procedure gives $(v_j^7)=(210, 240, 150, 240, 55, 210, 110, 165)$, with $\sum v_j^7=1380$. Adding the fixed charge $f_2=200$ gives a bound of 1580, and this node is fathomed. The set of open facilities in the optimal integer primal solution is thus $I^+=\{1, 2\}$.

4. COMPUTATIONAL RESULTS

We have implemented the dual-based solution approach in a FORTRAN IV computer program called DUALOC. (A listing of the DUALOC program is available from me upon request.) For comparison with results from other approaches, including those of [3, 12, 20] we have conducted computational experiments with DUALOC on several problems with dimensions $(m \times n)$ of (24×50), (49×200), (33×33), (57×57), and (100×100). Except where noted, all results are for the maximum/one-pass dual improvement option.

The first set of problems contains the 16 variants of the Kuehn-Hamburger problems [16] considered by Khumawala [12] and Sá [19]. Each of these problems has 24 potential warehouse locations and 50 demand locations, not including those warehouses available for zero fixed cost at factories. The first four have a factory and warehouse at Indianapolis, the second four at Jacksonville, the third four at both Baltimore and Indianapolis, and the last four a factory, but no warehouse, at Indianapolis. Table II gives the results for these problems. The initial dual and primal solutions are those obtained at the end of the first application of the dual ascent procedure, and the number of dual solutions is the number of times the dual ascent procedure completed a solution $\{v_j^+\}$. No problem required more than a single node; there was no branching. In seven of the problems, the initial application of the dual ascent procedure produced and verified the optimal solution immediately.

Table II includes comparative data on solution times for DUALOC and Khumawala's branch-and-bound procedure [12]. Even after allowing for the slower computer used by Khumawala, DUALOC seems to be faster for these problems by an order of magnitude or more. Schrage [20] applied his variable upper-bound linear programming code to these problems and reported computational times of 2–3 seconds on an IBM 370/168, a computer at least equal in speed to the IBM 360/91 used by DUALOC. Again DUALOC appears to be more than an order of magnitude faster.

We also have solved Khumawala's version [13] of the Feldman, Lehrer, and Ray (49×200) warehouse location problem [8]. There are two cases for this problem. The first, with a continuous, two-segment warehousing cost function (curve "A"), required 0.269 seconds computational time on the IBM 360/91 and two dual solutions. The second, with a typical fixed-charge warehousing cost function (curve "B"), took 0.376 seconds and nine dual solutions. Neither problem required branching. Khumawala

TABLE II
COMPUTATIONAL RESULTS FOR KUEHN-HAMBURGER PROBLEMS

Problem No.	Fixed charge, f_i	Objective value			No. of facilities opened	No. of nodes evaluated	No. of dual solutions	Computational time (sec)	
		Initial dual	Initial primal	Optimal*				DUALOC†	Khumawala‡
1	7,500	796,648	796,648	796,648	15	1	1	0.032	1.61
2	12,500	853,264	854,704	854,704	11	1	7	0.073	3.43
3	17,500	892,064	897,833	893,782	8	1	3	0.070	2.61
4	25,000	926,216	936,690	928,942	4	1	2	0.065	1.56
5	7,500	1,092,916	1,092,916	1,092,916	14	1	1	0.032	0.88
6	12,500	1,145,923	1,145,923	1,145,923	10	1	1	0.032	4.37
7	17,500	1,187,632	1,188,241	1,188,241	9	1	3	0.045	5.86
8	25,000	1,244,874	1,256,377	1,244,991	8	1	2	0.064	17.38
9	7,500	614,548	614,548	614,548	13	1	1	0.034	1.50
10	12,500	658,543	659,983	659,983	9	1	7	0.072	1.36
11	17,500	690,735	693,530	690,746	8	1	2	0.066	1.66
12	25,000	724,887	724,887	724,887	5	1	1	0.035	0.85
13	7,500	806,145	806,145	806,145	15	1	1	0.031	1.55
14	12,500	870,792	870,792	870,792	10	1	1	0.033	3.87
15	17,500	917,961	921,261	919,993	8	1	3	0.069	6.39
16	25,000	967,395	978,195	970,446	6	1	3	0.068	6.80

* Values do not exactly match those reported in [13] because of slight differences in rounding.
† On IBM 360/91, including output but excluding input time.
‡ On CDC 6500, CPU time as reported in [12]. Computational times on an IBM 360/65 are indicated in [12] to be "about the same."

