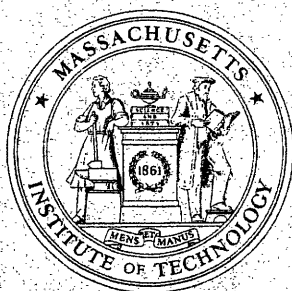# OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

A SURVEY
ON THE WAREHOUSE LOCATION PROBLEM

by

Joseph J. Cohen

OR 022-73   September, 1973

ABSTRACT

The warehouse location problem has assumed numerous formulations, and solutions have been devised using a variety of mathematical techniques. The development of this effort is examined and relevant models presented for evaluation.

I. Introduction

The warehouse location problem has many variations, so we state at the outset those characteristics we would ideally like a location model to possess. We then proceed to examine presently proposed models in relation to our ideal.

A good description of the problem is given by Feldman [7]. He states:

> The warehouse location problem involves the determination of the number and sizes of service centers (warehouses) to supply a set of demand centers. The objective is to locate and size the warehouses and determine which demand centers are supplied from which warehouses so as to minimize total distribution costs. This distribution cost is the total transportation cost, which is assumed linear, plus the cost of building and operating the warehouse.

Atkins and Shriver [1] propose several "ideal analysis requirements desired by management." Using their list as a basis, we will consider as requisite for a location model the following characteristics.

1. No restrictions on the number of possible locations to be considered.

2. Explicit consideration of production and/or storage capacity limitations.

3. Explicit consideration of locational impact on inventory requirements, customer service, and demand patterns.

4. Explicit accounting for possible economies of scale involved in warehouse costs.

5. Finally the model must achieve an optimal location pattern based on the above considerations in the shortest possible time and at minimum cost.

There are clearly wide disparities between our "ideal" model and what treatment methodologies are realistically capable of handling. We will explore these differences to ascertain how good various models are in spite of such differences.

Most location models are of the uncapacitated type, thus violating (2) above. This is a serious assumption to make, but the computation problems involved have often required it. Except where we specifically discuss capacitated models, it is assumed that all models are of the uncapacitated type, unconstrained in production or storage capacity.

The third characteristic mentioned brings to light the crucial interactions involved in a logistics distribution system. The total system nature of the problem needs considerations, for decisions in one part of the distribution system invariably have effects elsewhere. Matthews, et al. [16] describe a firms logistic system, the management of its physical distribution activities, as composed of the following major activities.

1. Warehousing: The managerial problems associated with the storage of goods. The major problem involved in warehousing for the individual firm is determining the number and locations of distribution points.

2. Inventory Control: The determining of the optimal trade-off between inventory levels and customer service.

3. Transportation: Selecting the means of transportation, routing shipments, scheduling and consolidation of shipments, and providng for an information system.

Interactions in such a system essentially imply tradeoffs between the various elements in the system. For example, transportation costs may be cut through use of a less expensive transportation mode, but at the expense of system flexibility and responsiveness to customer requirements, greater inventories, and larger investment requirements. How to weigh these tradeoffs is a crucial question.

In a competitive economy, producers must make their products conveniently available to their customers. This requires inventories at a sufficient number of locations to provide for prompt service. But the decisions regarding inventories and inventory locations simply cannot be made separate from the decisions concerning transportation modes to effect these deliveries. These decisions interact to determine the characteristics the system will ultimately have and are therefore mutually dependent.

Incorporating system interactions into a location model in an explicit form is difficult. One of the criticisms of early models, and of some recent ones as well, is the failure to include any of these interactions.

Finally we would hope that a means to deal with economies of scale in warehousing costs is included. Frequently the assumption is made that there is only a fixed charge associated with opening a warehouse. This is clearly not enough for our needs, for distinct differences are frequently involved in warehousing costs, depending on location.

II. Models and Techniques

We examine models based on the following classifications: (1) heuristic; (2) simulation; and (3) algorithmic (branch and bound, mixed integer programs). First, however, let us define a mathematical model of the problem from which to proceed. It is a general model that will be modified to describe other formulattions. We let

D_k = demand requirements at demand concentration center k

d_{ij} = distance from factory i to warehouse j

d_{jk} = distance from warehouse j to demand point k

s_{ij} = unit shipping costs from factory i to warehouse j

s_{jk} = unit shipping costs from warehouse j to demand point k

F_j = fixed cost associated with opening a warehouse

T_j = throughput of warehouse j
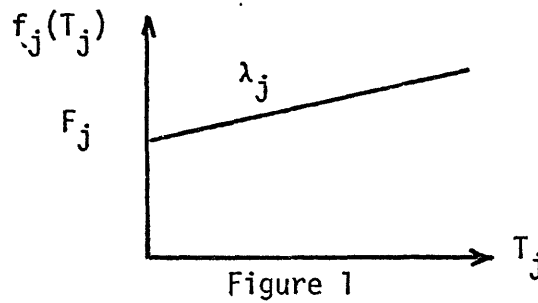
Then for every possible (i,j,k) combination let

$$c_{ijk} = d_{ij}s_{ij} + d_{jk}s_{jk}$$

Due to the uncapacitated assumption described earlier we can let

$$c_{jk} = \min_i (d_{ij}s_{ij}) + d_{jk}s_{jk}$$

since flow of goods will always be along the least cost route.

If we describe warehousing costs as composed of both a fixed and variable component as in Figure 1, then we can set $\bar{c}_{jk} = c_{jk} + \lambda_j$.

Figure 1

Now the total cost of supplying the demand points becomes

$$z = \sum_j \sum_k \bar{c}_{jk} D_k + \sum_j Y_j F_j$$

where $Y_j = \begin{cases} 0 \text{ if warehouse } j \text{ is closed} \\ 1 \text{ if warehouse } j \text{ is open} \end{cases}$, and we wish to minimize z.

