1970

An investigation into the effects of partitioning the facilities assignment problem by hierarchical clustering methods

Charles Henry Blackburn III
Lehigh University

AN INVESTIGATION INTO THE EFFECTS OF PARTITIONING THE FACILITIES ASSIGNMENT PROBLEM BY HIERARCHICAL CLUSTERING METHODS

by

Charles Henry Blackburn III

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Industrial Engineering

Lehigh University

1970

Certificate of Approval

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

May 13, 1970
Date

Professor in Charge

Chairman of the Department
of Industrial Engineering

Acknowledgements

The author wishes to express his appreciation to Professor A. F. Gould of the Industrial Engineering Department, Lehigh University, for his guidance and interest in the work presented here. Thanks are also due to Professors J. W. Adams and Sutton Monro for serving as members of the thesis advisory committee.

I wish to also thank Dr. E. S. Bahary and P. T. Lele of the Western Electric Company's Engineering Research Center for their time, knowledge and support.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Recently the branch and bound technique has been applied to the facilities-location assignment problem. However, the computing times for problems having more then ten facilities are still prohibitive.

This thesis considers the possibility of partitioning larger problems by clustering the facilities according to the hierarchy of their mutual flows. Two different methods of accomplishing this clustering are developed and evaluated. A model is developed to partition the problem by these methods and to use a branch and bound algorithm at two levels. One level arranges the clusters in an optional manner and the second level arranges the facilities within the clusters.

The number of partitions, the configuration of the partitions, the clustering methods, and the size of the problem are evaluated forr their effects on the final solution value.

It is found that increased problem size has a detrimental effect on the efficiency of the solution. As the number of partitions increases, the cost of the solution also increases. A combination of increased problem size and increased number of partitions also has a negative effect on the quality of the solution. The shape of the partitions of the floor plan has a significant effect with a compact arrangement being the best.

I - INTRODUCTION

The problem of determining optimal locations for the facilities within an industrial plant has created considerable interest among plant and factory engineers, industrial engineers, management scientists, and operations researchers. This continuing interest has led to the development of both optimal-producing and heuristic improvement algorithms. The optimum seeking techniques have the disadvantage of being computationally feasible for only relatively small problems. Conversely, the heuristic methods can accommodate problems of a larger dimension, but they usually produce sub-optimum results.

Recent developments[15,46] have proposed the use of the branch and bound technique, an optimum seeking method, for the solution of facilities assignment problems. While these developments are encouraging, test results indicate that the branch and bound algorithm still requires excessive computing time for problems with more than ten facilities.

The objective of this thesis is to determine the effectiveness of solving a large problem by decomposing it into several smaller ones. Each of these small problems will be solved by Turner's[46] branch and bound algorithm and then they will be integrated into a complete solution.

Obviously, an optimization of these sub-problems should result in a less than optimal solution for the entire problem, and the technique of partitioning will have a bearing on the final result.

Therefore, secondary objectives of this thesis are the evaluation of the effects of the partitioning technique, the number of partitions, and the configuration of the partitions on the final solution.

The method of decomposing a facilities assignment problem should be related to some common characteristic of the various components. Current textbooks[2,22,37] propose that the facilities be grouped by either a similarity of function or by common products made with the facilities. The resultant assignments are respectively designated as "process" and "product" oriented layouts. For a more general case, this thesis will assume that neither of these classifications are relevant, so that a "job shop" layout must be designed. The problem must now be partitioned by some relative degree of the inter-facility traffic.

It seems appropriate to begin the thesis with a discussion of the problem and the limitations of the branch and bound technique. This will be followed by a review of existing models and the current state of the art. Chapter IV will be devoted to a description of the branch and bound algorithm used in this thesis. Chapter V will discuss the aspects of partitioning the problem. The final two chapters will be used to present the experimental results and to convey the conclusions that may be derived from these results.

II - THE PROBLEM

A general statement of the facilities-location assignment problem is that there are N different facilities to be assigned to N different locations, each of which can accommodate one facility. The analyst is required to make an assignment that will optimize some sort of an objective function. This problem has many facets and has received a considerable amount of interest. However, to the knowledge of the author, the problem presented here has been relatively neglected.

Definition of Terms

A definition of the terms recurring throughout this paper is in order before proceeding further:

Facility - an indivisible stationary unit which performs some specific function. The usual connotation is that individual machines or people are facilities, but this is not necessary here. Groups of machines, groups of people, or departments may be considered to be one facility as well, depending on the scale of planning.

Location - the unit of area into which the total floor space has been, or can be divided.

Flow - the volume of traffic moving between individual facilities. This traffic may consist of either physical units (such as weight, quantities, or the number of trips) or intangibles such as communications.

Distance - the units of separation between locations. This is commonly considered to be the actual distance between locations, the time, the cost, or some index of the difficulty of transfer.

Layout - a plan for the arrangement of the facilities within an industrial plant or an office building.

The assignment of facilities to locations is only one aspect of designing a layout. There are many other relevant problems and factors in layout planning that are not within the modest scope of this paper.

The Layout Problem

There are several different degrees to the layout problem:

(1) Assigning additional disjoint facilities to an existing layout.

(2) Assigning additional interrelated facilities to an existing layout.

(3) Relayout, or relocating the existing facilities, with no additional facilities.

(4) Relayout with additional facilities to be added.

(5) Planning a layout for a new floor space.

It can be said in general that the complexity of the problem increases as a function of the number of facilities, the heterogeneity of the locations, and the conjunction of the inter-facility flows.

The simplest form of the problem is associated with homogeneous locations and independent facilities. The facilities-location assignment problem then becomes an assignment problem with N! feasible solutions from which the optimal must be found. However, the sheer number of possible solutions rapidly becomes unmanageable for a problem of really small size. As an example, the assignment of ten facilities to ten locations has 10!, or 3,628,800, feasible solutions. The problem increases in complexity if other factors, such as heterogeneous locations, must be considered.

Simplifying Assumptions

In order to reduce the complexity of the problem, some combination of the following assumptions is used in the existing techniques and models.

(1) The flow data is deterministic.

(2) Cost data and flow data are available for conditions that are probably unknown.

(3) One and only one facility may be assigned to a location.

(4) Any facility may be assigned to any location.

(5) The flows of a wide variety of materials can be equated to equivalent units.

(6) The distances of travel can be approximated by direct or rectangular distances between the centers of the facilities.

(7) An effective measure for the value of an assignment can be determined.

The existing models consider the frequency of traffic to be deterministic, which is an unrealistic assumption. Obviously, the volume of production is a random variable which is a function of time. Currently, the addition of stochastic processes to the assignment models is infeasible, so expected values must be used if it is necessary to recognize the probabilistic nature of the data.

Most models assume that the various flows of widely divergent classes of materials can be equated to consistent units. Actually, very little research has been devoted to this area. If the materials are similar or reasonably homogeneous, common units of measurement are satisfactory. Usually the materials are diverse in nature and handling characteristics, which requires some conversion between the various volumes.

Muther[40] describes the Mag Count as a measure of the transportability of any item in any condition. He describes the basic unit, one mag, as an item that: can be conveniently held in one hand, is reasonably solid, is compact, is only slightly susceptible to damage, and is reasonably clean and firm. He then lists tables to modify the Mag count according to variations in these characteristics.

The common assumption that the distances of travel are approximated by either the direct distances or rectangular distances between centers of facilities is appropriate for all but vertical movement. Normally, the mode of travel in an industrial plant is along a series of orthoginal aisles, which can be represented by the rectangular distance between centers of locations.

Undoubtedly, the most important aspect of an assignment model is the method of evaluating the effectiveness of a solution. Traditionally the solutions have been judged by a measure of the flow movement. Usual criteria are (1) total distance that the traffic moves, (2) the total number of handlings, or (3) some weighted total of either of these first two.

Reis and Anderson[42] proposed the use of "importance factors" to weight certain flows that are found to be relatively important. They define these factors as "any factor other than volume of product or distance to be moved that is to be considered in determining a good plant layout from a materials handling point of view." These importance factors would be used, for example, to give priority to hazardous moves, counterflow, or cross-flow. While these factors are not widely used, they are compatible with the existing models and could be included easily.

Statement of the Problem

The most commonly used criterion is the first one listed above, the total distance that the traffic moves. This allows the problem to be formulated as a linear assignment model:

$$\text{Minimize} \quad \sum_{i=1}^{n} \sum_{j \neq i}^{n} f_{ij} d_{kl}$$

where:

N = the number of facilities and the number of locations.

$f_{ij} d_{kl}$ = the "cost" of assigning the i^{th} facility to the k^{th} location and the j^{th} facility to the l^{th} location.

f_{ij} = the flow of traffic between the i^{th} facility and the j^{th} facility ($f_{ij} = 0$ for $i = j$).

d_{kl} = the distance between the k^{th} location and the l^{th} location ($d_{kl} = 0$ for $k = l$).

Subject to these constraints:

(1) Only one facility may be assigned to a location.

(2) All facilities must be assigned to locations.

(3) All facilities are indivisible.

An additional assumption implied by this formulation is that the terminal costs of loading and unloading the material may be disregarded with no loss of significance. This assumption may be justified by considering these costs as being applicable to any arrangement of facilities. Therefore, these costs are constants that will not affect the assignments of the facilities and may be disregarded.

The linear assignment problem is not an easy one to solve due to the sheer magnitude of the number of feasible solutions. As will be shown in Chapter III, many attempts have been made to develop algorithms for its solution. Optimum seeking methods have been found, but they are computationally infeasible when more than ten or fifteen facilities are to be assigned. Some excellent heuristic algorithms have been developed that can solve problems with a relatively large number of facilities, but they generally produce sub-optimal results.

Recently, there have been some interesting developments in the use of a new optimum technique, branch and bound. This technique was applied to the facilities assignment problem by Gavett and Plyter,[15] and their algorithm was subsequently improved by Turner.[46] The author recently tested the running times of this algorithm on a large set or problems using the IBM-360/50 computer. These running times versus the problem size are shown on Figure I. It can be seen

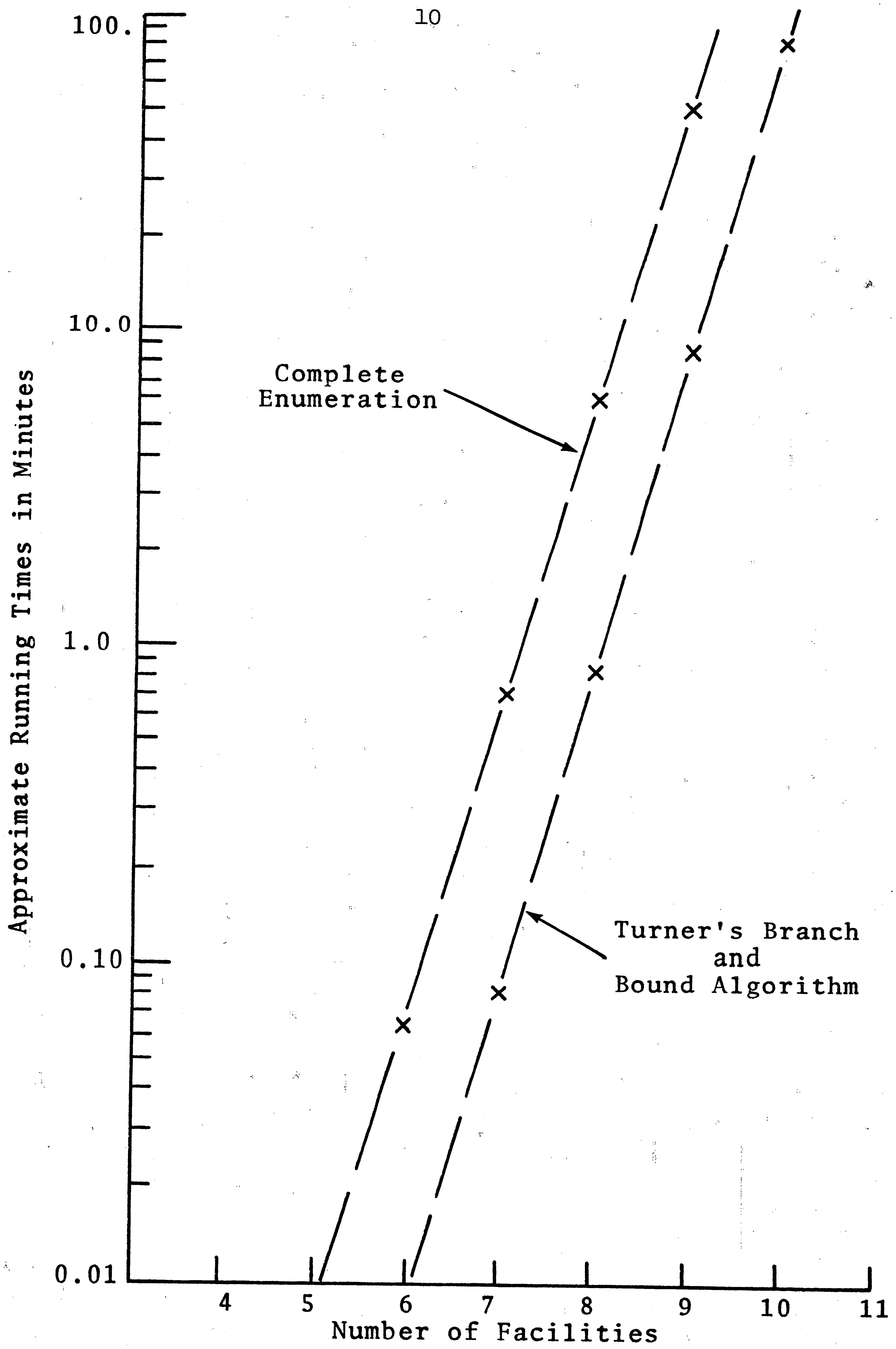


FIGURE I
Approximate Running Times on the IBM 360-50 Computer

that problems with more than ten facilities generally require prohibitive computing times. The times required to solve the same problems by complete enumeration are also shown, since these times are the upper extremes for the algorithm when it must examine every possible solution.