[13] reports computational times of 2–5 seconds on an IBM 360/75 for *heuristic* solutions to this problem. Hence DUALOC appears to find optimal solutions more rapidly than these other procedures can provide approximate solutions.

The other problems have c_{ij} values taken from data for traveling salesman problems. The (33×33) and (57×57) problems are from [11], and the (100×100) problem is from [15]. Demands are identical at all locations, and the absence of demand concentrations probably makes these problems more difficult than the warehouse location problems.

Tables III, IV, and V give the results for these problems, with ten fixed-charge values for each. The initial dual solution solved six of the (33×33) problems. Only one (with f_i=2000) required branching, and a single branch gave the optimal solution. Schrage [20] has reported an integer solution to the linear programming relaxation for this problem; hence a solution through the dual formulation should be possible without branching. However, the dual ascent and adjustment procedures did not attain the dual optimum for this problem. Branching was unnecessary for all of the (57×57) and all but two of the (100×100) problems. In every case, the initial dual solution from the dual ascent procedure is within 1% of the optimal objective value.

Tables III, IV, and V include comparative data on solution times for DUALOC, Schrage's code, and the CFN dual approach [3]. Comparisons are somewhat imperfect since many of Schrage's and all of CFN's results are for the p-median location problem, which specifies the number of facilities and omits facility fixed charges. However, Schrage has reported similar times for several problems solved either with fixed charges or with the corresponding number of facilities specified. An examination of the structure of the CFN approach suggests that results for fixed-charge problems would be similar to those with the number of facilities specified. It would appear, then, that the comparisons are reasonable.

The specialized DUALOC procedure is at least an order of magnitude faster for these problems than Schrage's more general routine. DUALOC is 10 to 50 times faster than the CFN approach for several problems, and the comparison is close only for two problems: (33×33), with f_i=3000; and (100×100), with f_i=2900. For some problems with more than just a few facilities opened, e.g., (100×100) with f_i=1150 or 2000, DUALOC obtains and verifies the optimal solution in little more time than CFN take to find an initial approximate solution.

DUALOC with the one-pass level of dual improvement typically consumed about the same amount of computational time as with the maximum or maximum/one-pass levels of improvement, but the maximum level more frequently revealed the linear programming optimum at the initial node and was able to terminate without branching. The no-improvement level usually required more time and extensive branching.

TABLE III

COMPUTATIONAL RESULTS FOR (33 × 33) LOCATION PROBLEM

Fixed charge, f_i	Objective value			No. of facilities opened	No. of nodes evaluated	No. of dual solutions	Computational time (sec)		
	Initial dual	Initial primal	Optimal				DUALOC*	Schrage†	CFN‡
184	6024	6024	6024	31 or 32	1	1	0.023	1.15	
295	8673	8673	8673	17 or 18	1	1	0.024	1.39	(0.4)
500	11267	11267	11267	10	1	1	0.022		
1000	14832	14832	14832	6	1	1	0.024		
1500	17815	17832	17832	6	1	4	0.037		
2000	20161	21402	20363	4	3	37	0.187	2.36	(0.4)
2500	22092	23137	22127	3	1	3	0.053		
3000	23465	25810	23474	2	1	2	0.052	(1.39)	(0.1)
4000	25474	25474	25474	2	1	1	0.034	(1.39)	(0.1)
5000	27474	27474	27474	2	1	1	0.034	(1.39)	(0.1)

* On IBM 360/91, including output but excluding input time.

† On IBM 370/168, including input-output time, as reported in [20]. Times in parentheses are for problems with zero fixed charges and fixed number of facilities.

‡ On IBM 370/168, excluding input-output time, as reported in [3]. Times are for problems with zero fixed charges and fixed number of facilities.

TABLE IV

COMPUTATIONAL RESULTS FOR (57 × 57) LOCATION PROBLEM

Fixed charge, f_i	Objective value			No. of facilities opened	No. of nodes evaluated	No. of dual solutions	Computational time (sec)		
	Initial dual	Initial primal	Optimal				DUALOC*	Schrage†	CFN‡
50	2821	2821	2821	55	1	1	0.052	(2.81)	(0.5)
200	9142	9142	9142	29	1	1	0.046	(3.97)	(2.2)
500	15221	15335	15261	13	1	11	0.179		
1000	20150	20651	20307	9	1	14	0.174		
1500	23916	24534	23943	7	1	5	0.105		
2000	27176	28593	27222	6	1	6	0.118		
2500	29953	30022	30022	5	1	5	0.091		
3000	31910	33604	32136	4	1	14	0.168	(10.58)	(1.1)
4000	35462	36452	35547	3	1	10	0.161		
5000	38247	38962	38547	3	1	23	0.227		

* On IBM 360/91, including output but excluding input time.