If we further let X_{jk} = the fraction of D_k supplied from warehouse j, then our objective function becomes

(I)
$$\min z = \sum_j \sum_k \bar{c}_{jk} X_{jk} + \sum_j F_j Y_j$$

$$\text{s.t.} \quad \sum_j X_{jk} = 1$$

$$0 \leq X_{jk} \leq Y_j \leq 1$$

$$Y_j = \begin{cases} 0 \text{ if warehouse } j \text{ is closed} \\ 1 \text{ if warehouse } j \text{ is open} \end{cases}$$

While only of historical interest, one of the earliest models was that proposed by Balinski and Mills in the form of an integer program (see [2]). They proposed an approximating technique for solving problem (I) above. Instead of addressing explicitly the fixed cost and throughput costs of operating a warehouse, they approximated the warehousing cost function by the average unit cost of operating at some high level of throughput. They set $A = \sum_k D_k$ and then let

$$\hat{c}_{jk} = c_{jk} + \frac{\lambda_j A + F_j}{A}$$

so $\hat{c}_{jk} = \bar{c}_{jk} + F_j/A$

Then the formulation becomes a simple transportaton problem.

(II)

$$\min z = \sum_j \sum_k \hat{c}_{jk} X_{jk}$$

$$\text{s.t.} \quad \sum_j X_{jk} = 1$$

$$0 \le X_{jk} \le 1$$

Since this will not satisfy the integer requirements of (I), it will obviously provide a lower bound on the objective function of that problem. They then evaluate the objective function of the original problem with the X_{jk}'s determined from (II) and show the optimal solution to (I) is somewhere between these two values.

Kuehn and Hamburger in a critique of the Balinski-Mills formulation question the closeness to optimality of their solution. Kuehn-Hamburger conclude that (1) the Balinski-Mills model is not well designed to handle decreasing marginal cost functions generally assumed for warehouse locations problems, and (2) the existence of fixed costs tend to increase deviation from optimality by locating more warehouses than necessary.

Prior to the Balinski-Mills model another approximation technique appeared. This was due to Baumol and Wolfe [3]. Their model was the first attempt to treat the non-linearities of the warehousing cost function, while applying linear programming to allocate warehouse territories. They assumed a strictly concave cost function which does not allow a fixed component, as in Figure 2. They approximated this function by piecewise linear approximations to the degree of accuracy desired.

Figure 2
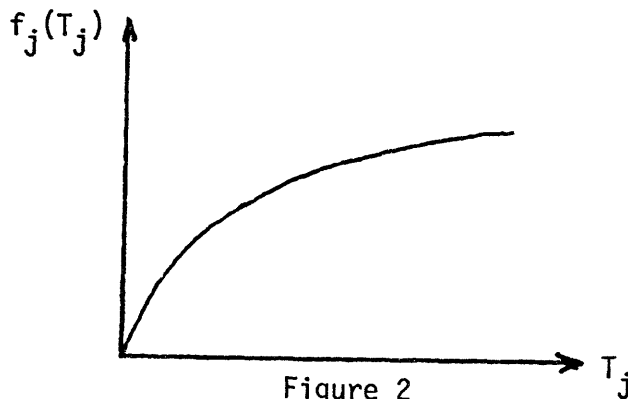
Their algorithm involved an iterative procedure including solution of an ordinary transportation problem at each stage. Again it was shown that their model tended to locate more than the optimal number of warehouses. In spite of this the Baumol-Wolfe model is noteworthy in that it introduced non-linearities into the cost function, recognized the need to supply the model with a list of potential sites, and noted the fallacy of constant demand unaffected by warehouse (i.e., the iteractions described earlier).

Heuristic Techniques

Kuehn and Hamburger [11] prefer to describe heuristic programming an an approach to problem solving where the emphasis is on working towards optimum solution procedures rather than optimum solutions. Heuristic techniques are most often used when the goal is to solve a problem whose solution can be described in terms of acceptability characteristics rather than by optimizing rules. In general, heuristics may be thought of as rules of thumb selected on the basis that they will aid in problem solving.

Heuristics were a logical development to follow the earlier formulations of the warehouse problem. The main problem faced is combinatorial in nature. Excessively long searches and infeasible storage requirements were characteristic of the problems facing model builders. Heuristics were designed to provide a reduction in search requirements. Furthermore, in combinatorial problems of this sort, there is frequently not a sharp global maximum or minimum, so that a "good" solution found by using heuristics is often very close to optimum.

Kuehn and Hamburger were among the first to apply heurstic programming techniques to the warehouse problem. They describe its advantages in the solution of this class of problems as (1) providing considerable flexibility in the specification (modeling) of the problem to be solved, (2) useful to study large scale problems, that is, complexes with several hundred potential warehouse sites and several

thousand shipment destinations and (3) economical in use of computer time.

Their heuristic program is discussed in detail in their paper [11], and else - where. We present their description of the major features in the technique.

> The program consists of two parts: (1) the main program, which locates warehouses one at a time until no additional warehouses can be added to the distribution network without increasing total costs, and (2) the bump and shift routine, entered after processing in the main program is complete, which attempts to modify solutions arrived at in the main program by evaluating the profit implications of dropping individual warehouses or of shifting them from one location to another. The three principal heuristics used in the main program are:
> 1. Most geographical locations are not promising sites for a regional warehouse; locations with promise will be at or near concentrations of demand.
> 2. Near optimum warehousing systems can be developed by locating warehouses one at a time, adding at each stage of the analysis that warehouse which produces the greatest cost savings for the entire system. (The use of this heuristic reduced the time and effort expended in evaluating pattersn of warehouse sites.)
> 3. Only a small subset of all possible warehouse locations need to be evaluated in detail at each stage of the analysis to determine the next warehouse site to be added.

A detailed flow diagram of their program is included in Figure 3.

Taken together the three heuristics provide substantial savings in computation time. Kuehn and Hamburger assert that use of these heuristics make the use of actual cost data computationally efficient, thereby avoiding errors associated with using, for example, air miles as a basis for shipping costs. Solution of large scale problems involving several factories, multiple products, and more than a thousand concentrations of demand now becomes feasible.