Figure I gives an approximate running time of 100 minutes for a problem with ten facilities to be assigned. However, the total running time for two problems with five facilities each would be so small that the measurement is insignificant. Obviously, there is a tremendous time saving to be realized by dividing a large problem into smaller sets.

When this idea is projected into the region of larger problems, the benefit to be accrued is computational feasibility. Extrapolating the plot of Figure I to a problem with twelve facilities gives an estimated running time of 10^4 minutes (6.8 days). Continuing to a problem with fifteen facilities gives an estimated time of 10^7 minutes (19 years) to find a solution. Yet two problems with eight facilities each would require only two minutes of total computing time.

The second reason for partitioning large problems is the saving of computer memory. The branch and bound algorithm requires a large memory to store the lists that are needed to follow a solution. This requirement grows at an exponential rate with the problem size. However, this limitation is not nearly as severe as the constraint of computing time.

Since the time required to solve large problems by the branch and bound algorithm appears to be prohibitive, the question of partitioning has merit.

Can a large problem be solved by partitioning with no significant loss in quality? If a problem can be efficiently partitioned, what are the best rules to follow in choosing a partitioning strategy? These are the questions to which this thesis is addressed.

III - REVIEW OF EXISTING TECHNIQUES

The problem of assigning facilities to optimal locations is certainly well known and widely discussed. Moore[37] states that:

> "The problem of arranging an industrial process has been in existence as far back as the Industrial Revolution. When Taylor first developed his concept of scientific management, industrialists had been wrestling with the problem of arranging facilities for years. Although plant layout evolved as a distinct industrial function relatively recently, it was a dominant factor of production throughout the development of the factory system."

A problem of such longevity, complexity, and importance should obviously be the motivation for numerous articles and papers of significance, and a detailed review of these documents would require a volume in itself. Consequently, this chapter will be concerned only with the major trends in methodology and the more commonly accepted algorithms. These methods can be broadly divided into three general classifications: (1) the traditional manual methods, (2) the "graphic" techniques, and (3) mathematical models.

Manual Methods

The oldest methods, which are intuitive and manual in nature, consist of simulating a layout by arranging templates or other physical models on a scaled floor plan of the area to be used. The procedure includes an analyzation of the high volume traffic and then designing the layout around these activities.

The problem is sometimes partitioned, if possible, by grouping facilities according to similar functions or similar product. The quality of the solution is dependent upon the complexity of the problem and the expertise of the designer. No particular individuals are credited with the development of these methods, which are well described in plant layout textbooks.[2,22,37,40]

Graphic Methods

Throughout the 1950's manufacturing layout methodology progressed into the use of graphic and schematic analysis. The procedures involve the collection of data concerning the interfacility flows of traffic for some period of time. This data is then presented in matrix form as a "from-to-chart," "cross chart," or "flow matrix." Figure II depicts an example of such a matrix taken from Apple.[2]

FROM \ TO	STORES	MILL	LATHE	DRILL	BORE	GRIND	PRESS	HONE	SAW	INSP.	TOTALS
STORES		2	8			1	4		2		17
MILL			1	2			1			1	5
LATHE		2		4			1		1	3	11
DRILL		1			1		2	1		5	10
BORE				1							1
GRIND				1						1	2
PRESS				2						6	8
HONE										1	1
SAW			2			1					3
INSP.											0
TOTALS	0	5	11	10	1	2	8	1	3	17	58

Figure II - Example of a From-To-Chart

The analyst is left to his judgment and ingenuity for the determination of a set of assignments that will minimize the volume of non-adjacent facilities flows. Link analysis,[40] travel charting,[37] and operation sequence analysis[6] were subsequently developed to assist the analyst in this selection of assignments. While these techniques are great improvements over preceding methodology, they become cumbersome as the number of facilities increase. Vollmann and Buffa[48] state that these methods "become virtually unmanageable when the number of departments becomes at all large (say above 10) unless the flow has a dominant pattern."

Mathematical Models

Interest in mathematical treatments of the facilities-location assignment problem has been divided between heuristic, sub-optimal methods and true optimum seeking approaches. The heuristic methods can also be sub-classified into either improvement or construction types. The improvement algorithms are designed to improve on a given starting arrangement, while the construction algorithms are self-starting and require no initial solution.

There have been very little quantitative comparisons of the heuristic versus the optimal techniques. While Nugent, Vollmann, and Ruml[41] made a study along these lines in 1967, they were primarily interested in the relative efficiencies of the major sub-optimal techniques. It is their opinion that optimal procedures are not computationally feasible for large problems. This opinion appears to be well founded.

The heuristic approaches can solve problems that have a large number of facilities in a short time, but the results are admittedly sub-optimal. Conversely, the optimum producing methods are severely limited by the size of the problem that they can solve in a reasonable length of time.

Heuristic Methods

The heuristic methods consider the problem as a form of combinatorial analysis. Since the total number of feasible solutions is a factorial function of the number of facilities to be assigned, the problem is usually attacked in some manner that reduces the combinations to be evaluated. The trade-off is absolute optimality versus computational feasibility.

Wimmert[50] proposed one of the first of these methods in 1958. His model constructed an assignment cost matrix by the dot product of a monotonic non-increasing flow vector and a monotonic non-decreasing distance vector. His method of selecting assignments from the elements of this matrix was based upon an incorrect parameter of the matrix, and the model was subsequently disproved by Conway and Maxwell.[8] However, Conway and Maxwell did note that an optimum, although not necessarily feasible, solution could be found along the main diagonal of a matrix constructed in this manner.

Hillier[19] developed a heuristic improvement procedure in 1963 that is based on the assumption of travel along a system of orthogonal aisles. This allows changes in the X direction to be considered independently of those in the Y axis. Each facility can then be evaluated according to the effects of moving it either up or down (Figure III and IV).

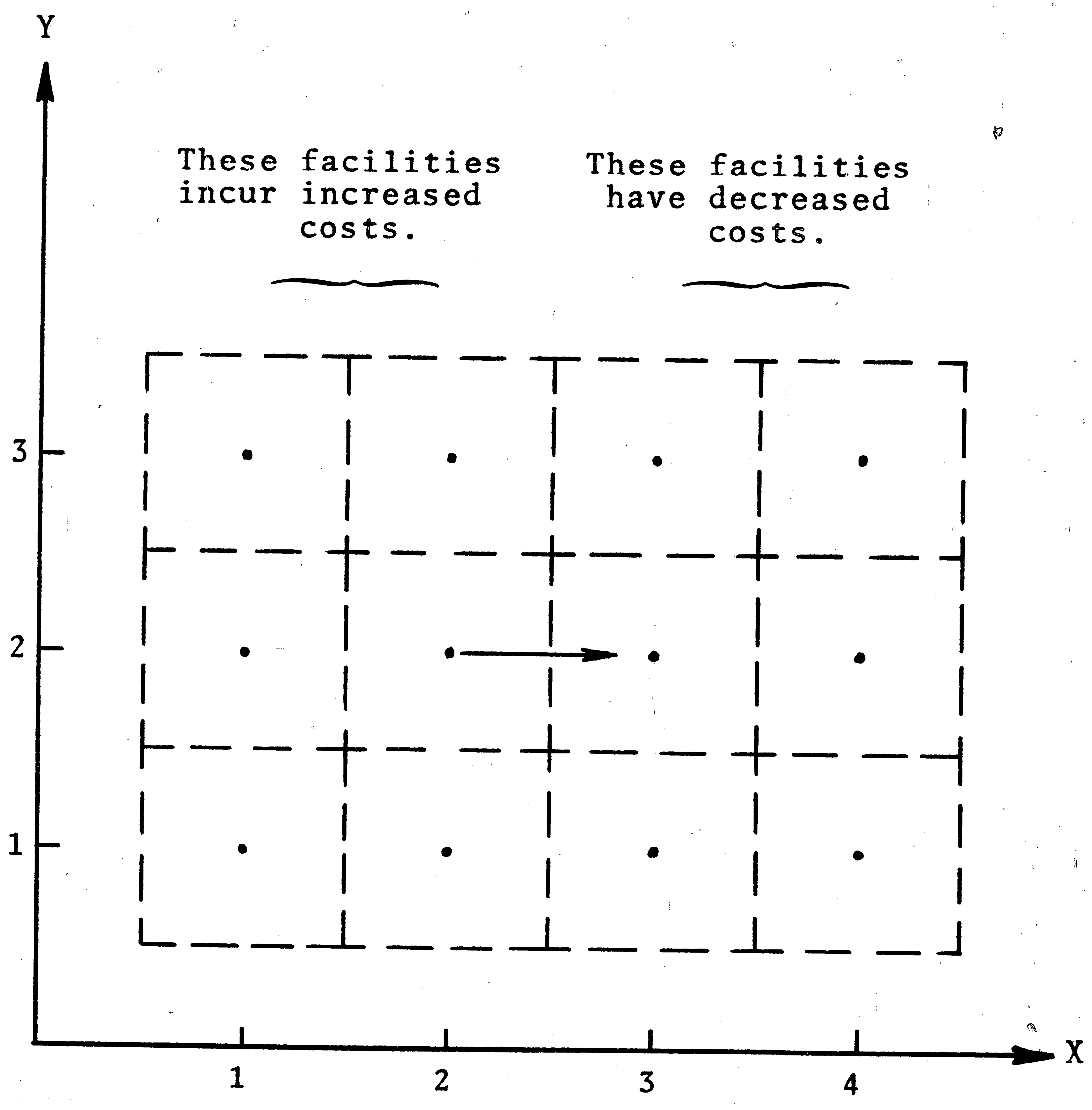


Figure III - The Effects of Moving a Facility to the Right

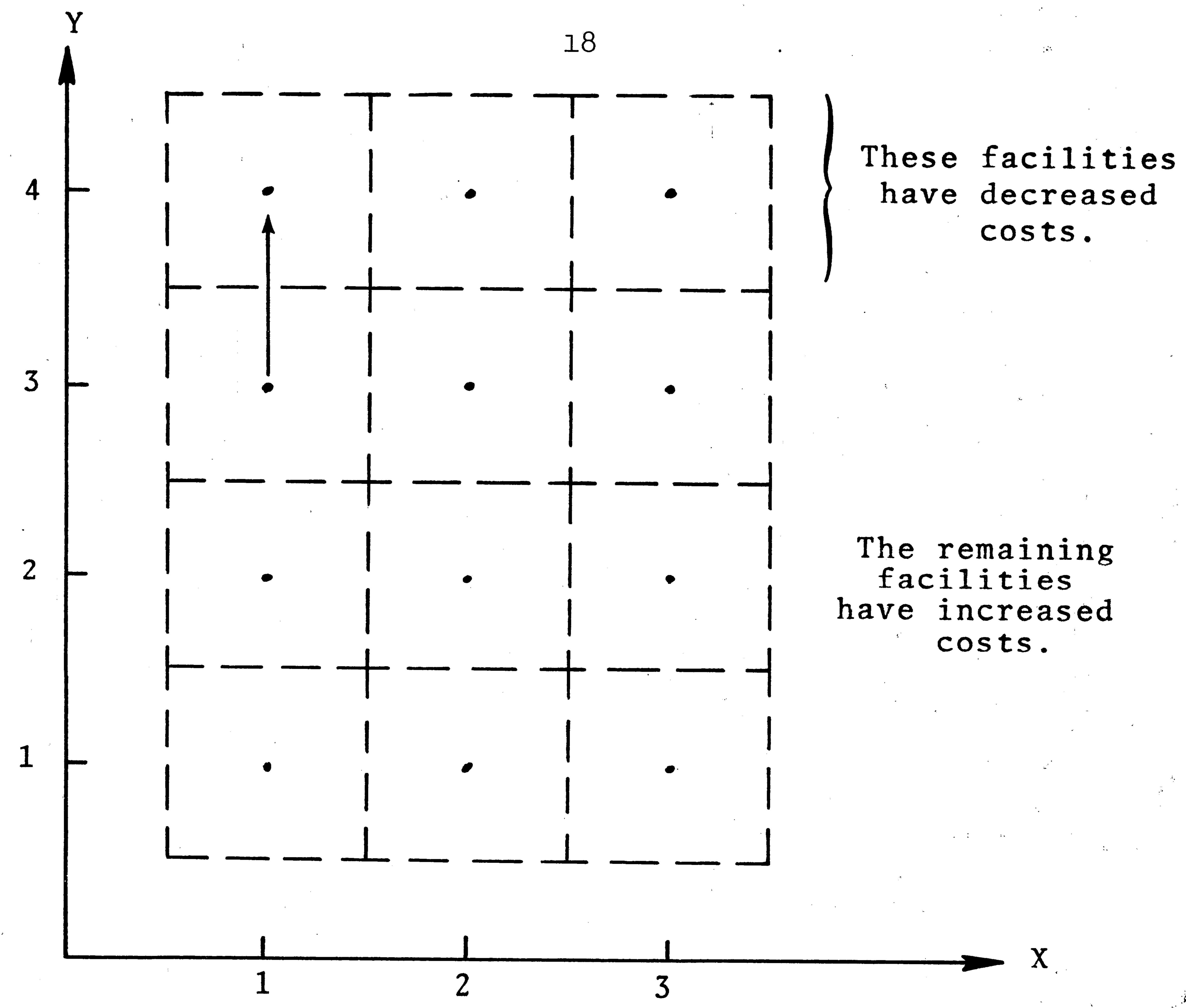


FIGURE IV - The Effects of Moving a Facility Upwards

The objective is to find the particular assignment of facilities which minimizes

$$S = 1/2 \sum_{x_1=1}^{x_n} \sum_{y_1=1}^{y_n} \sum_{x_2=1}^{x_n} \sum_{y_2=1}^{y_n} F(x_1, y_1; x_2, y_2)(|x_2-x_1| + |y_2-y_1|)$$

the total distance that materials must travel.

Every facility is evaluated for the effects of changing its location by only one position and the results are tabulated in a

"Move Desirability Table." This complete table is then searched for the interchange of two facilities that will generate the greatest reduction in S. This exchange is made and a subsequent table is computed. The process continues until there is no possible improvement to be made by a one-step transfer.