† On IBM 370/168, including input-output time, as reported in [20]. Times are for problems with zero fixed charges and fixed number of facilities.

‡ On IBM 370/168, excluding input-output time, as reported in [3]. Times are for problems with zero fixed charges and fixed number of facilities.

TABLE V

COMPUTATIONAL RESULTS FOR (100 × 100) LOCATION PROBLEM

Fixed charge, f_i	Objective value			No. of facilities opened	No. of nodes evaluated	No. of dual solutions	Computational time (sec)	
	Initial dual	Initial primal	Optimal*				DUALOC†	CFN‡
1000	35926	37992	35965	17	1	10	0.236	
1150	38314	40168	38479	16	1	18	0.389	(7.2)
2000	49916	50468	50103	12	1	21	0.377	(7.0)
2900	58106	61987	58613	8	13	329	3.383	(4.8)
3000	58837	62887	59407	7	21	279	3.136	
4000	65942	68679	66407	7	1	41	0.606	
5000	72777	76715	73073	6	1	55	0.727	
6000	78159	91442	78720	5	1	71	0.906	
7000	83018	95224	83720	5	1	79	0.986	
8000	87263	97233	87889	4	1	47	0.668	(3.4)

* Values do not exactly match those reported in [3] because of slight differences in rounding.

† On IBM 360/91, including output but excluding input time.

‡ On IBM 370/168, excluding input-output time, as reported in [3]. Times are for problems with zero fixed charges and fixed number of facilities.

However, the two difficult (100×100) problems were notable exceptions. Table VI gives computational results for these problems with the four levels of dual improvement. For a fixed charge of 2900, one-pass improvement took the least time; for a fixed charge of 3000, no dual improvement gave the best performance even with considerable branching. The maximum improvement level required the most time for both problems. The number of primal improvements is the number of times DUALOC provided an improved primal solution after the initial solution. Increasing the dual improvement level reduced the number of intermediate im-

TABLE VI

COMPARISON OF DUAL IMPROVEMENT LEVELS FOR THE (100 × 100) PROBLEM

Fixed charge, f_i	Level of dual improvement	No. of nodes evaluated	No. of dual solutions	No. of primal improvements	Computational time* (sec)
2900	none	307	307	17	4.696
	one-pass	13	172	9	2.196
	maximum/one pass	13	329	5	3.383
	maximum	13	542	3	5.518
3000	none	181	181	10	2.847
	one-pass	21	276	6	3.087
	maximum/one pass	21	279	6	3.136
	maximum	11	412	4	4.183

* On IBM 360/91, including output but excluding input time.

proved primal solutions generated. Both of these problems have many near-optimal integer solutions—for the case with f_i=2900, DUALOC detected 75 solutions with objective values within 1% of the optimal value.

The solution times for DUALOC exclude the time for sorting the c_{ij} into non-decreasing order for each j, a data-preparation task that is performed only once for each set of c_{ij}. Since most applications would require several runs with this data, assigning this time to a single run is inappropriate. In our computational experiments we have not exploited a possible efficiency for multiple runs of the same problem with different fixed-charge sets: if none of the fixed charges is decreased, the previous initial dual solution provides a feasible starting solution for the new problem.

5. CONCLUSIONS

We have seen how simple ascent and adjustment procedures applied to a condensed dual formulation and coupled with a branch-and-bound phase where necessary can solve the uncapacitated facility location problem. The ease of obtaining integer primal solutions from the dual solutions makes the dual-based approach an attractive one. Since many problems have multiple optimal dual solutions, an optimal dual solution often is not difficult to find. The computational results suggest that this solution approach may be close to the ultimate in efficiency for the problems solved.

The results reported here also have some interesting implications for the theory of resource allocation. For a different but related problem, Koopmans and Beckmann [14] have drawn pessimistic conclusions about the existence of a price system that would lead to optimal location decisions. Manne [17] has given a multiple-price system for the problem we examine, but his prices do not guarantee an optimal solution. For 46 of the 48 large problems solved here, we know that solution of a linear programming problem gives an optimal integer solution. Therefore, the dual solution provides a set of prices that will lead to an optimal primal integer solution. Although such a solution does not always exist, as in our second example problem and the one given by Balinski [1], this seems to be more the exception than the rule for the problems we have investigated.