Kuehn and Hamburger's formulation has been demonstrated to give solutions close to optimum in most instances. It is both fast and efficient. Kuehn and Hamburger assert it is capable of handling realistic complexities of the type described earlier (capacity restrictions, interactions, economies of scale). Unfortunately they do not show how their program might be adjusted to handle these problems. It is notwithstanding an important contribution on which much additional work is based.

FLOW DIAGRAM

1. Read in:
   a) The factory locations.
   b) The M potential warehouse sites.
   c) The number of warehouse sites (N) evaluated in detail on each cycle, i.e., the size of the buffer.
   d) Shipping costs between factories, potential warehouses and customers.
   e) Expected sales volume for each customer.
   f) Cost functions associated with the operation of each warehouse.
   g) Opportunity costs associated with shipping delays, or alternatively, the effect of such delays on demand.

2. Determine and place in the buffer the N potential warehouse sites which, considering only their local demand, would produce the greatest cost savings if supplied by local warehouses rather than by the warehouses currently servicing them.

3. Evaluate the cost savings that would result for the total system for each of the distribution patterns resulting from the addition of the next warehouse at each of the N locations in the buffer.

4. Eliminate from further consideration any of the N sites which do not offer cost savings in excess of fixed costs.

5. Do any of the N sites offer cost savings in excess of fixed costs?

Yes

6. Locate a warehouse at that site which offers the largest savings.

No

7. Have all M potential warehouse sites been either activated or eliminated?

No

Yes

8. Bump-Shift Routine
   a) Eliminate those warehouses which have become uneconomical as a result of the placement of subsequent warehouses. Each customer formerly serviced by such a warehouse will now be supplied by that remaining warehouse which can perform the service at the lowest cost.
   b) Evaluate the economics of shifting each warehouse located above to other potential sites whose local concentrations of demand are now serviced by that warehouse.
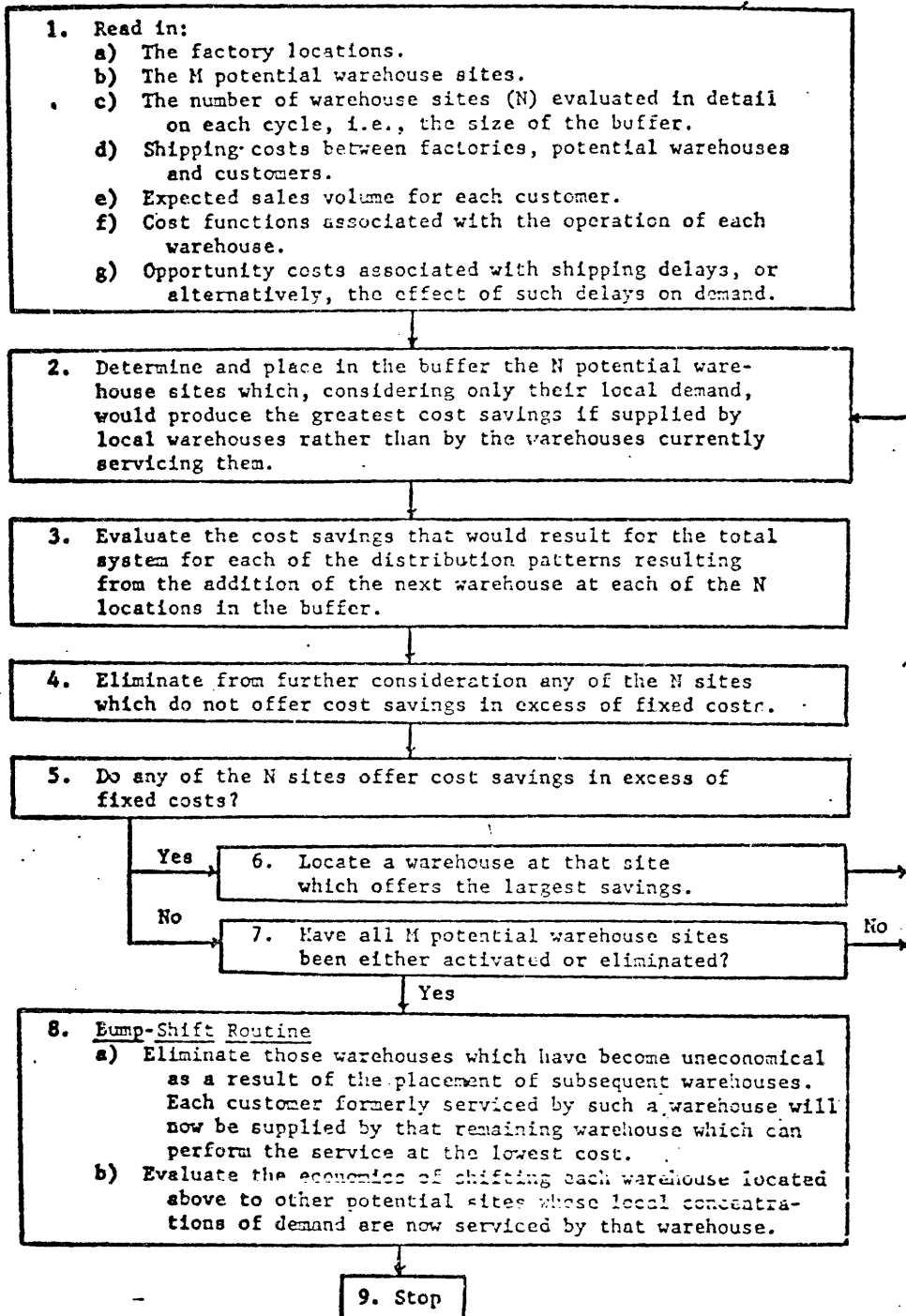
9. Stop

Figure 3. Heuristic Program of Kuehn and Hamburger
Source: Management Science (July, 1963)

Shortly after Kuehn and Hamburger's paper, A.S. Manne [15] proposed a somewhat similar approach for solving a class of fixed charge problems. His assumptions were very stringent, assuming a single product, single time period, no uncertainty in demand forecast, and no capacity limitations. He acknowledged at the outset that "it remains to be seen whether the performance of this technique is seriously downgraded as a result of introducing more realistic assumptions." Manne's algorithm is called the Steepest Ascent One Point Move Algorithm, SAOPMA. Manne points out that the uncapacitated structure of his formulation makes it an easy matter to calculate the X_{jk}'s of problem (I) given any set of warehouses that are fixed open or closed (Y_j = 1 or 0). Demand center k is simply supplied by that warehouse with the minimum value of $\overline{C}_{jk}$.