Nugent[41] found this algorithm to be fast and effective. While the results are near-sighted and therefore sub-optimal, they are comparitively good. Like most improvement techniques, the quality of the final solution will depend on the quality of the initial solution.

Armour and Buffa[3] introduced Craft (Computerized Relative Allocation of Facilities) in 1963. This is also an improvement technique that evaluates the pairwise exchange of facilities for decreases in the total distance travelled.

The basis of the method lies in the fact that a consideration of facilities by pairs results in a reduction of the number of combinations. Consider for example a problem that has ten facilities. There are 3,628,800 permutations if the facilities are taken singly:

$$N! = 10! = 3,628,800$$

when the same problem is considered by pairs there are only:

$$\binom{N}{R} = \frac{N!}{R!(N-R)!} = \frac{10!}{2!(8!)} = 45$$

permutations. This is a truly magnificent reduction in the number of combinations to be evaluated.

During each iteration the algorithm evaluates the effects of each of the $\binom{N}{R}$ possible pairwise combinations and selects the most effective exchange. This exchange is made and the process is repeated until no profitable exchange exists.

Craft has been found to be useful due to its ability to accommodate heterogeneous departments, mandatory assignments of facilities to certain locations, and problems with up to forty departments. As with other improvement techniques, the quality of the solution is dependent upon the quality of the initial solution. In the experiment by Nugent[41] it was found that a random sample of initial solutions could produce a solution superior to the one from straight Craft.

In 1966 Hillier and Connors[20] presented a modified version of Hillier's[19] 1963 algorithm. This modification consists mainly of relaxing the restriction against multiple step moves. The new algorithm computes an "N-Step Move Desirability Table" and makes the most effective interchange of facilities shown on this table. Nugent found that this algorithm runs longer on a computer than the original method, yet produces only slightly better results.

Suganami[44] has developed an algorithm for the traveling salesman problem that is essentially an application of Craft. His method is applicable to the facilities -location assignment problem, but it has not been evaluated with problems of a significant size.

A recent addition to the heuristic methods is Corelap (Computerized Relationship Layout Planning).[33] This is a construction method with flexible boundary constraints, a feature that makes the method useful for designing a new plant around the layout. Each facility is ranked according to the desirability of locating it adjacent to the others. This ranking is tabulated as the "total closeness rating," which is used as a priority for assignments. The method has not been sufficiently evaluated by comparison due to its recent introduction.

Optimum Seeking Procedures

The interested reader is referred to Moore[36] for an excellent review of the general mathematical techniques that have been used on the problem. Moore lists six general approaches:

1. The Level Curve Concept
2. The Assignment Model
3. The Piece-wise Linear Model
4. The Integer Programming Model
5. The Quadratic Integer Model
6. The Quadratic Programming Model

Commonly accepted models have not been developed in each of these areas. While a valid formulation may be possible, current technology can produce no efficient algorithm. An example of such a case is the quadratic formulation.

Quadratic Formulation

One of the first mathematical approaches to the facilities-location assignment problem was proposed by Koopmans and Beckman[26] in 1957. Their contribution suggested a quadratic formulation:

$$\text{Minimize } \sum_i \sum_g \sum_j \sum_h X_{ij} X_{gh} f_{qi} d_{hj}$$

Subject to the restrictions:

$$\sum_{i=1}^{N} X_{ij} = 1 \qquad j = 1, 2 \ldots N$$

$$\sum_{j=1}^{N} X_{ij} = 1 \qquad i = 1, 2 \ldots N$$

where: $X_{ij} = \binom{1}{0}$ if facility j $\binom{\text{is}}{\text{is not}}$ in location i

$X_{gh} = \binom{1}{0}$ if facility h $\binom{\text{is}}{\text{is not}}$ in location g

$X_{ij}, X_{gh} = 0 \text{ or } 1 \ (i, j, g, h = 1, 2 \ldots N)$

$f_{qi}, d_{hj} \geq 0$

Neither an algorithm or a practical method of solution were proposed.

Since then, Gilmore[16] and Lawler[31] have developed optimum seeking algorithms based on a similar formulation, but using the branch and bound technique for solution. These algorithms appear to be computationally feasible for only small problems.

The Assignment Model

In 1961 Moore[38] proposed that the Hungarian assignment method as described by Kuhn[27] and Munkres[39] be used for locating additional facilities in an existing layout. The model assumes that costs of assigning each facility to each of the candidate locations can be readily determined. These costs are compiled as an effectiveness matrix and the Hungarian method is employed to determine the optimum set of assignments.

While the Hungarian assignment method is extremely efficient, the procedure becomes unmanageable when the new facilities are highly interrelated. This increase in dependence makes it difficult to evaluate the costs of assignments and trends towards complete enumeration. However, for problems with facilities that are loosely related, the method is very efficient.

The Level Curve Concept

Moore,[36] McHose,[35] and Vergin and Rogers[47] discuss optimal seeking techniques that are based on the level curve concept. These methods attempt to find a location for an additional facility that will minimize the aggregate distance from the existing machines.

$$\text{Minimize } C = \sum_{i=1}^{n} f_i d_i \qquad \text{(where n is the number of existing machines)}$$

If the mode of travel is via a system of rectangular aisles and the new machine is located at point (x,y), then the distance to existing machine i at point (x_i, y_i) is:

$$d_i = |x-x_i| + |y-y_i|$$

So that:

$$C = \sum_{i=1}^{n} f_i (|x-x_i| + |y-y_i|)$$

or:

$$C = \sum_{i=1}^{n} f_i (|x-x_i|) + \sum_{i=1}^{n} f_i (|y-y_i|)$$

Since the median values of the sets of x_i and y_i will minimize the variations about themselves, it can be shown that the two summations will be minimized at the median values in the x and y directions.

The problem becomes more complex when the mode of movement is on a direct line between facilities. Then movement in the x direction is not independent of movement in the y direction. Since:

$$d_i = [(x-x_i)^2 + (y-y_i)^2]^{1/2}$$

then:

$$C = \sum_{i=1}^{n} f_i[(x-x_i)^2 + (y-y_i)^2]^{1/2}$$

To find a minimum the expression must be partially differentiated with respect to both x and y and then equated to zero.

$$\partial C/\partial x = \sum_{i=1}^{n} f_i(x-x_i)/[(x-x_i)^2 + (y-y_i)^2]^{1/2} = 0$$

$$\partial C/\partial y = \sum_{i=1}^{n} f_i(y-y_i)/[(x-x_i)^2 + (y-y_i)^2]^{1/2} = 0$$

Unfortunately, these equations cannot be solved simultaneously, but most be solved by iteration. Vergin and Rogers[47] use a Fortran program to iterate on these equations, while Moore[36] used the computer to plot curves of equal costs. The methods are inefficient for problems of large size, but are excellent for adding a small number of facilities to an existing layout.

Branch and Bound Methods

One of the newest approaches towards an optimal solution of the facilities-location assignment problem is the branch and bound technique. It is a semi-enumerative approach, since the solution set is searched in a manner that tends to reduce the combinations that are evaluated. The Nugent[41] study evaluated branch and bound algorithms developed by Gilmore,[17] Lawler,[31] and Gavett and Plyter[15] and concluded that "no computationally feasible optimal-producing procedure exists at present."

However, Turner[46] has shown subsequently that the branch and bound techniques are worthy of further study. Turner's algorithm, which is a modification of the method of Gavett and Plyter,[15] is used as a subroutine in the model developed for this thesis. Consequently, a discussion of his algorithm will be included as Chapter IV.

The algorithms of Gilmore[17] and Lawler[31] are so similar, although they were developed independently, that Nugent was led to call them the "Gilmore-Lawler algorithms." A cost effectiveness matrix is constructed using the pairwise inter-facility flows and the pairwise inter-location distances. As shown on page 18, pair-wise consideration of n objects produces $\frac{n!}{2!(n-2)!}$ or $\frac{n(n-1)}{2}$ combinations. Therefore, there are $\frac{n^2-n}{2}$ pairs of facilities to be assigned to $\frac{n^2-n}{2}$ pairs of locations and an assignment matrix of dimension $(\frac{n^2-n}{2}, \frac{n^2-n}{2})$ results.

As suggested by Conway and Maxwell,[8] the pairwise flows are ranked in monotonic non-increasing order and the distances in monotonic non-decreasing order. The dot product of these two vectors produces a matrix whose main diagonal holds the optimal assignment. However, as also noted by Conway and Maxwell[8] this assignment is not necessarily feasible.

"The n elements which describe an assignment will be such that there will be one in each column and one in each row of the matrix. The sum of these n elements will be the value of the assignment. Unfortunately, the converse is not true: any n elements with one in each column and one in each row do not necessarily correspond to an assignment."

Gilmore[16,17] suggests the formation of "partial permutations" by assigning a pair of facilities (he used modules) to a pair of locations. The assignment matrix is then reduced in dimension and the process is repeated to determine a lower cost bound. The search is obviously inefficient in design and requires much computing time.

Land[28] proposes the construction of the cost matrix in the same manner, but employs a Hungarian-type approach to the selection of assignments. The method appears to be interesting, but it has not been evaluated by any comparative tests.

IV - THE BRANCH AND BOUND ALGORITHM

Little, Murty, Sweeney and Karel[34] are credited with the formal development of the branch and bound technique in 1963. While various features of the technique have been used by previous authors, Little et al formulated the complete procedure as a method for solving the traveling salesman problem. They also noted the general applicability of the technique to many other combinatorial problems. This generality has proven to be true, as shown by the survey of branch and bound methods made by Lawler and Wood[32].

Branch and bound is essentially an intelligently structured search of the space of all feasible solutions. This space is repeatedly partitioned into successively smaller subsets which are then evaluated for the bounding cost of solutions within each subset. Those subsets having a cost bound within the limits of the currently best solution are kept active for further consideration. Those subsets having a cost bound exceeding that of the current solution are excluded from further consideration and are considered as being implicitly enumerated. The entire procedure continues until a feasible solution is found which has a cost no greater than any bound.

It is convenient to represent the process by a decision tree similar to that shown in Figure V. The points of decision at the junction of the branches are commonly called nodes.

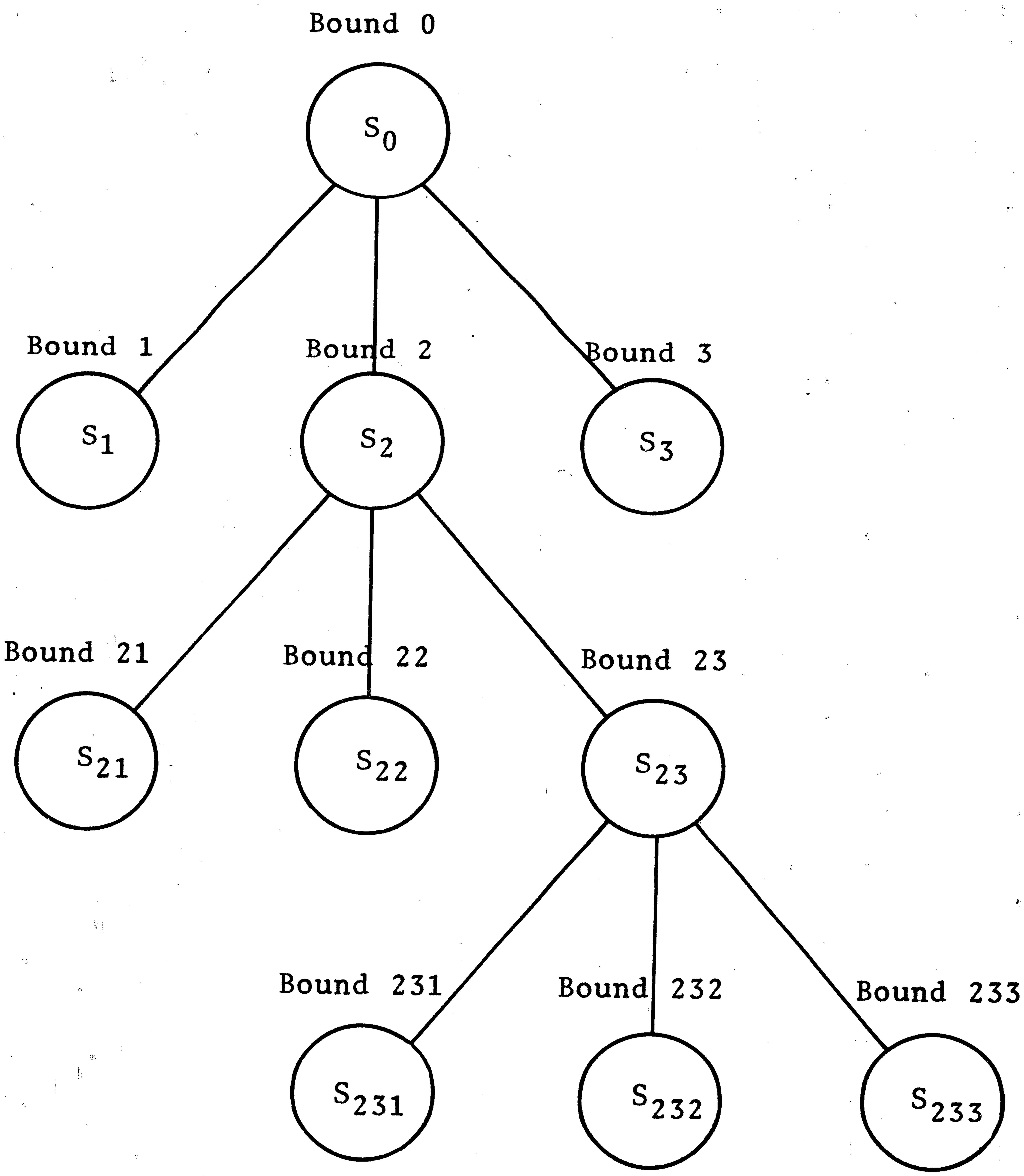


FIGURE V

A Branch and Bound Tree

The Gavett and Plyter Algorithm

Agin[1] has defined a branch and bound algorithm by these procedural requirements:

> "A set of rules for (1) branching from nodes to new nodes, (2) determining lower bounds for the new nodes, (3) choosing an intermediate node from which to branch next, (4) recognizing when a node contains only infeasible or non-optimal solutions and (5) recognizing when a final node contains an optimal solution."