ACKNOWLEDGMENTS

I wish to thank Basheer Khumawala for providing data for several of the problems and Sanjay Deshpande for assistance in preparing the data.

REFERENCES

1. M. L. BALINSKI, "Integer Programming: Methods, Uses, Computation," *Management Sci.* **12**, 253–313 (1965).

2. O. BILDE AND J. KRARUP, "Sharp Lower Bounds and Efficient Algorithms for the Simple Plant Location Problem," *Ann. Discrete Math.* **1,** 79–88 (1977).
3. G. CORNUEJOLS, M. L. FISHER, AND G. L. NEMHAUSER, "Location of Bank Accounts to Optimize Float: An Analytic Study of Exact and Approximate Algorithms," *Management Sci.* **23,** 789–810 (1977).
4. P. S. DAVIS AND T. L. RAY, "A Branch-Bound Algorithm for the Capacitated Facilities Location Problem," *Naval Res. Logist. Quart.* **16,** 331–344 (1969).
5. G. DIEHR, "An Algorithm for the p-Median Problem," Working Paper No. 191, Western Management Science Institute, University of California, Los Angeles, June 1972.
6. M. A. EFROYMSON AND T. L. RAY, "A Branch-Bound Algorithm for Plant Location," *Opns. Res.* **14,** 361–368 (1966).
7. D. ERLENKOTTER, "Facility Location with Price-Sensitive Demands: Private, Public, and Quasi-Public," *Management Sci.* **24,** 378–386 (1977).
8. E. FELDMAN, F. A. LEHRER, AND T. L. RAY, "Warehouse Location under Continuous Economies of Scale," *Management Sci.* **12,** 670–684 (1966).
9. A. M. GEOFFRION, "Integer Programming by Implicit Enumeration and Balas' Method," *SIAM Rev.* **9,** 178–190 (1967).
10. A. M. GEOFFRION, "A Guide to Computer-Assisted Methods for Distribution Systems Planning," *Sloan Management Rev.* **16,** 17–42 (1975).
11. R. L. KARG AND G. L. THOMPSON, "A Heuristic Approach to Solving Travelling Salesman Problems," *Management Sci.* **10,** 225–248 (1964).
12. B. M. KHUMAWALA, "An Efficient Branch and Bound Algorithm for the Warehouse Location Problem," *Management Sci.* **18,** B718–B731 (1972).
13. B. M. KHUMAWALA, "An Efficient Heuristic Procedure for the Uncapacitated Warehouse Location Problem," *Naval Res. Logist. Quart.* **20,** 109–121 (1973).
14. T. C. KOOPMANS AND M. J. BECKMANN, "Assignment Problems and the Location of Economic Activities," *Econometrica* **25,** 53–76 (1957).
15. P. KROLAK, W. FELTS, AND G. MARBLE, "A Man-Machine Approach toward Solving the Traveling Salesman Problem," *Comm. ACM* **14,** 327–334 (1971).
16. A. A. KUEHN AND M. J. HAMBURGER, "A Heuristic Program for Locating Warehouses," *Management Sci.* **9,** 643–666 (1963).
17. A. S. MANNE, "Plant Location Under Economies-of-Scale—Decentralization and Computation," *Management Sci.* **11,** 213–235 (1964).
18. C. S. REVELLE AND R. W. SWAIN, "Central Facilities Location," *Geographical Anal.* **2,** 30–42 (1970).
19. G. SÁ, "Branch-and-Bound and Approximate Solutions to the Capacitated Plant-Location Problem," *Opns. Res.* **17,** 1005–1016 (1969).
20. L. SCHRAGE, "Implicit Representation of Variable Upper Bounds in Linear Programming," *Math. Programming Study* **4,** 118–132 (1975).
21. M. SIMONNARD, *Linear Programming,* Prentice-Hall, Englewood Cliffs, N.J., 1966.
22. K. SPIELBERG, "Algorithms for the Simple Plant-Location Problem with Some Side Conditions," *Opns. Res.* **17,** 85–111 (1969).
23. R. W. SWAIN, "A Parametric Decomposition Approach for the Solution of Uncapacitated Location Problems," *Management Sci.* **21,** 189–198 (1974).