Manne adopts two heuristics comparable to Kuehn and Hamburger. He assumes that one need consider only a small subset of all possible locations sets from which to select the final solutions. He assumes further that you can get a reasonably good solution by only considering the status of one particular warehouse at a time. Here, however, his algorithm diverges from that of Kuehn and Hamburger.

A geometric insight into the problem is helpful. Each of the possible combinations of open or closed warehouses may be considered as one of the vertices of a unit hypercube (i.e., for n possible locations, there are 2^n possible warehouse locations vectors $(Y_1, Y_2, \ldots, Y_n)$ where each Y_j can take on values of either 0 or 1). As mentioned previously, for a given vertex it is trivial to allocate the demand concentrations among the open warehouses and to compute the associated value of the objective function at that vertex. For large n, the 2^n possible combinations make complete enumeration impossible. As is the case with other heuristics, SAOPMA attempts to reduce the number of vertices examined before a good solution is obtained.

SAOPMA starts at an arbitrary lattice point or vertex and evaluates all possible one point moves to adjacent vertices. It selects that move which offers the greatest total cost reduction, or terminates if no improvement is found. By a one point move is meant a single element alteration the n-element location vector. (For example, take n = 3. If we are at vertex(1,0,1) warehouses 1 and 3 are open. Possible one point moves are to (0,0,1), (1,1,1), and (1,0,0), but no others.)

Manne provides an interesting economic interpretation of the model, but we shall not delve into that here. He concludes from his computational exprerience that SAOPMA is capable, under the present assumptions, of obtaining solutions within 2-6% of optimum.

More recently, Feldman, Lehrer, and Ray (FL&R) [7] proposed extensions of the work done by Kuehn and Hamburger, Manne, and others. They comment that previous attempts to explicitly consider economies of scale were represented by fixed charges associated with opening a warehouse as in our problem (I) formulation. FL&R, however, are concerned with problems where the economies of scale affect warehousing costs over the entire range of warehouse sizes. Therefore, they consider the problem (I) where the warehousing cost is a strictly concave function as in the Balinski-Mills model (Figure 2). To do this they develop additional heuristics for assigning customers to those warehouses that have been opened. This assignment for the non-linear cost function $f_j(T_j)$ is no longer a trivial assignment, as in the case $f_j(T_j) = F_j + \lambda_j T_j$.

FL&R were able to incorporate allowances for different regional warehousing costs, a more realistic assumption than had previously been provided for. They also developed two solution routines for their problem. The first is an add routine, similar to Kuehn and Hamburger in which warehouses are added one at a time. The other is a drop routine in which all warehouses are initially assumed

open and are then closed one at a time. They justify the development of separate routines on the basis of practical considerations. Only infrequently is a distributor interested in designing a system from the ground up. Often the problem facing management is how to consolidate present facilities. Use of a particular routine will clearly depend on the specifics of a problem. A further justification for the drop routine is the problem of handling infeasible routes, routes for which no transportation costs are specified or to which one may wish to assign infinite costs.

FL&R describe their heuristic for handling the non-linear warehousing costs in the following manner:

> The approach taken is to evaluate, for each customer the total incremental cost, including transportation and operating components, associated with shipments from each of the potential suppliers. For the problem we are considering the transportation aspects are trivial; unit costs are presumed independent of a shipment size. Incremental warehousing costs, on the other hand, are usually a decreasing function of warehouse throughput.
> The method for assigning customers to warehouses, based upon both transportation and warehouse cost requires the determination of some "reference level" for warehouse size. For the purpose of providing these required reference levels we define each warehouse's local customer set (LCS) as consisting of those customers to whom the warehouse is closest on a pure transportation cost basis. The warehouse can then be said to have a local volume that is the sum of the individual demands in its LCS. The quantity thus determined is taken as a preliminary measure of the extent to which a warehouse is centrally located, and is used to get an idea of which portion of the cost curve is most relevant to decisions involving this warehouse.
> Once local volumes have been established, warehouse-customer allocation is independent of those made for all other customers. We simply examine the incremental cost of supplying a given customer ifrom each of the available warehouses, assuming that these warehouses have, independent of the allocation in question, throughput levels equal to their local volumes.
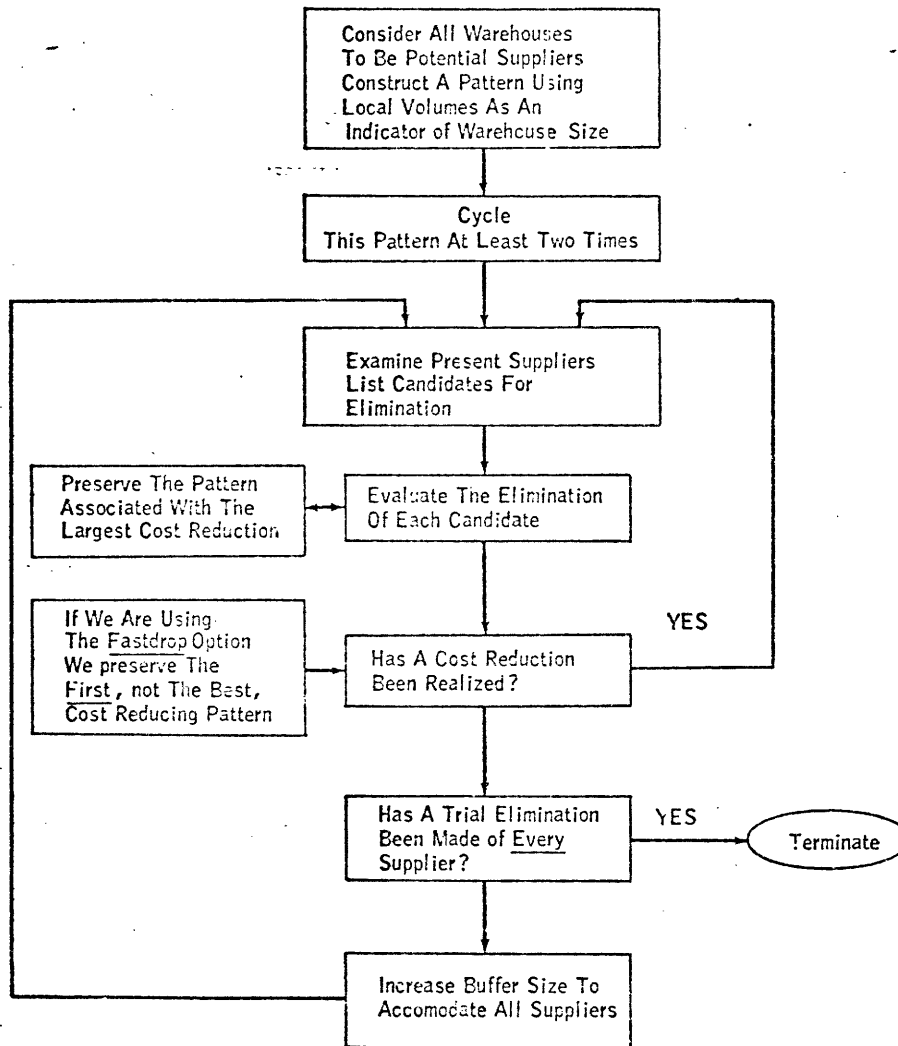
Figure 4. The Drop Approach of Feldman, Lehrer, and Ray
Source: Management Science (May, 1966)

Use of the LCS concept identifies that portion of the warehousing cost curve a particular warehouse is expected to operate on. Their "drop" routine has some built-in characteristics that make it particularly attractive. After the initial stage, in which the local volumes described above are used, they are no longer required for subsequent analysis. At each stage of the solution a far better indication of ultimate warehouse size exists -- the actual throughputs characterizing the pattern in use at that time. Effective use is therefore made of information available at each stage in proceeding toward an ultimate solution. A detailed flow diagram of the drop approach is given in Figure 4 (page 12).

In testing their routines, FL&R found they were able to generate location patterns (using problems tested earlier by K&H) that were in no cases higher than those found by Kuehn and Hamburger. However, they also noted they were unable to obtain any large improvements, which speaks well for the simple heuristics we have previously discussed. There may not be much to be gained by conducting elaborate searches as is often proposed.

They then tested their codes on larger problems with both a cost curve of the form $f_j(T_j) = F_j + \lambda_j T_j$ (Figure 1), and with a cost curve of the form (Figure 5),

$$f_j(T_j) = \delta_{j1} T_j \qquad T_j \leq t_1$$
$$= F_j + \delta_{j2} T_j \qquad T_j \geq t_1$$

They found that the "optimal" solutions obtained were very sensitive to the form of the cost function used. They recommend that oversimplification of the cost function be avoided in the formulation phase.

Figure 5
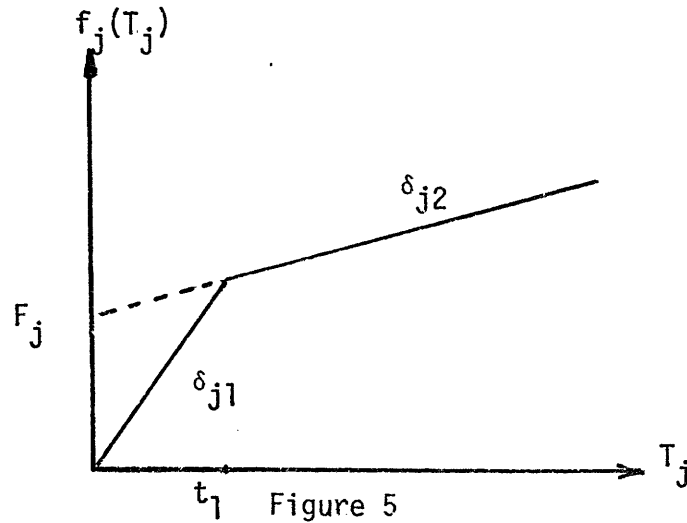
As mentioned previously the drop routine provides for its own updating of the reference level used in determining which warehouses are to leave the solution. The add routine does not do this. It seems in fact that FL&R propose to use the reference level based on initial LCS throughout the entire add routine. It appears this may introduce distortions into the problem, in that at any given stage one warehouse may be picked as superior to another based on LCS, when in fact it is far from superior. Their problem testing, however, does not bear this out. It would be informative to develop a means of updating the reference level in the add routine.

The importance of FL&R's paper is for the complexities it adds in the allocation of demand phase of the problem. They treat the non-linear cost function by approximating it, to the desired degree of accuracy, by a piece-wise linear curve. Using reference levels based on the concept of a local customer set (LCS), they develop suitable heuristics. They have provided a means for developing good location patterns for rather complex problem formulation.