Gavett and Plyter[15] developed such an algorithm to obtain optimal solutions for the facilities assignment problem. Turner[46] has subsequently modified their algorithm in order to reduce the time required to reach a solution. A description of the initial algorithm will be followed by Turner's modifications.

The first step in the algorithm is the development of a ranked cost matrix. The facilities to be assigned and the locations to be filled are both considered in pairs, so that the elements of the cost matrix represent the costs of assigning a pair of facilities to a pair of locations. The flow between each pair of facilities is totaled, so that the flow from facility A to facility B and that from B to A are represented by one value. The flow values are ranked into a monotonically non-decreasing vector (F_R) and the pairwise inter-location distances are ranked into a monotonically non-increasing vector (D_R). The multiplication of these two vectors produce the ranked cost matrix (C_R).

$$F_R = [f_1, f_2, f_3, \ldots f_k]$$

where $f_1 \leq f_2 \leq f_3 \quad \ldots \leq f_k$

and $k = \frac{N(N-1)}{2}$

$$D_R = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \cdot \\ \cdot \\ \cdot \\ d_k \end{bmatrix} \qquad \text{where } d_1 \geq d_2 \geq d_3 \quad \ldots \geq d_k$$

$$C_R = D_R F_R = \begin{bmatrix} d_1f_1 & d_1f_2 & d_1f_3 & \cdots & d_1f_k \\ d_2f_1 & d_2f_2 & d_2f_3 & \cdots & d_2f_k \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ d_kf_1 & d_kf_2 & d_kf_3 & \cdots & d_kf_k \end{bmatrix}$$

As Conway and Maxwell[8] noted, the minimal set of assignments can be found along the main diagonal of such a matrix:

$$Z(X)^* = d_1f_1 + d_2f_2 + d_3f_3 \ldots + d_kf_k$$

$$= \sum_{i=1}^{k} d_i f_i$$

$$= \sum_{i=1}^{k} C_R\,(i,i)$$

However, this solution disregards the feasibility constraints. Since the distance elements (d_i) represent the distances between pairs of locations and the flow elements (f_i) represent the flows between pairs of facilities, the elements of C_R are not independent. It is highly probable that the diagonal solution will attempt to assign more than one facility to a location, which is infeasible. However, the value of this set of assignments is useful as the lower cost bound for the complete set of solutions.

If the diagonal solution proves to be infeasible, the next step in the algorithm is a reduction of the matrix to facilitate the selection of alternate assignments. A reduced matrix is defined as having non-negative elements and at least one zero in each row and each column. The values of elements of the reduced matrix represent the differential costs of not selecting the zero elements.

To construct this matrix from the ranked cost matrix each diagonal element is subtracted from all the elements in its respective column and row. This will normally result in negative elements below the diagonal, which may be corrected by subtracting the minimum element of that row from all elements of the row. The process is continued until the two requirements of non-negativity and the zero elements have been met. Due to the method of constructing the initial ranked cost matrix, the reduced matrix will at least have zeroes along the diagonal. The sum of the constants used in reducing the matrix will equal the diagonal sum, $Z(X)^*$. The reduced cost matrix from Gavett and Plyter's example is shown in Figure VI.

Facility Pair		2-4	1-4	2-3	3-4	1-3	1-2
Flow Vector (F_R)		4	13	15	23	25	28
Location Pair	Distance Vector (D_R)						
A-C	7	0	9	11	27	37	55
A-B	6	0	0	0	8	16	31
B-D	6	0	0	0	8	16	31
B-C	5	11	2	0	0	6	18
A-D	2	68	32	24	0	0	3
C-D	1	89	44	34	2	0	0

Figure VI

Reduced Cost Matrix from Gavett and Plyter's Example

The next procedure is that of selecting an optimal set of assignments from the reduced cost matrix using branch and bound. Obviously, the initial node in the decision tree is the set of diagonal assignments with a lower bound equal to Z(X)*. The next node must be chosen by a branching rule and then evaluated for the cost bound.

The rule used by Gavett and Plyter considers the cost elements nearest the main diagonal and selects the maximum value among these. The cost bound for the assignment at this node is easily computed by adding the element just selected to the bound of the previous node. The alternate node selection is evaluated by selecting the remaining feasible assignments and evaluating their bounds in the same manner. Naturally the lower of the two is chosen for further evaluation, if the bound is within the current active bound.

Turner's Algorithm

In their study of algorithms for the facilities assignment problem, Nugent, Vollmann and Ruml[41] concluded that the Gavett and Plyter algorithm was not computationally feasible for problems of a reasonable size. They wrote:

> "In research for this paper total enumeration computer runs made on the smaller problems were accomplished on the GE-265 in times comparable to those reported by Gavett and Plyter using their branch-and-bound formulation on the IBM-7074, a machine about 20 per cent faster than the GE-265. The Gavett-Plyter procedure is clearly computationally infeasible for only but the smallest problems."

Turner[46] subsequently worked towards increasing the efficiency of the algorithm by improving the methods of making assignments and of selecting subsequent nodes.

He improved the efficiency of making assignments by considering single facilities rather than pairs. With the pairwise assignments there are K^K possibilities to be considered when feasibility is disregarded, as it is at this point. By considering single facilities the possibilities are reduced to N^N, but a problem is created in determining the cost of the assignment. Turner overcame this with a recursive relationship for the successively assigned facility costs.

The method of selecting subsequent nodes in the decision tree was improved by adopting a more forward looking branching method.

The unassigned facilities are ranked by the relative amounts of flow associated with them and the busiest are assigned first. This tends to place those facilities with the higher activities near the center of the floorspace and those with lower activity near the perimeter. A similar short cut will be used in the partitioning techniques described in the next chapter.

Turner reports that the combination of these modifications greatly reduces the computer time spent on each node and results in a significant increase in efficiency. A flow diagram of his algorithm is included at the end of this thesis as Appendix A.

V - THE PARTITIONING MODEL

A feasible method of decomposing the problem into smaller subsets is needed, and a possible procedure is to group the facilities by a hierarchy of mutual flows. Those facilities with a high degree of inter-flow will be considered for grouping first, and those having lower degrees of flow will be subsequently assigned. The distance between locations will be temporarily ignored, since no effort will be made to associate facilities with locations during the decomposition. Once all of the facilities have been allocated to groups, the groups can be arranged in an optimal manner. Finally the facilities can be optimally arranged within their respective groups.

A search of the literature [9,12,23,30,45,49] produced few answers for the questions concerning grouping technique. Most of the work done with hierarchical grouping concerns statistical sampling and the minimization of intra-group variance. While none of these techniques directly applied to the question at hand, they suggested two methods for the grouping of facilities. These methods are (1) grouping with cumulative external flow and (2) grouping with non-cumulative flow.

The Cumulative Flow Method

When individual facilities have been grouped, they are then regarded as a single entity and the treatment of the flows between external facilities and the group must now be re-evaluated. The cumulative flow method considers the group to be a sum of its parts, and therefore sums the flows from each external facility to all members of the group.

Consider the following example of six facilities that are to be clustered into two groups of three each.

To / From	1	2	3	4	5	6
1	0	5	2	4	1	0
2	5	0	3	0	5	3
3	2	3	0	0	0	0
4	4	0	0	0	1	6
5	1	5	0	1	0	10*
6	0	3	0	6	10*	0

The flow between facilities 5 and 6 has the highest value, so these facilities are candidates for a group. They are combined into one entity, facility 5-6, and the dimension of the matrix is reduced by one.

To / From	1	2	3	4	5-6
1	0	5	2	4	1
2	5	0	3	0	8*
3	2	3	0	0	0
4	4	0	0	0	7
5-6	1	8*	0	7	0

Note that the flow from 4 to 5-6 is now seven units, which is the sum of flows 4 to 5 (1 unit) and 4 to 6 (6 units). The highest flow now is that from 2 to 5-6, so facility 2 is a candidate for the group containing 5-6. There is room available in the group, so this combination is made.

To / From	1	3	4	2-5-6
1	0	2	4	6
3	2	0	0	3
4	4	0	0	7*
2-5-6	6	3	7*	0

The highest flow remaining is between 4 and 2-5-6. Since group 2-5-6 is complete, the flow from 4 to 2-5-6 is blocked from the matrix and the process continues.

To / From	1	3	4	2-5-6
1	0	2	4	6*
3	2	0	0	3
4	4	0	0	X
2-5-6	6*	3	X	0

Again the highest flow exists between an external facility and the full group, so this flow is also blocked from the matrix. Eventually the two groups of facilities 1-3-4 and 2-5-6 will be completed.

There is never an advantage in the unseating of a previously assigned member of the group in favor of a subsequent facility, as the assignments are made in the order of descending hierarchy. No subsequent facility can be more favorable for the group than one already assigned. A flow diagram for this procedure is given in Appendix B.

The Non-Cumulative Flow Method

This method of clustering considers the attraction between individual facilities rather than that between a facility and the group. When facilities are grouped, the external flows are established by the highest single flow to any member of the group. Therefore, subsequent assignments to the group are made on the basis of the degree of flow between two facilities. Obviously, there is no difference between this method and the cumulative one if only two facilities are to be assigned to a group.

Consider the flow matrix of the previous example for the formation of two groups.

To / From	1	2	3	4	5	6
1	0	5	2	4	1	0
2	5	0	3	0	5	3
3	2	3	0	0	0	0
4	4	0	0	0	1	6
5	1	5	0	1	0	10*
6	0	3	0	6	10*	0

The flow between facilities 5 and 6 is the highest, so they are grouped into one entity.

To / From	1	2	3	4	5-6
1	0	5	2	4	1
2	5	0	3	0	5
3	2	3	0	0	0
4	4	0	0	0	6*
5-6	1	5	0	6*	0

Note that the flow from 4 to 5-6 is now six units rather than seven as before. This value was selected as the higher of flows 4 to 5 (1 unit) and 4 to 6 (6 units). Again the matrix is reduced in dimension by the combining of facilities. The next selection is the flow from 4 to 5-6, so facility 4 is added to the group.

To / From	1	2	3	4-5-6
1	0	5*	2	4
2	5*	0	3	5*
3	2	3	0	0
4-5-6	4	5*	0	0

The next level of flow is five units and there are two flows at this value. Facility 2 cannot be added to the cluster of 4-5-6, so the flow between 2 and 4-5-6 is blocked from the matrix. This leaves the highest flow between 1 and 2, which are joined as a new group.

To / From	1-2	3	4-5-6
1-2	0	3	5*
3	3	0	0
4-5-6	5*	0	0

Now the highest flow exists between groups 1-2 and 4-5-6, which must obviously be blocked from the matrix. This leaves only the flow from 3 to 1-2, so the final group is complete.

Note that this method has formed groups of facilities 1-2-3 and 4-5-6, where the cumulative method formed groups of 1-3-4 and 2-5-6. The methods will usually produce different results, except for the case of two facilities per group. A comparison of the effects of the two methods on the final results of a problem will be made in the next chapter. A flow diagram of the method is given in Appendix B.

The Complete Model

The groups resulting from the decomposition can be optimally arranged by the branch and bound technique. This technique can be applied to these groups at two levels. One is a "macro" level to find optimal relative positions for the groups, while the second is a "micro" level to arrange the facilities within their respective groups.

The number of groups to be formed and the relative configuration of their positions must be submitted by the user. The model assumes that all traffic between groups originates and terminates at the centroid of the area, which is similar to the assumption made concerning the traffic between the individual facilities.

The model proceeds to partition the facilities into groups, to locate the groups, and to assign the facilities within each of the groups. The resultant layout is the tested for possible improvement by rotating each group about its centroid and also by creating and evaluating the mirror image of each group. The justification of this improvement technique lies in the fact that the relative positions chosen by the branch and bound algorithm are not altered, while the distance between communicating facilities that are located in different groups may be reduced. This possibility for improvement results from the assumption that macro travel originates and terminates at the centroids of the groups' areas.

After the layout has been improved, if possible, and then evaluated, the final assignments and the total cost are written out on the computer's printer. A flow diagram of the model is given in Appendix C.

VI - EVALUATION OF CLUSTERING FACTORS

Five factors associated with clustering were selected for the evaluation of their effects on the solution to a facilities-location assignment problem. These five factors appear to have the most relevance to the partitioning of a job shop layout problem:

Clustering Method - This factor pertains to the two methods that were discussed in the previous chapter. These methods are (1) the clustering of facilities with cumulative external flows and (2) with non-cumulative external flows.

Divisibility Factor - This concerns the number of facilities to be assigned and whether this number is a prime or a non-prime integer. Non-prime numbers of facilities (such as 5 or 13) are not conducive to partitioning. This factor creates considerable difficulty in the selection of partitions and their geometric configurations.

Number of Facilities - This factor concerns the size of the complete problem. It should be of considerable interest to know how the problem size effects the quality of the solution.

Number of Partitions - This pertains to the number of sub-sets that are created from the complete problem by the partitions.

Configuration of Partitions - This factor defines the geometric configuration of the arrangement of the partitioned locations. The shape of these locations should have a significant effect on the final result. This factor is defined as being either linearly or centrally configured. Figure VII shows an example of how eight locations can be partitioned into either linear or central configurations that accept two groups of four facilities each.