Simulation

Simulation, as described by Shycon and Maffei [19] allows one to operate some particular phase of a business on paper, or in a computer, for a period of time, and by this means test various alternative strategies and systems. Shycon develops a large scale simulation system to handle the problem of planning physical distribution for a multi-source, multi-product system. They attempt to determine the number and locations of warehouses, and the assignment of demands, the same problem we have been discussing. Rather than proceeding systmatically from a starting point toward an optimum solution, their simulation provides a means to look at proposed alterations in the distribution system. The decision maker must supply these alterations and decide when he is satisfied with the solution obtained.

Simulation models are often very detailed and attempt to consider many of the interactions in the system. They also require an enormous amount of data collection, substantial development time, and often large investment sums. The question of how good are the results obtained is never clearly determined, for there is often no standard to compare against. For the H.J. Heinz Company, the simulation model provided a solution which offered some cost savings, but which was only a small percentage (see [19]).

A much more recent development is the Distribution System Simulator (DSS) proposed by Connors, et al. [4]. DSS produces a customized simulation model for a large scale physical distribution system. The use of DSS as described by Connors et al. is given below.

> The user of DSS responds to a questionnaire which contains the options that he can use to develop a model of his distribution system. The user specifies the characteristics of the desired model by answering true or false to questions expressed in English. The options allow the analyst to take into account each of the major factors involved in the operation of a distribution system: the characteristics of customers' demand for products, buying patterns of customers, order filling policies, replenishment policies, redistribution policies, transportation policies,

distribution channels, factor locations, production capa-
bilities, and other significant elements. These options
are essentially inventory an! product movement-oriented --
beyond this, DSS provides the capability, through user
functions, to incorporate other vehicle scheduling algorithms,
forecasting techniques, production schedules, and pricing
mechanisms which are outside the scope of the options.

DSS actually generates the computer program required to perform the simula-
tion, and specifies the data input to be provided by the user. It is a remark-
able software package for analyzing a distribution system.

Hax [8], however, comments that DSS has several serious failings. He des-
cribes these as follows:

(a) It fails to support plant and warehouse location decisions,
as well as decisions regarding expansions or improvements in
the production and distribution facilities. By approaching
the problem from a simulation point of view, instead of
using an optimization approach, DSS is hopeless in this
respect.

(b) It ignores entirely the production process and the very dif-
ficult questions affecting the production planning process
interacting with a complex distribution process. DSS treats
the manufacturing plants as a source of unlimited inventory,
which is naive and overly simplistic.

(c) It does not provide an integrative approach to the logistics
decision process. Essentially, DSS treats each stocking point
as if it were independent from the rest of the system. The
difficult problem faced in a multi-level, multi-item distribu-
tion situation is the optimum allocation of the total available
inventory among the various stocking points. Basic issues to
be resolved are: where to stock a given item, and what strat-
egies to use in the stock allocation, replenishment and tran-
shipment processes. A simulation approach to deal with these
issues seems highly inadequate to me.

Hax suggests that the customized approach of Connors is basically sound, but
needs development to include certain other issues, These issues include providing
a framework for a hierarchical system of decision making, concentrating in effec-
tive ways in which to accomplish the partitioning, linkage, aggregation, and dis-
aggregation of the decision process and to measure its overall performance.

Simulation of itself does not appear a most useful tool with regards to the
warehouse location problem. It has many other applications in which its capabili-
ties seem more appropriate.

Branch & Bound Techniques

Branch and bound techniques are algorithmic procedures designed to generate optimal solutions. Many of the techniques currently available are based on the branch and bound code of Effroymson and Ray (E&R) (see [6]). The Facilities Location Planner (FLP), an analytic approach based on this technique, has been developed into a computer procedure to aid management in analysis of decision alternatives (see [1]). Unfortunately, the FLP model fails to deal with complex capacity limitations. This is mitigated somewhat by the fact that FLP is considered a long range investment planning aid in which such constraints are not serious shortcomings.

Initially Effroymson and Ray formulated their problem by first using the formulation of our problem (I), repeated here for convenience

$$\min z = \sum_j \sum_k c_{jk} X_{jk} + \sum_j F_j Y_j$$

$$\sum_j X_{jk} = 1$$

$$0 \leq X_{jk} \leq Y_j \leq 1$$

$$Y_j = 0, 1$$

Then added the following definitions, letting

N_k = set of warehouses that can supply customer k

P_j = set of customers that can be supplied by warehouse j

n_j = number of elements in P_j

The branch and bound algorithm is described in detail in the literature [6], [13]. The idea is to solve a sequence of linear programming problems, not necessarily satisfying the integer restrictions, that give progressively lower bounds on the value of the solution to the mixed integer problems. Khumawala provides a good description of the process in one of his earlier papers [9].

Problem (I) is first solved as a linear program (without the integer restrictions on the Y's). Let Z_0 be the solution that problem. If all the Y's are integer, then the problem is solved. If some Y_k is fractional, then: (a) the restriction $Y_k = 0$ is added to the problem and the problem is again solved. Let Z_1 be the solution; clearly $Z_1 \geq Z_0$; (b) the restriction $Y_k = 1$ is added to the problem and the problem is again solved; also, $Z_2 \geq Z_0$. Then $\bar{Z} = \min (Z_1, Z_2)$ is a new lower bound on Z. This procedure has resulted in the construction of a tree whose nodes are represented by the Z's and the corresponding value of the fixed Y's. If a node is reached where all the Y's are integers the LP solution at this node forms an upper bound on Z. A node where all the Y's are integers will be called a terminal node, as opposed to a nonterminal node, where at least on Y is fractional; the LP solution at a terminal node will be referred to as a terminal solution. Branching continues from nonterminal nodes, whose LP solutions are less than the current upper bound; i.e., a fractional Y at such a nonterminal node is constrained 0 and 1 and the linear programs are solved at the two additional nodes. The b & b algorithms continues in this manner, updating the bounds at each stage. Of course, no branching takes place from an infeasible node, the node at which the LP solution is infeasible. The algorithm stops when a nonterminal node, whose LP solution is less than the current upper bound, cannot be found. The current upper bound is then the optimal solution. The tree constructed in this manner will be called the b & b tree for the problem and its size will be measured by the number of nodes in the tree.

Effroymson and Ray note, as we have also observed, that due to the uncapacitated nature of the formulation, the optimal solutions (allocation) to the linear program at each node is trivial. They continue by defining K_1, K_0 and K_2 as the set of indices of warehouses that are fixed open, fixed closed and free at the node. Then the optimal LP solution at the node is:

$$X_{jk} = 1 \text{ if } C_{jk} + g_j/n_j = \min_{j \epsilon (k_1 \cup k_2)} [C_{jk} + g_j/n_j],$$

$$= 0 \text{ otherwise,}$$

$$Y_j = 0 \qquad\qquad j \epsilon K_0,$$

$$= \Sigma_{k \epsilon P_j} X_{jk}/n_j \qquad j \epsilon K_2,$$

$$= 1, \qquad\qquad j \epsilon K_1,$$

where

$$g_j = F_j \qquad\qquad j \epsilon K_2,$$

$$= 0 \qquad\qquad j \epsilon K_1.$$

At this point E&R propose three simplifications for use at each node. These simplifications can significantly reduce the number of possible branches which need to be evaluated at each node, thereby reducing the size of the branch & bound tree. These are (1) a minimum bound for opening a warehouse, (2) a means of reducing n_j, the number of customers that can be supplied from warehouse j, and (3) a maximum bound on the cost reduction for opening a warehouse. Details of the computations involved in the simplification are not of specific interest to us here, but may be found in [6] or [9].

The E&R algorithm cycles through these three simplifications at each node until no further openings or closings can be effected. As E&R note, the main problem with branch and bound is computational. If a large number of linear programs have to be solved, and computation time for each is high, the method becomes prohibitively expensive.

Khumawala, using the same formulation as E&R, was able to develop several additional means to guide the branch and bound algorithm more efficiently. He describes these further refinements in [9] as

(1) Formal rules for selecting the free warehouse at a given stage from among those in the set of free warehouses at that stage. These he calls "branching decision rules."

(2) A means of using much of the information already available at a stage to solve the linear program (I) without the integer requirements, as rapidly as possible.

(3) Programming related improvements resulting in efficient storage use.

The branching decision rules that Khumawala tested were eight in number. They are all based on information obtained in performing the simplifications described by E&R. As such they require no additional computation or storage. He tested the eight rules and found one superior to the others in almost every case. This he calls the "largest omega" rule.

Omega (Ω) is defined as the sum of the minimum savings for supplying each customer that can be made, if a given warehouse is opened, considered over all the fixed open warehouses at a node. If this sum fails to exceed the fixed cost, F_j, of opening a warehouse, then such a warehouse is closed and eliminated from further consideration. Mathematically we let

$$\omega_{jk} = \text{Min}_{r \epsilon (N_k \cap K_1)} [\text{max } (C_{rk} - C_{jk}, 0)]$$

and define

$$\Omega_j = \sum_{k \epsilon P_j} \omega_{jk} - F_j$$

We then select as the next warehouse to enter the set of open warehouses, that which has the largest omega.

Khumawala[a] also proposed a means for lessening storage requirements through "judiciously deleting nodes no longer necessary for the algorithm. The storage used for these nodes is effectively used over and over again for the new nodes that are generated as the algorithm proceeds."

In a very recent paper [10], Khumawala has proceeded still further in his work on the facilities location problem. Noting that the efficiency of the branch and bound method is approximately exponential with the number of potential warehouses considered, he developed a set of "simple and computationally inexpensive heuristic rules." He states that his results show the computational efficiency of these rules to be substantially less than linear while still providing excellent solutions.

He uses all of the eight branching decision rules alluded to above and with each branching rule he now associates a heuristic decision rule. Simply his heuristic rules are as follows: "Trace the preferred path which would result from the application of the particular decision rule." In essence, he takes a given problem and applies in turn each of the branching decision rules and their associated heuristics. The solution is the best of the eight solutions obtained.

Testing his program on several problems from the literature he finds solutions that are optimal or very near optimal in all cases. Computation time is fast and efficient. He suggests that although it is computationally feasible to use all eight rules, certain problems are solved more readily by individual heuristics than others. When this is the case certain branching rules could be eliminated.

In another branch and bound approach, Spielberg [20] suggested algorithms based on both opening and closing of warehouses. His justifications for the closing or drop routine were comparable to those of FL&R described earlier. Spielberg's computational results indicate that many problems easily solvable by one algorithm become very difficult when attempted by use of the other algorithm. This suggests the starting point need be determined as each particular situation warrants. In a later paper [21], he proposes a "generalized search origin," whereby the search may be started at any convenient point, be it known apriori, generated randomly, or generated dynamically during the search. He feels his results are satisfactory, but require further development.

One further approach at the problem deserves mention. This is a general integer linear programming algorithm tested by Ocampo-Gaviria in a recent Master's thesis. He used an IBM program, MPSX, to solve the warehouse location problem. The program has two stages. The first stage solved the problem with all integer variables considered continuous. The second stage then searches integer solutions using a branch and bound technique with heuristic rules. The search continues as long as improvement is realized in the objective function. Then optimality must be proven.

The conclusion drawn from his work was that "MPSX will find 'good' integer solutions relatively fast, but may take a long time to prove optimality of the solution." This is seen as the major failing of the approach, although optimality is by no means our sole criterion.

Capacitated Models

Comparatively little work has been done with the capacitated location problem, although it lends itself quite well to mathematical formulation. If we define K_j as the capacity of the jth warehouse then

$$\sum_k X_{jk} \leq Y_j K_j$$

Sa [18] has proposed an exact branch and bound treatment of the capacitated warehouse location problem. He developed an approximate algorithm for solving the case when many integer variables are involved and tested his technique on problems from the literature. His findings indicate that the algorithm is effective for large problems, but the branch and bound routine is greatly restrained by problem size.