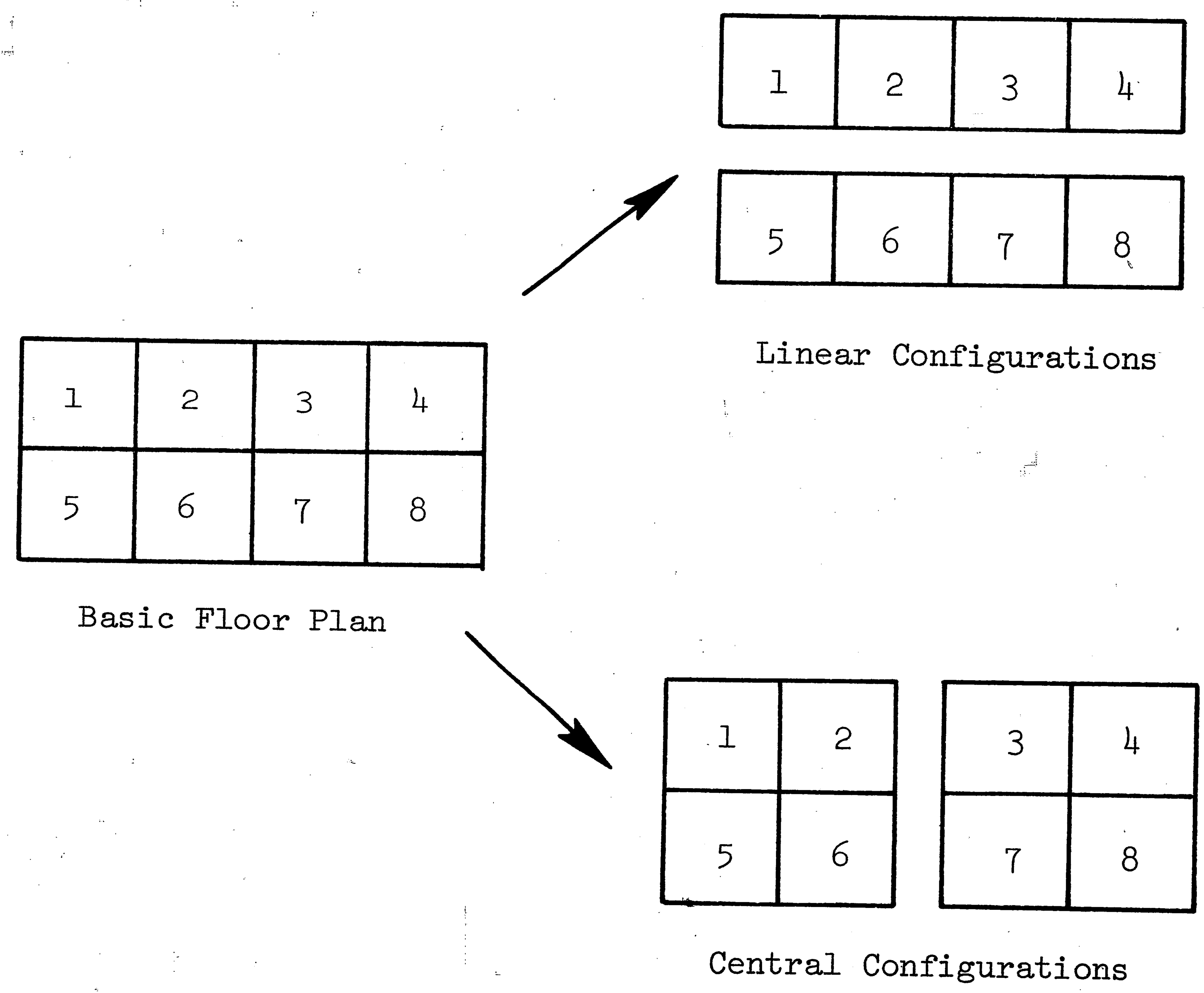


FIGURE VII

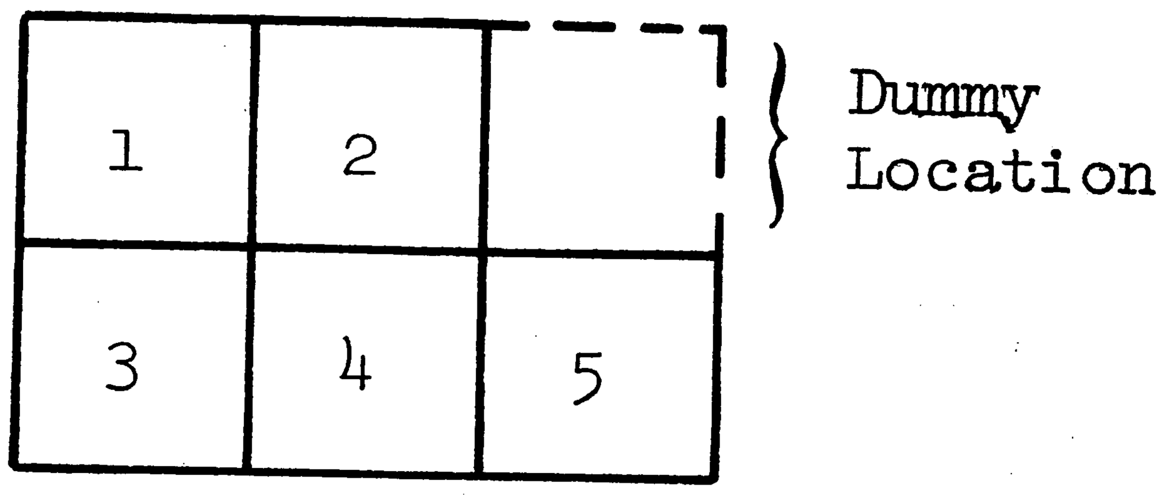
An Example of Linear and Central Configurations

The Experimental Design

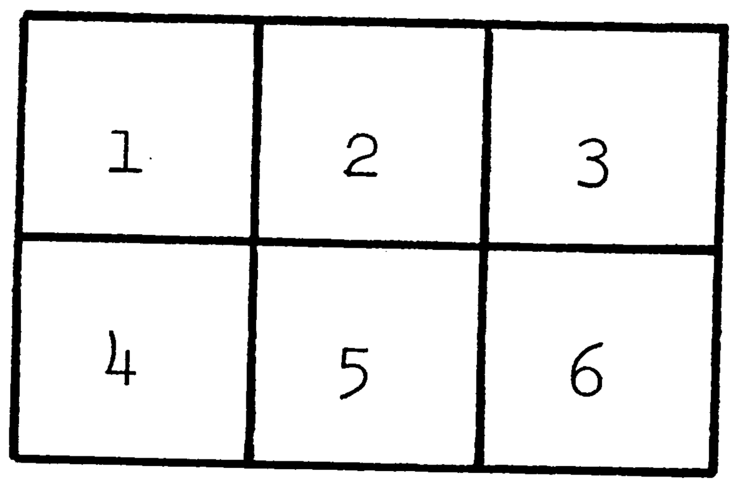
It is obvious that the partitioning of facilities assignment problems is unnecessary for problems having less than nine or ten facilities, since these small problems can be solved in reasonable times by the branch and bound method. For the purposes of this study, it is desirable to compare the results of a partitioned problem to the true optimum solution, and optimum solutions are unknown for the large problems. Therefore, it was decided to create small problems and to solve them twice, once by the branch and bound technique and once with the partitioning model.

Four basic floor plans of five, six, seven, and eight locations were chosen to evaluate the model. These floor plans (Figure VIII) are similar to those used by Nugent, et al[41] in their study. The floor plans for five and seven locations, non-prime divisibility factors, were augmented with dummy locations to facilitate multiple partitioning. This was required due to the assumption of equal sized partitions that is made by the model.

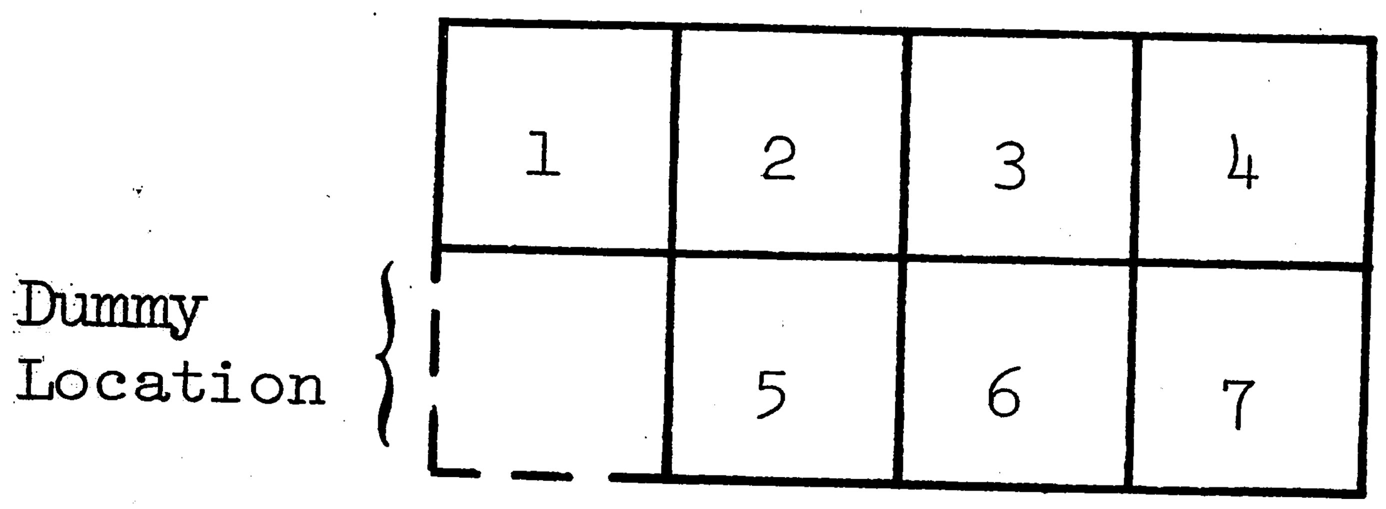
A normalization of the results is needed to provide a common basis for the comparison of problems of different size. This normalization was chosen as the ratio of the partitioned solutions to the true optimum solutions:



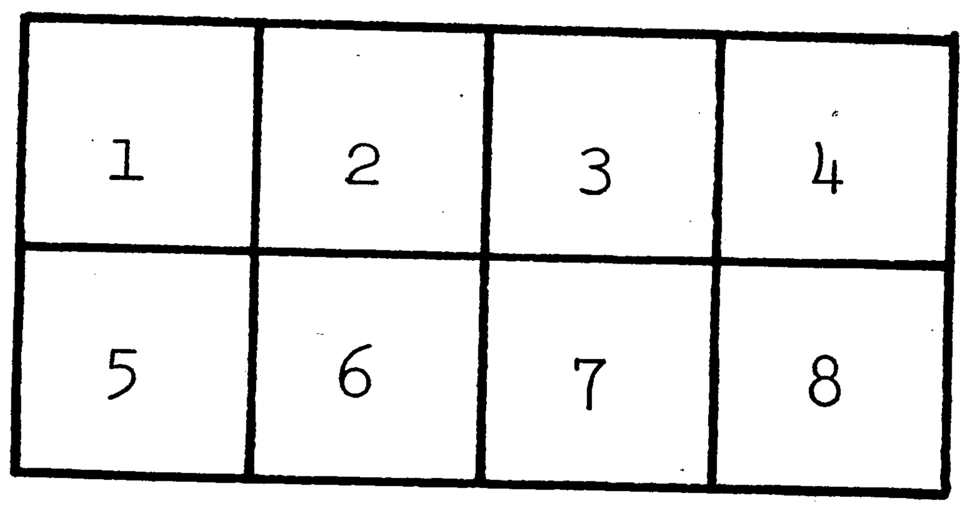
Five Location
Floor Plan



Six Location
Floor Plan



Seven Location
Floor Plan



Eight Location
Floor Plan

FIGURE VIII

Floor Plans Used in Experiment

$$\text{Inefficiency Ratio} = \frac{\text{Partitioned Solution}}{\text{True Optimum Solution}}$$

A ratio of unity implies that the two solutions of the same problem were equal. Ratios higher than one indicate the measure of additional cost incurred due to the partitioning of the problem.

One hundred problems were created for each of the four floor plans by generating the inter-facility flows. These flows were created with the IBM-1130 computer and a random number generator. The values of these flows were uniformly distributed between 0 and 10 units to approximate the job shop layout problem. These problems were first solved by the branch-and-bound algorithm to produce the optimal solutions, and then they were solved by partitioning to yield the inefficiency ratios.

The one hundred inefficiency ratios that resulted from each experimental session were averaged to give a mean inefficiency ratio. According to the Central Limit Theorem, these mean ratios can be assumed to be normally distributed due to the large sample size. This normality allows an evaluation of the results by analysis of variance techniques.

Analysis of Variance

The presence of several factors at different levels of treatment suggested the possibility of using the Yates technique for the analysis of variance. This technique is a simplified mechanical method for determining the total effects of the factors involved and their interactions in a two level factorial experiment. Therefore, the experiment was structured to evaluate the five factors at two levels each, which made it a 2^5 factorial experiment.

The application of the Yates technique requires alphabetical aliases for each of the factors involved:

A - Partitioning method

B - Divisibility factor

C - Number of facilities

D - Number of partitions

E - Configuration of partitions

Table I lists the factors, their levels of treatment, and corresponding aliases that were used in the experiment. Note that some factors of the treatment, such as algorithm used, are not actually used at different levels, but that the definition is arbitrary. This is a useful attribute of the Yates technique.

Table II gives the mean inefficiency ratios that were observed for each of the combinations of the treatments. This experimental data was then processed by the Yates technique and the results are recorded on Table III.