Davis and Ray [5] take a slightly different approach using a mixed integer programming method for solving the same problem. They further use a decomposition principle to solve the dual of an associated continuous problem at each branch and bound iteration. The rationale was an attempt to strike a balance between the number of branches to be examined and the cost of computation.

The problem in capacitated formulations is in computation time and cost. Additional work is clearly desirable in these areas.

III. Public versus Private Warehousing

Concurrent with decisions regarding warehouse location, management must consider the issue of public versus private warehouses. Each offers particular advantages under specific conditions. Consideration of these alternatives and conditions under which the firm's logistic system is operating should be thorough and complete before commitments are made.

Private warehousing offers several potential economic advantages. The choice of private warehousing allows management to design to internal specifications consistent with operational requirements. Size, storage requirements, and special considerations can be tailored to effect internal efficiency. In addition, such facilities are particularly attractive when throughput is stable and seasonality is not a factor.

Public warehousing offers the obvious advantage of flexibility. A firm may choose to alter warehouse location patterns with changes in demand patterns, market conditions, and seasonality effects. Commitments can be made for short or long term use of such facilities. Public warehouses offer a variety of services including receiving, handling and shipment, and loading operations. Other services are frequently available for the convenience and option of the leasee. In short public warehousing offers a full range of facilities, manpower, and equipment on a long or short term basis to support a customer's local sales and distribution activity.

The relative costs of the two alternatives differ. Public warehousing will almost always cost more per unit of storage and handling than a well managed private warehouse. This is natural since the owners of public warehouses make their profit on the operation of a warehouse. However, the differences in cost between the two are often difficult to discern due to the nature of accounting practices. Careful examination should be given to such analysis to determine which alternative is economically preferable.

Magee [14] provides a summary of circumstances under which public or private warehousing may be favored. It is reproduced here.

Circumstances Favoring

	Private Warehousing	Public Warehousing
Market		
Geographic Location	Concentrated, Stable	Dispersed, Shifting
Level of Demand	Stable	Volatile
Seasonality	Uniform	Seasonal
Product Line		
Promotions	Limited or Unimportant	Frequent and Important
Physical Characteristics	Special Handling or Storage	Standardized Packaging, Handling, Storage
Transporation Media		
Technique Used	Stable	Changing

Source: Magee, J.F., _Industrial Logistics_

IV. Conclusions

Substantial effort has been directed at solving the warehouse location problem in its various forms. The development of this effort has been traced, analyzed, and evaluated. Computational aspects emerge as a central problem in most every approach. This then is the problem to be solved.

Suitable heuristics have and will continue to facilitate the solution to this problem. Furthermore, it seems clear that branch and bound techniques are also essential to the effort. It is through further joint application of these methodologies that productive effort should be directed.

REFERENCES

1. Atkins, R.J. and R.H. Shriver, "New Approach to Facilities Location," Harvard Business Review (May-June, 1968).

2. Balinski, M.L. and H. Mills, "A Warehouse Problem," Mathematica (April, 1960).

3. Baumol, W.J. and P. Wolfe, "A Warehouse Location Problem," Operations Research, Vol. 6 (March-April, 1958).

4. Connors, M.M., Coray, C., Cuccaro, C.J., Green, W.K., Low, D.W., and Markowitz, H.M., "The Distribution System Simulator," Management Science, Vol. 18 (April, 1972).

5. Davis, P.S. and T.L. Ray, "A Branch-Bound Algorithm for the Capacitated Facilities Location Problem," Naval Research Logistics Quarterly, Vol. 16 (September, 1969).

6. Effroymson, M.A., and T.L. Ray, "A Branch and Bound Algorithm for Plant Location," Operations Research, Vol. 14 (May-June, 1966).

7. Feldman, E., Lehrer, F.A., and Ray, T.L., "Warehouse Locations Under Continuous Economies of Scale," Management Science, Vol. 12 (May, 1966).

8. Hax, A.C., "A Comment on the Distribution System Simulator," Unpublished, MIT (June, 1973).

9. Khumawala, B.M., "An Efficient Branch and Bound Algorithm for the Warehouse Location Problem," Management Science, Vol. 18 (August, 1972).

10. _____, "An Efficient Heuristic Procedure for the Uncapacitated Warehouse Location Problem," Naval Research Logistics Quarterly, Vol. 20 (March, 1973).

11. Kuehn, A.A. and M.J. Hamburger, "A Heuristic Program for Locating Warehouses," Management Science, Vol. 9 (July, 1963).

12. Liberatore, J.A., "Development and Evaluation of a Modified Heuristic Program for Locating Warehouses," MIT Master's Thesis (September, 1970).

13. Little, J.D.C., K.G. Murtz, D.W. Sweeney, and C. Karel, "An Algorithm for the Traveling Salesman Problem," Operations Research, Vol. 11 (November-December, 1963).

14. Magee, J.F., Industrial Logistics, New York: McGraw-Hill (1968).

15. Manne, A.S., "Plant Location Under Economies of Scale-Decentralization and Computations," Management Science, Vol. 11 (November, 1964).

16. Matthews, J.B., Jr., R.D. Buzzel, T. Levitt, and R. Frank, Marketing: An Introductory Analysis, New York: McGraw-Hill (1964).

17. Ocampo-Gaviria, A., "Design of a Distribution System," MIT Master's Thesis (June, 1973).

18. Sā, G., "Branch and Bound and Approximate Solutions to the Capacitated Plant Location Problem," Operations Research, Vol. 17 (November-December, 1969).

19. Shycon, H.N., and R.B. Maffei, "Simulation - Tool for Better Distribution," Harvard Business Review (November-December, 1960).

20. Spielberg, K., "Algorithms for the Simple Plant Location Problem with some Side Conditions," Operations Research, Vol. 17 (November-December, 1969).

21. _____, "Plant Location With Generalized Search Origin," Management Science, Vol. 16 (November, 1969).