To provide an estimation of the variance, the four and five factor interactions are assumed to be zero. This permits the total of the squares for these six effects (indicated in Table III by an *) to be used for the residual sum of squares. Dividing this total by six (the number of 4 and 5 level interactions) gives a pooled estimate of the variance. This estimate is then used to calculate the "Mean Square Ratios" for the other sources (Table IV). The significant ratios are denoted by asterisks and are keyed to the critical F ratios

Factor Levels with Corresponding Alaises

Divisibility Factor	B_0 Not a Prime Number				B_1 Prime Number			
Number of Facilities	C_0 5		C_1 7		C_0 6		C_1 8	
Number of Partitions	D_0 2	D_1 3	D_0 2	D_1 4	D_0 2	D_1 3	D_0 2	D_1 4
Configuration	E_0 Linear		E_1 Central		E_0 Linear		E_1 Central	
Algorithm	A_0 Cumulative		A_1 Non-Cumulative		A_0 Cumulative		A_1 Non-Cumulative	

TABLE I

EXPERIMENTAL FACTOR LEVELS WITH CORRESPONDING ALIASES

		A_0				A_1			
		B_0		B_1		B_0		B_1	
		C_0	C_1	C_0	C_1	C_0	C_1	C_0	C_1
E_0	D_0	1.0466	1.0863	1.0546	1.0691	1.0502	1.0771	1.0528	1.0586
	D_1	1.0205	1.0762	1.0363	1.0554	1.0205	1.0762	1.0363	1.0554
E_1	D_0	1.0618	1.0549	1.0423	1.0517	1.0554	1.0607	1.0453	1.0611
	D_1	1.0254	1.0472	1.0381	1.0410	1.0254	1.0472	1.0381	1.0410

TABLE II

LAYOUT SOLUTION INEFFICIENCY RATIOS WITH CORRESPONDING FACTOR LEVELS

Treatment Combination	Inefficiency Ratio	Average Effect	Sum of Squares
(1)	1.0466	--	--
A	1.0502	-2.125×10^{-4}	0.003×10^{-4}
B	1.0546	-32.374×10^{-4}	0.838×10^{-4}
AB	1.0528	2.249×10^{-4}	0.004×10^{-4}
C	1.0836	191.750×10^{-4}	29.414×10^{-4}
AC	1.0771	-0.125×10^{-4}	0.000
BC	1.0691	-79.875×10^{-4}	5.104×10^{-4}
ABC	1.0586	-2.749×10^{-4}	0.006×10^{-4}
D	1.0205	-153.499×10^{-4}	18.849×10^{-4}
AD	1.0205	2.125×10^{-4}	0.003×10^{-4}
BD	1.0363	36.125×10^{-4}	1.044×10^{-4}
ABD	1.0363	-2.249×10^{-4}	0.004×10^{-4}
CD	1.0762	56.999×10^{-4}	2.599×10^{-4}
ACD	1.0762	0.125×10^{-4}	0.000
BCD	1.0554	-58.874×10^{-4}	2.773×10^{-4}
ABCD	1.0554	2.749×10^{-4}	0.006×10^{-4}*
E	1.0618	-82.999×10^{-4}	5.511×10^{-4}
AE	1.0554	16.874×10^{-4}	0.227×10^{-4}
BE	1.0423	8.124×10^{-4}	0.052×10^{-4}
ABE	1.0453	14.000×10^{-4}	0.156×10^{-4}
CE	1.0549	-100.499×10^{-4}	8.080×10^{-4}
ACE	1.0607	23.375×10^{-4}	0.437×10^{-4}
BCE	1.0517	66.124×10^{-4}	3.498×10^{-4}
ABCE	1.0611	-4.499×10^{-4}	0.016×10^{-4}
DE	1.0254	-8.749×10^{-4}	0.061×10^{-4}
ADE	1.0254	-16.874×10^{-4}	0.227×10^{-4}
BDE	1.0381	-20.624×10^{-4}	0.340×10^{-4}
ABDE	1.0381	-14.000×10^{-4}	0.156×10^{-4}*
CDE	1.0472	-24.749×10^{-4}	0.490×10^{-4}
ACDE	1.0472	-23.375×10^{-4}	0.437×10^{-4}*
BCDE	1.0410	-21.875×10^{-4}	0.380×10^{-4}*
ABCDE	1.0410	4.499×10^{-4}	0.016×10^{-4}*
TOTAL	33.6060		
GRAND MEAN	1.0501		

* - 4 and 5 level interactions used for the residual sum of squares

TABLE III

RESULTS OF THE EXPERIMENT AND OF APPLYING THE YATES TECHNIQUE

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	MEAN SQUARE RATIO
A	0.003×10^{-4}	1	0.003×10^{-4}	0.021
B	0.838×10^{-4}	1	0.838×10^{-4}	4.955*
C	29.414×10^{-4}	1	29.414×10^{-4}	173.848***
D	18.849×10^{-4}	1	18.849×10^{-4}	111.408***
E	5.511×10^{-4}	1	5.511×10^{-4}	32.572***
AB	0.004×10^{-4}	1	--	--
AC	0.000	1	--	--
AD	0.003×10^{-4}	1	--	--
AE	0.227×10^{-4}	1	0.227×10^{-4}	1.346
BC	5.104×10^{-4}	1	5.104×10^{-4}	30.166***
BD	1.044×10^{-4}	1	1.044×10^{-4}	6.170**
BE	0.052×10^{-4}	1	--	--
CD	2.599×10^{-4}	1	2.599×10^{-4}	15.362***
CE	8.080×10^{-4}	1	8.080×10^{-4}	47.756***
DE	0.061×10^{-4}	1	--	--
ABC	0.006×10^{-4}	1	--	--
ABD	0.004×10^{-4}	1	--	--
ABE	0.156×10^{-4}	1	--	--
ACD	0.000	1	--	--
ACE	0.437×10^{-4}	1	0.437×10^{-4}	2.583
ADE	0.227×10^{-4}	1	0.227×10^{-4}	1.346
BCD	2.773×10^{-4}	1	2.773×10^{-3}	16.389***
BCE	3.498×10^{-4}	1	3.498×10^{-3}	20.674***
BDE	0.340×10^{-4}	1	0.340×10^{-3}	2.011
CDE	0.490×10^{-4}	1	0.490×10^{-4}	2.896
4 AND 5 FACTOR INTERACTIONS	1.031×10^{-4}	6	0.172×10^{-4}	
TOTALS	80.751×10^{-4}	31		

Significance limits for F: $F_{1,6,0.90} = 3.78$*

$F_{1,6,0.95} = 5.99$**

$F_{1,6,0.99} = 13.75$***

TABLE IV

ANALYSIS OF VARIANCE FOR SOLUTION INEFFICIENCIES listed at the bottom of Table IV.

From these results, it can be seen that there were ten sources of significant effects within the range of treatments that were evaluated. Of the five main factors, only the partitioning method was insignificant. The divisibility factor (B) is mildly significant at a level of ten per cent, while the factors of problem size (C), number of partitions (D), and the configuration (E) have strongly significant effects on the quality of the final solution.

The divisibility factor and the number of partitions have a significant interaction (BD) at the five per cent level. This is understandable, since it has been noted that a non-prime number of locations is difficult to partition effectively.

The number of facilities interacts strongly with the divisibility factor (source BC), the number of partitions (CD), and the configuration (CE). Apparently the quality of the partitioned solution will decrease significantly as the problem size increases, if the decomposition is not limited to a minimum and the partitions are not centrally configured. This conclusion is supported by the significance of the three factor interactions, BCD and BCE, at the one per cent level.

These adverse effects on the quality of the solution can best be explained by comparing partitioning to the ranked cost matrix of the Gavett-Plyter algorithm. While the set of assignments falling on the diagonal of this matrix yields an absolute minimum cost, some of these assignments must be traded off to achieve feasibility. The hierarchical clustering methods overlook such global considerations in the pursuit of a local objective, namely the proximity of highly related facilities.

Once these clusters have been formed, the trade-offs for overall efficiency are precluded. During this experiment an attempt was made to include a subroutine that would make such an exchange of facilities between groups, if the cost could be reduced. However, this routine was an assignment algorithm in itself and caused a prohibitive usage of time, so it was deleted from the model.

It can be concluded that the best partitioning strategy is to create the minimum number of clusters that will maintain computational feasibility within the groups. The number of facilities in a group should never exceed ten and smaller groups are naturally faster to work with.

As was expected, the configuration of the partitioned floor plan has a very significant effect on the quality of the solution. Since the facilities are clustered by the degree of mutual flow, they should be closely located, and configuring the locations in a linear arrangement tends to separate the facilities rather than to join them. A partitioned assignment problem should use configurations that are compact and are uniformly distributed about their centroids.

Comparison to Heuristic Techniques

Now that the proper choices of partitioning strategy are known, the partitioning technique can be compared to the better heuristic methods of solving the same problem. The results of the Nugent[41] study can be used as the vehicle for such a comparison.

The problems with twelve and thirty facilities were chosen from

Problem Size	Partitioned Solutions Cumulative	Partitioned Solutions Non-Cumulative	Hillier's 1963 Algorithm	Hillier-Connors Algorithm	CRAFT
12	315	313	317	310	296
30	3359	3459	3267	3206	3189

TABLE V

Cost of Final Layout Versus Problem Size

Problem Size	Partitioning	Hillier's 1963 Algorithm	Hillier-Connors Algorithm	CRAFT
12	37 sec.*	55 sec.	19 sec.	70 sec.
30	204 sec.*	398 sec.	285 sec.	3150 sec.

* Times normalized to the GE-265 computer

TABLE VI

Computer Solution Times Versus Problem Size those used in that study. The twelve facility problem was partitioned into two groups of six, and the thirty facility problem was decomposed into five groups of six. These partitions allowed the maximum sizes of groups with centrally configured locations. The solutions to these problems with the corresponding results as reported by Nugent are given in Table V.

The partitioned solutions are definitely competitive for the twelve facility problem, but the increased amount of partitioning that is necessary to accomodate the large problem causes the solution to run from 2.8% to 8.4% higher than the heuristic solutions. It must be concluded that the heuristic techniques are as good as, or better than, the partitioning technique, when quality is the main criterion.

The computing times for these problems are given in Table VI. Note that the times for the partitioned solutions have been normalized to approximate the running times on a General Electric-265 computer, which was the machine used in the Nugent study. The partitioning method is not only competitive in the lower range; it is clearly superior in the upper ranges of problem size. This advantage is relatively unimportant for problems in the twenty to thirty facility range, but it could mean that partitioning is the only feasible technique for truly large problems.

A problem with one hundred facilities, ten groups of ten facilities each, is the approximate upper limit for computational feasibility with the partitioned branch and bound technique. Such a problem should require approximately twenty hours of computer time.

VII - CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

The branch and bound technique can produce optimum solutions for facilities assignment problems that fall within the current limit of computational feasibility, which is approximately ten facilities. This limit can be extended to problems with one hundred facilities by partitioning them with hierarchical clustering techniques similar to those used in this thesis. A problem of this magnitude can be expected to require from fifteen to twenty hours on the IBM-360/50 computer.

In the lower range of ten to twenty facilities the partitioning technique equals the speeds of the heuristic algorithms of Hillier, Hillier and Connors, and Buffa. In the range above twenty facilities the speed of the partitioning technique is superior to these same algorithms.

This increase in speed and the resultant increase in problem capacity is accompanied by a decrease in the quality of the final solution. In the lower range of problem size, ten or twelve facilities, the partitioned solution is competitive with the heuristic solutions and there is no significant difference between them. At the highest point tested, a thirty facility problem, the heuristic solutions average from 2.8 per cent to 8.4 per cent better than the partitioned solution. This superiority of the heuristic algorithms can be expected to increase with problems of even larger size. However, the heuristic methods will rapidly become computationally infeasible while the partitioning technique remains usable.

There are two factors concerning partitioning strategy that the analyst should consider when solving a problem by the partitioned branch and bound technique. First, the problem should be decomposed into the minimum number of subsets that will maintain computational feasibility. Secondly, the number of partitions should be chosen to facilitate the configuration of centrally arranged locations to receive these facilities. The configuration of these locations should be compactly and uniformly arranged about the centroids of the groups in order to minimize the cost of the final layout.

It was found that a problem with a non-prime number of facilities is difficult to partition effectively, and must be imbedded in a larger problem that has a prime number. This is a minor difficulty and should present no serious problems in actual layout practice.

Recommendations for Further Study

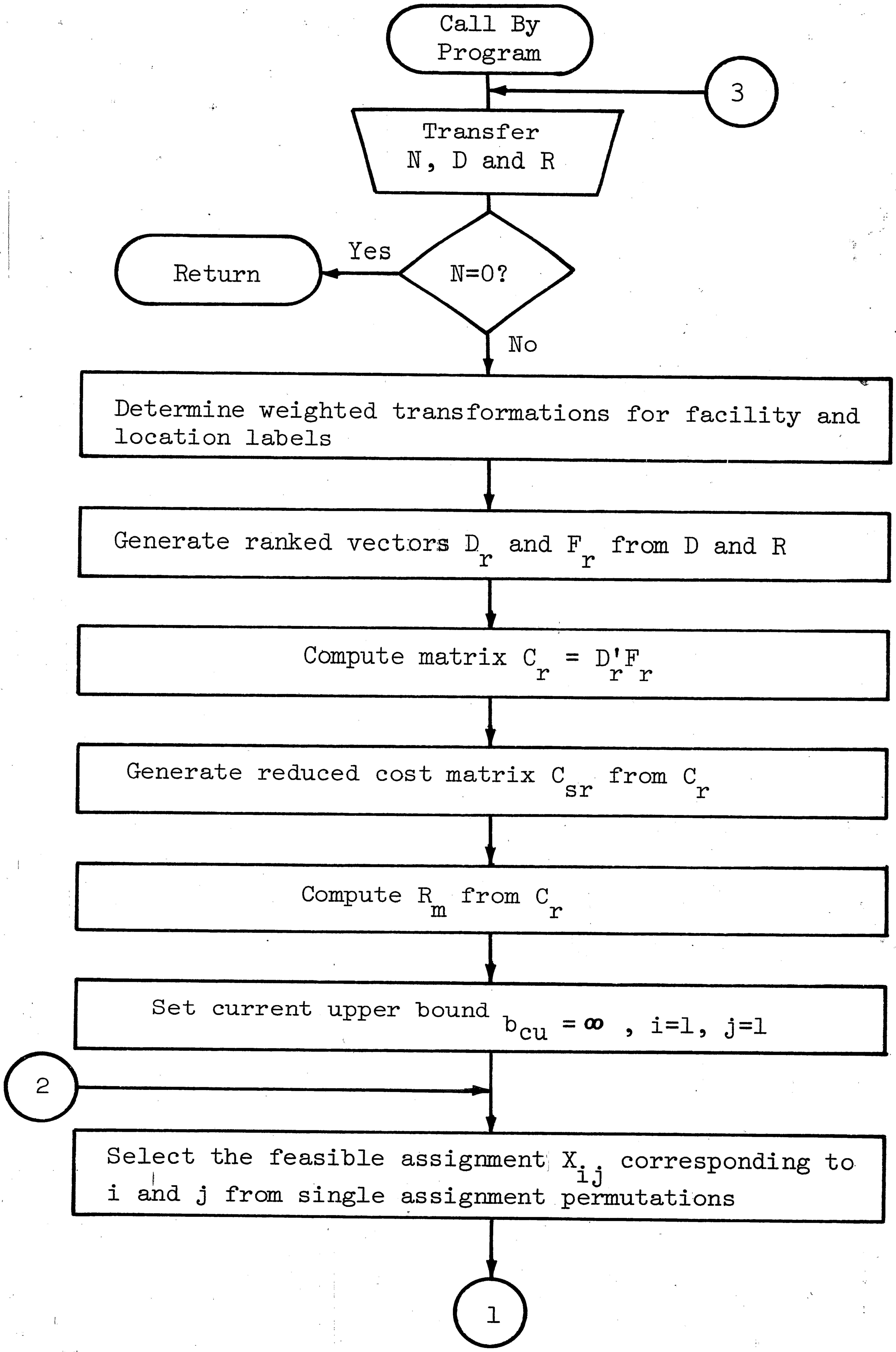
Since this study evaluated the factors effecting the solutions achieved by decomposing facilities assignment problems, the density of the inter-facility flow matrix was purposely overlooked. This factor is an attribute of the problem that is not controllable by the analyst, so it was considered to be outside the scope of the study. The worse possible case was assumed with approximately ninety percent of the facilities communicating with each other. However, it seems intuitively appealing to assume that the quality of the final solution will improve as the density of the flow matrix decreases. Research in this area may prove or disprove this hypothesis, and may determine a relationship between partitioning policy and the percentage of zeros in the flow matrix.

A second area for further study stems from the limiting factor of divisibility that was found in this study. Since it was found that certain odd numbers of facilities are excessively costly to partition, the concept of unequal partition sizes seems very attractive. It should be rather easy to modify the model used in this thesis to provide such a feature.

A third area of similar research is related to, and could be a combination of, the first two recommendations. This suggestion concerns modifying the model to permit the clusters to continue growing in size as long as there is a high degree of flow between the group and other facilities. Of course an upper limit on the cluster size would have to be specified to remain within the time limitations of the branch-and-bound algorithm.

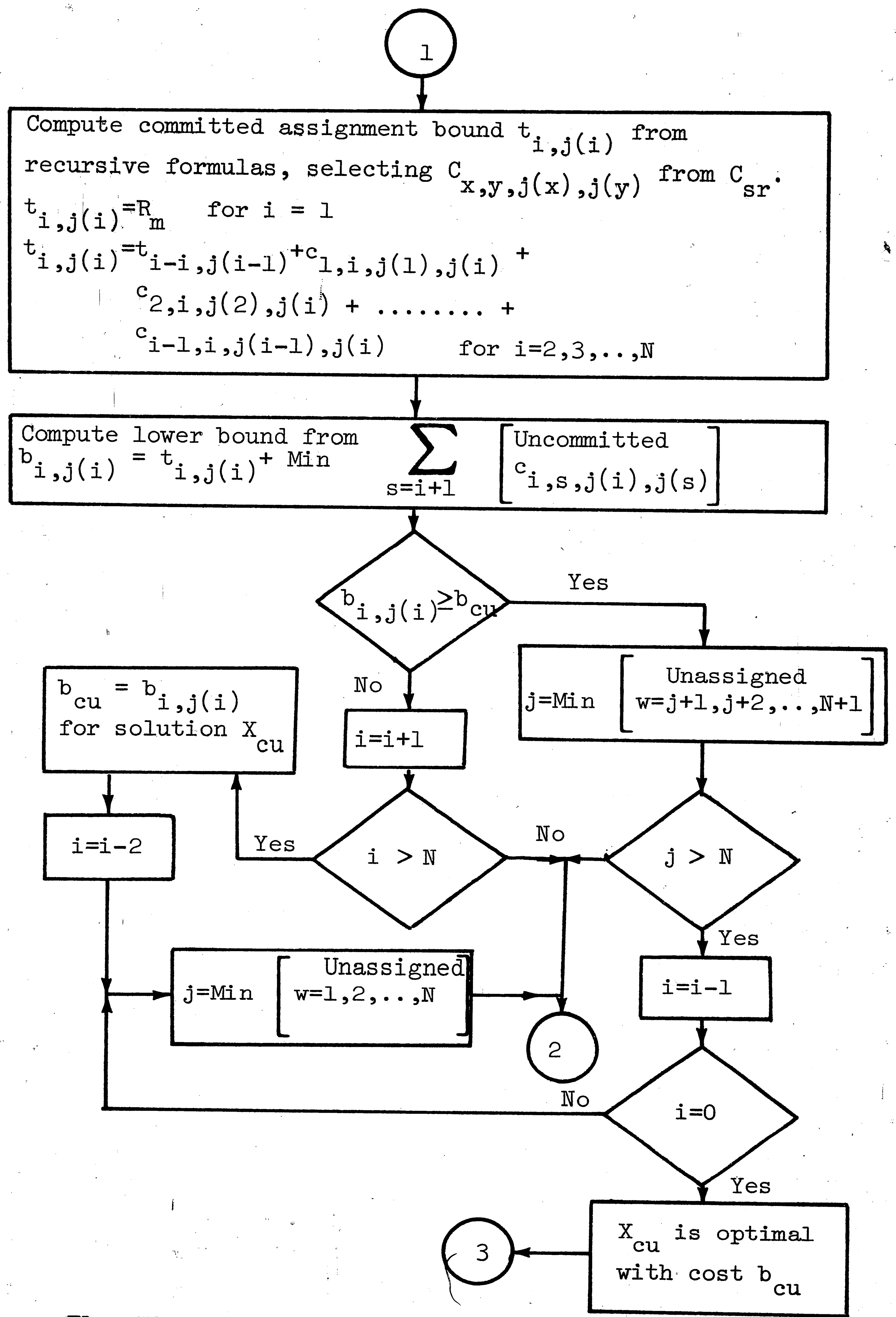
APPENDIX A

Flow Diagram for Turner's Branch and Bound Algorithm



Flow Diagram for Turner's Branch and Bound Algorithm

Part 1

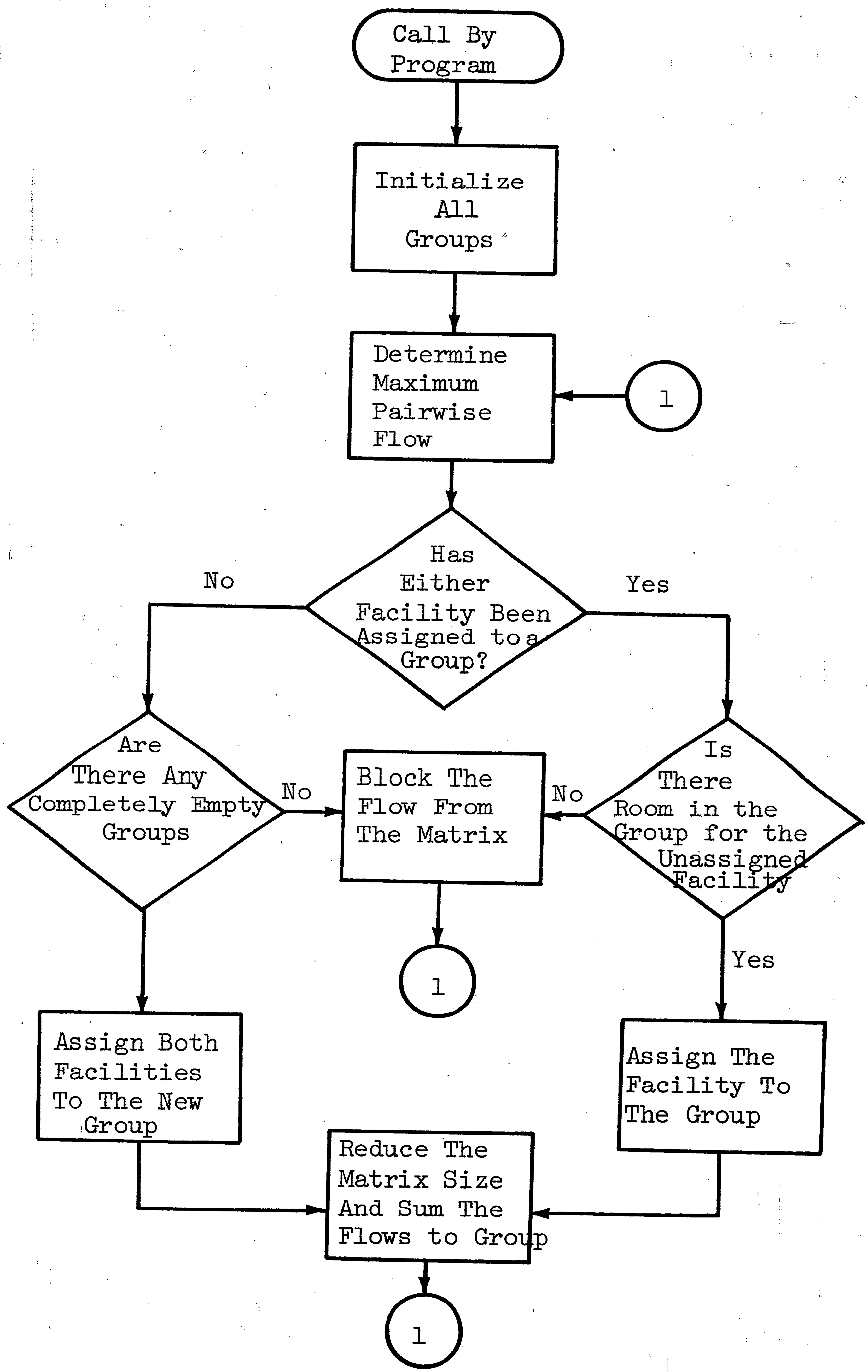


Flow Diagram for Turner's Branch and Bound Algorithm

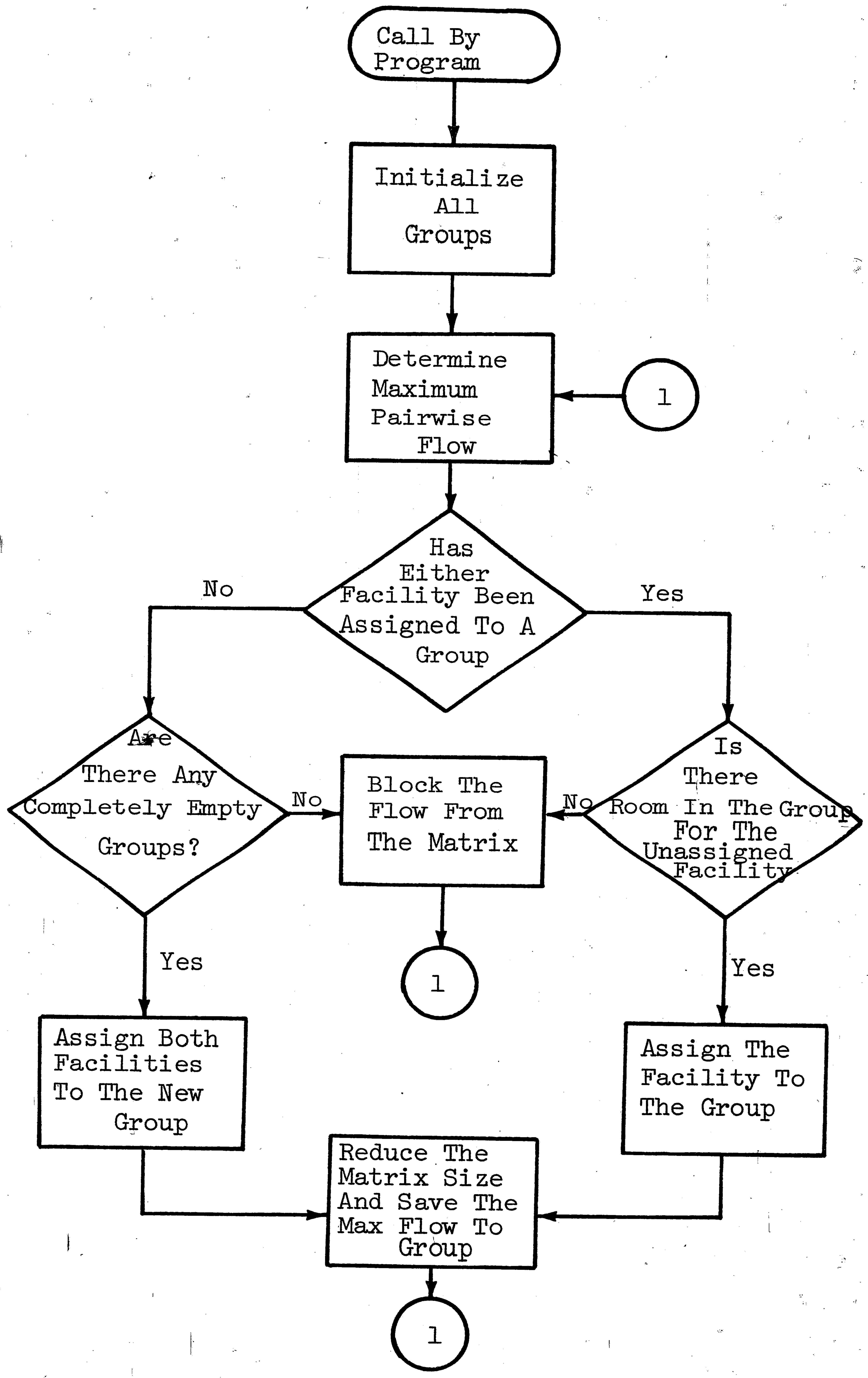
Part 2

APPENDIX B

Flow Diagrams for the Grouping Subroutines



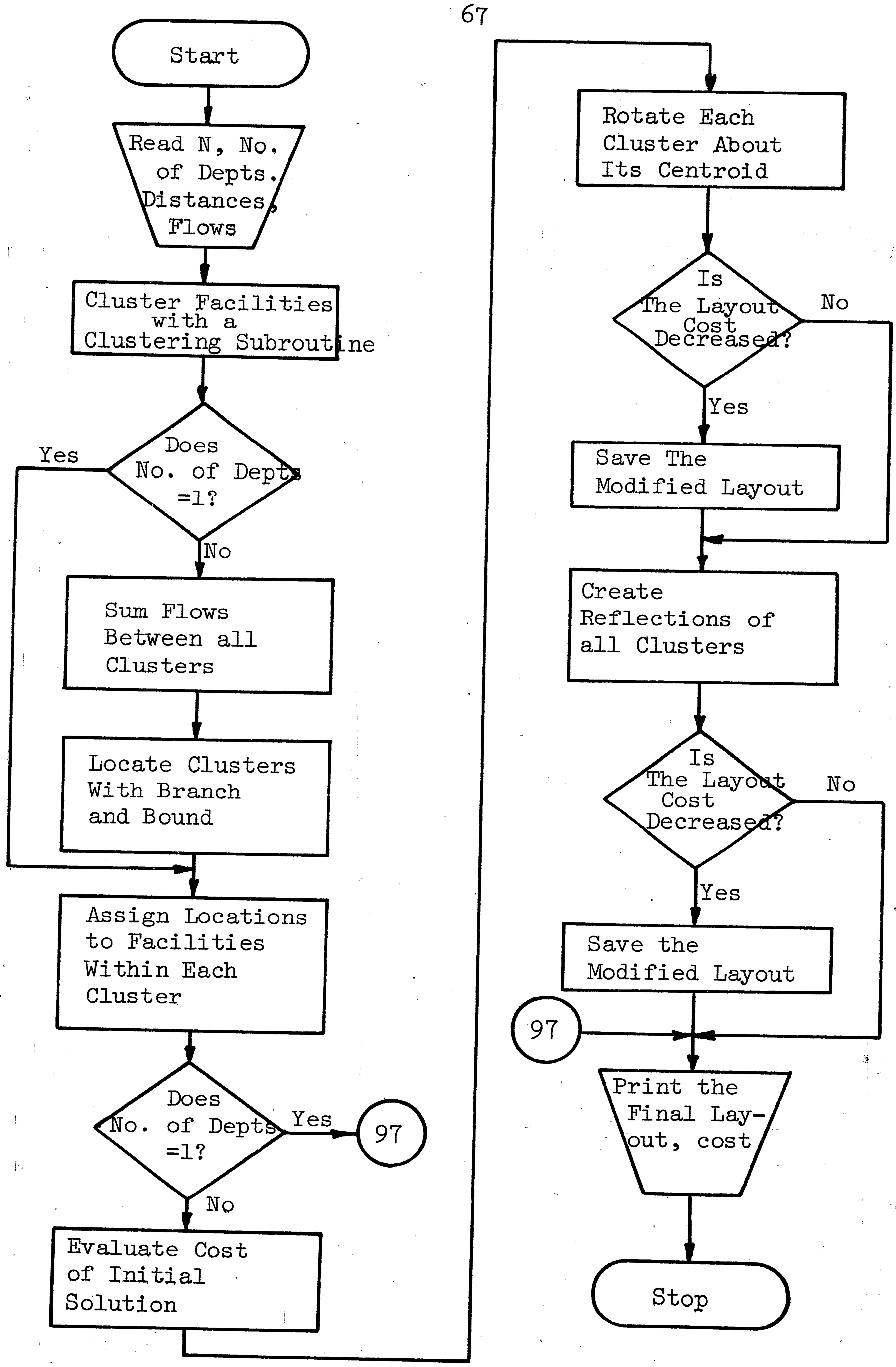
Flow Diagram For The Cumulative Flow Method Of Grouping



Flow Diagram For The Non-Cumulative Flow Method Of Grouping

APPENDIX C

Flow Diagram for the Mainline
Partitioning Program



Flow Diagram for the Partitioning Mainline Program

BIBLIOGRAPHY

1. Agin, Norman, "Optimum Seeking with Branch and Bound," Management Science, Vol. 13, No. 4, December, 1966, pp. 176-185.

2. Apple, James M., Plant Layout and Materials Handling, Ronald Press, New York, 1963.

3. Armour, G. C. and E. S. Buffa, "A Heuristic Algorithm and Simulation Approach to Relative Location of Facilities," Management Science, Vol. 9, No. 2, January, 1963, pp. 294-309.

4. Balinski, M. L., "Fixed Cost Transportation Problems," Naval Research Logistics Quarterly, Vol. 8, No. 1, March, 1961, pp. 41-54.

5. Buffa, E. S., Production-Inventory Systems, Planning and Control, Richard D. Irwin, Inc., Homewood, Illinois, 1968.

6. Buffa, E. S., "Sequence Analysis for Functional Layouts," The Journal of Industrial Engineering, Vol. 6, No. 2, March-April, 1955, pp. 12, 13, 25.

7. Buffa, E. S., G. C. Armour, and T. E. Vollmann, "Allocating Facilities with Craft," Harvard Business Review, Vol. 42, March-April, 1964, pp. 136-158.

8. Conway, R. W. and W. L. Maxwell, "A Note on the Assignment of Facility Location," The Journal of Industrial Engineering, Vol. 12, No. 1, January-February, 1961, pp. 34-36.

9. Cox, D. R., "Note on Grouping," Journal of the American Statistical Association, Vol. 52, 1957, pp. 543-547.

10. Dantzig, G. B., "On the Significance of Solving Linear Programming Problems with Some Integer Variables," Econometrica, Vol. 28, No. 1, January, 1960, pp. 30-44.

11. Driebeek, N. J., "An Algorithm for the Solution of Mixed Integer Programming Problems," Management Science, Vol. 12, No. 7, March, 1966, pp. 576-587.

12. Fisher, Walter D., "On Grouping for Maximum Homogeneity," Journal of the American Statistical Association, Vol. 53, 1958, pp. 789-798.

13. Fleisher, R. E., On the Solution of the General Single-Stage Location-Allocation Problem by a Branch and Bound Algorithm, Unpublished Master's Degree Thesis, Lehigh University, Bethlehem, Pennsylvania, 1967.

BIBLIOGRAPHY (Cont.)

14. French, C. E., G. W. Isaacs, and R. M. Peart, "Optimizing Systems Having Components with Discontinuous Cost Functions," Operations Research Quarterly, Vol. 9, No. 4, July, 1961, pp. 468-478.

15. Gavett, J. W., and N. V. Plyter, "The Optimal Assignment of Facilities to Locations by Branch and Bound," Operations Research, March-April, 1966, pp. 210-232.

16. Gilmore, P. C., "A Solution to the Module Placement Problem," International Business Machines Research Report RC-430, April 26, 1961.

17. Gilmore, P. C., "Optimal and Suboptimal Algorithms for the Quadratic Assignment Problem," The Journal of the Society for Industrial and Applied Mathematics, Vol. 10, No. 2, June, 1962, pp. 305-313.

18. Harris, R. D., and R. K. Smith, "A Cost Effectiveness Approach to Facilities Layout," The Journal of Industrial Engineering, Vol. 19, No. 6, June, 1968.

19. Hillier, F. S., "Quantitative Tools for Plant Layout Analysis," The Journal of Industrial Engineering, Vol. 14, No. 1, January-February, 1963, pp. 33-40.

20. Hillier, F. S., and M. M. Connors, "Quadratic Assignment Problem Algorithms and the Location of Indivisible Facilities," Management Science, Vol. 13, No. 1, September, 1966, pp. 42-57.

21. Hillier, F. S. and G. J. Lieberman, Introduction to Operations Research, Holden-Day, Inc., San Francisco, 1968.

22. Ireson, W. G., Factory Planning and Plant Layout, Prentice-Hall, Inc., 1952.

23. Johnson, S. C., "Hierarchial Clustering Schemes," Psychometrika, Vol. 32, No. 3, 1967, pp. 241-254.

24. Klein, M. and S. H. Milberg, "The Application of Linear Programming to Materials Handling," Modern Materials Handling, Vol. 10, No. 2, February, 1955, pp. 80-84.

25. Kolesar, P. J., "A Branch and Bound Algorithm for the Knapsack Problem," Management Science, Vol. 13, No. 6, May, 1967, pp. 723-735.

BIBLIOGRAPHY (Cont.)

26. Koopmans, T. C., and M. Beckmann, "Assignment Problems and the Location of Economic Activities," Econometrica, Vol. 25, No. 1, January, 1957, pp. 53-76.

27. Kuhn, H. W., "The Hungarian Method for the Assignment Problem," Naval Research Logistics Quarterly, Vol. 2, No. 2, March-June, 1955, pp. 83-97.

28. Land, A. H., "A Problem of Assignment with Interrelated Costs," Operations Research Quarterly, Vol. 14, 1963, pp. 185-199.

29. Land, A. H., and A. G. Doig, "An Automatic Method of Solving Discrete Programming Problems," Econometrica, Vol. 28, No. 3, July, 1960, pp. 497-520.

30. Land, F. M., "Notes on a Problem of Multiple Classifications," Psychometrika, Vol. 17, No. 3, September, 1952.

31. Lawler, E. L., "The Quadratic Assignment Problem," Management Science, Vol. 9, No. 4, July, 1963, pp. 586-599.

32. Lawler, E. L., and D. E. Wood, "Branch and Bound Methods: A Survey," Operations Research, Vol. 14, No. 4, July-August, 1966, pp. 699-719.

33. Lee, R. C., and J. M. Moore, "Corelap-Computerized Relationship Layout Planning," The Journal of Industrial Engineering, Vol. 18, No. 3, March, 1967, pp. 195-200.

34. Little, John D. C., K. G. Murty, D. W. Sweeney, and Caroline Karel, "An Algorithm for the Traveling Salesman Problem," Operations Research, Vol. 11, 1963, pp. 972-989.

35. McHose, A. H., "A Quadratic Formulation of the Activity Location Problem," The Journal of Industrial Engineering, Vol. 12, No. 5, September-October, 1961, pp. 334-337.

36. Moore, J. M., Mathematical Models for Optimizing Plant Layout, Unpublished Doctorial Dissertation, Stanford University, Stanford, California, 1965.

37. Moore, J. M., Plant Layout and Design, The Macmillan Co., New York, 1962.

BIBLIOGRAPHY (Cont.)

38. Moore, J. M., "Optimal Locations for Multiple Machines," The Journal of Industrial Engineering, Vol. 12, No. 5, October, 1961, pp. 307-313.

39. Munkres, James, "Algorithms for the Assignment and Transportation Problems," The Journal of the Society for Industrial and Applied Mathematics, Vol. 5, No. 1, March, 1957, pp. 32-38.

40. Muther, Richard, Systematic Layout Planning, Industrial Education Institute, Boston, Massachusetts, 1961.

41. Nugent, C. E., T. E. Vollmann, and John Ruml, "An Experimental Comparison of Techniques for the Assignment of Facilities to Locations," Operations Research, January-February, 1968, pp. 150-173.

42. Reis, I. L., and G. E. Anderson, "Relative Importance Factors in Layout Analysis," The Journal of Industrial Engineering, Vol. 11, No. 4, July-August, 1960, pp. 312-316.

43. Shetty, C. M., "A Solution to the Transportation Problem with Non-Linear Costs," Operations Research, Vol. 7, September-October, 1959, pp. 571-580.

44. Suganami, Saburo, "Schedule of Rooming Assignments," The Journal of the Operations Research Society of Japan, Vol. 8, No. 2, February, 1966, pp. 91-107.

45. Thorndike, Robert L., "Who Belongs in the Family," Psychometrika, Vol. 18, 1953, pp. 267-276.

46. Turner, L. C., An Investigation of Branch and Bound Techniques for Solving the Facilities-Location Assignment Problem, Unpublished Master's Degree Thesis, Lehigh University, Bethlehem, Pennsylvania, 1969.

47. Vergin, R. C., and J. D. Rogers, "An Algorithm and Computational Procedure for Locating Economic Facilities," Management Science, Vol. 13, No. 6, February, 1967, pp. 240-254.

48. Vollmann, T. E., and E. S. Buffa, "The Facilities Layout Problem in Perspective," Management Science, Vol. 12, No. 10, June, 1966, pp. 450-468.

49. Ward, J. H., "Hierarchial Grouping to Optimize an Objective Function," Journal of the American Statistical Association, Vol. 58, 1963, pp. 236-244.

BIBLIOGRAPHY (Cont.)

50. Wimmert, R. J., "A Mathematical Method of Equipment Location," The Journal of Industrial Engineering, Vol. 9, No. 6, November-December, 1958, pp. 498-505.

VITA

PERSONAL HISTORY

Name:	Charles Henry Blackburn III
Birth Place:	Birmingham, Alabama
Parents:	Charles and Edna Blackburn
Wife:	Brenda Sue (Nelson) Blackburn
Children:	Susan Lea, Michael Sean, and Terri Lynn

EDUCATIONAL BACKGROUND

Auburn University Bachelor of Science in Electrical Engineering	1952-1957
Lehigh University Candidate for Master of Science in Industrial Engineering	1968-1970

HONORS

Alpha Pi Mu
Edcen Club
Eta Kappa Nu
Phi Eta Sigma
Pi Tau Pi Sigma
Tau Beta Pi

PROFESSIONAL EXPERIENCE

Schlumberger Well Surveying Corp. Laurel, Mississippi Shreveport, Louisiana Many, Louisiana Field Engineer	1957-1965
Western Electric Co., Inc. Shreveport, Louisiana Planning Engineer	1965-1968

VITA CONT.

PROFESSIONAL EXPERIENCE

Western Electric Co., Inc.
Princeton, New Jersey
Research Engineer 1968-1970